



Guia do Desenvolvedor

# AWS Lake Formation



# AWS Lake Formation: Guia do Desenvolvedor

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 clientes ou que deprecie ou desprestigie a Amazon. Todas as outras marcas comerciais que não são propriedade da Amazon pertencem aos respectivos proprietários, os quais podem ou não ser afiliados, estar conectados ou ser patrocinados pela Amazon.

---

# Table of Contents

O que é o AWS Lake Formation? .....	1
Características do Lake Formation .....	2
Ingestão e gerenciamento de dados .....	2
Gerenciamento de segurança .....	3
Compartilhamento de dados .....	4
Como funciona .....	5
Fluxo de trabalho de gerenciamento de permissões do Lake Formation .....	5
Permissões de metadados .....	7
Gerenciar o acesso às do armazenamento .....	10
Compartilhamento de dados entre contas no Lake Formation .....	12
Componentes do Lake Formation .....	13
Lake Formation .....	13
API e interface de linha de comando do Lake Formation .....	13
Outros produtos da AWS .....	13
Terminologia do Lake Formation .....	14
Data lake .....	14
Acesso aos dados .....	14
Modo de acesso híbrido .....	14
Blueprint .....	15
Fluxo de trabalho .....	15
Catálogo de dados .....	15
Dados subjacentes .....	16
Entidade principal .....	16
Administrador do data lake .....	16
Integrações de serviços da AWS com o Lake Formation .....	16
Recursos adicionais do Lake Formation .....	18
Blogs .....	18
Palestras técnicas e webinars .....	19
Arquitetura moderna .....	19
Recursos de data mesh .....	19
Guias práticas recomendadas .....	19
Introdução ao Lake Formation .....	20
Introdução .....	21
Concluir as tarefas iniciais de configuração do AWS .....	21

Inscrever-se para uma Conta da AWS .....	21
Criar um usuário administrativo .....	22
Conceder acesso programático .....	23
Configurar o AWS Lake Formation .....	24
Configurar recursos do Lake Formation usando o modelo AWS CloudFormation .....	25
Crie um administrador de data lake .....	26
Altere o modelo de permissão padrão ou use o modo de acesso híbrido .....	31
Atribuir permissões aos usuários do Lake Formation .....	32
Como configurar um local no Amazon S3 para o data lake .....	34
(Opcional) Configurações externas de filtragem de dados .....	35
(Opcional) Conceder acesso à chave de criptografia do Catálogo de Dados .....	36
(Opcional) Crie uma função do IAM para fluxos de trabalho .....	36
Atualizar as permissões de dados AWS Glue para o modelo do Lake Formation .....	38
Atualizar as permissões de dados para o modelo Lake Formation .....	38
Etapa 1: listar permissões existentes .....	40
Etapa 2: configurar permissões do Lake Formation .....	42
Etapa 3: dar permissões do IAM .....	42
Etapa 4: mude seu modelo de permissões do Lake Formation .....	43
Etapa 5: proteja os novos recursos do Catálogo de Dados .....	46
Etapa 6: fornecer aos usuários uma nova política do IAM .....	47
Etapa 7: limpar políticas do IAM existentes .....	48
Configurando endpoints da VPC da Amazon (AWS PrivateLink) .....	48
Considerações sobre endpoints da VPC do Lake Formation .....	49
Criando um endpoint da VPC de interface para o Lake Formation .....	49
Criando uma política de endpoint da VPC para o Lake Formation .....	50
Tutoriais .....	52
Criação de um data lake a partir de uma AWS CloudTrail fonte .....	53
Público-alvo .....	54
Pré-requisitos .....	55
Etapa 1: Criar um usuário de analista de dados .....	56
Etapa 2: adicionar permissões para ler AWS CloudTrail registros à função do fluxo de trabalho .....	57
Etapa 3: Criar um bucket do Amazon S3 para o data lake .....	57
Etapa 4: Registrar um caminho do Amazon S3 .....	58
Etapa 5: Conceder permissões de local de dados .....	58
Etapa 6: Criar um banco de dados no Catálogo de Dados .....	59



Etapa 7: Conceder permissões de dados .....	59
Etapa 8: usar um esquema para criar um fluxo de trabalho .....	61
Etapa 9: executar o fluxo de trabalho .....	62
Etapa 10: Conceder SELECT nas tabelas .....	63
Etapa 11: Consultar o data lake usando Amazon Athena .....	64
Criação de um data lake a partir de uma fonte JDBC .....	64
Público-alvo .....	65
Pré-requisitos .....	66
Etapa 1: criar um usuário de analista de dados .....	66
Etapa 2: criar uma conexão em AWS Glue .....	68
Etapa 3: criar um bucket do Amazon S3 para o data lake .....	68
Etapa 4: registrar um caminho do Amazon S3 .....	69
Etapa 5: conceder permissões de localização de dados .....	69
Etapa 6: criar um banco de dados no Catálogo de dados .....	70
Etapa 7: conceder permissões de dados .....	70
Etapa 8: usar um esquema para criar um fluxo de trabalho .....	71
Etapa 9: executar o fluxo de trabalho .....	72
Etapa 10: conceder SELECIONAR nas tabelas .....	73
Etapa 11: consultar o data lake usando Amazon Athena .....	74
Etapa 12: consulte os dados no data lake usando o Amazon Redshift Spectrum .....	74
Etapa 13: conceder ou revogar permissões do Lake Formation usando o Amazon Redshift Spectrum .....	79
Como configurar permissões para formatos de tabela aberta no Lake Formation .....	79
Público-alvo .....	80
Pré-requisitos .....	80
Etapa 1: Provisionar os recursos .....	82
Etapa 2: Configurar permissões para uma tabela do Iceberg .....	83
Etapa 3: Configurar permissões para uma tabela do Hudi .....	90
Etapa 4: Configurar permissões para uma tabela do Delta Lake .....	92
Etapa 5: Limpar os recursos do AWS .....	95
Como gerenciar um data lake usando o controle de acesso baseado em tags .....	95
Público-alvo .....	97
Pré-requisitos .....	98
Etapa 1: Provisionar os recursos .....	98
Etapa 2: registre sua localização de dados, crie uma ontologia LF-Tag e conceda permissões .....	99

Etapa 3: Criar bancos de dados do Lake Formation .....	103
Etapa 4: Conceder permissões de dados .....	113
Etapa 5: Executar uma consulta no Amazon Athena para verificar as permissões .....	115
Etapa 6: Limpar os recursos do AWS .....	116
Como proteger os data lakes com controle de acesso em nível de linha .....	117
Público-alvo .....	117
Pré-requisitos .....	118
Etapa 1: Provisionar os recursos .....	119
Etapa 2: Consulta sem filtros de dados .....	120
Etapa 3: Configurar filtros de dados e conceder permissões .....	122
Etapa 4: Consulta com filtros de dados .....	124
Etapa 5: Limpar os recursos do AWS .....	126
Compartilhe seus dados com segurança usando o Lake Formation .....	126
Público-alvo .....	127
Definir as configurações do Lake Formation .....	128
Etapa 1: Provisionar seus recursos usando modelos AWS CloudFormation .....	131
Etapa 2: Pré-requisitos de compartilhamento entre contas do Lake Formation .....	133
Etapa 3: Implementar o compartilhamento entre contas usando o método de controle de acesso baseado em tags .....	136
Etapa 4: Implementar o método de recurso nomeado .....	143
Etapa 5: Limpar os recursos do AWS .....	146
Compartilhamento de recursos do Catálogo de Dados com recursos externos da Contas da AWS usando um controle de acesso refinado .....	147
Público-alvo .....	148
Pré-requisitos .....	149
Etapa 1: Forneça acesso refinado a outra conta .....	150
Etapa 2: Forneça acesso refinado a um usuário na mesma conta .....	152
Permissões de integração ao Lake Formation .....	153
Visão geral das permissões do Lake Formation .....	154
Métodos para controle de acesso de alta granularidade .....	156
Controle de acesso a metadados .....	159
Controle de acesso a dados subjacente .....	163
Referência de personas e permissões do IAM do Lake Formation .....	168
AWS Lake Formation personas .....	168
AWS políticas gerenciadas para Lake Formation .....	170
Permissões sugeridas por personas .....	177

Alterando as configurações padrão do seu data lake .....	187
Permissões implícitas do Lake Formation .....	190
Referência de permissões do Lake Formation .....	192
Permissões do Lake Formation por tipo de recurso .....	193
Lake Formation concede e revoga comandos AWS CLI .....	195
Permissões do Lake Formation .....	200
Integrar o Centro de Identidade do IAM .....	214
Pré-requisitos .....	215
Conectar o Lake Formation ao Centro de Identidade do IAM .....	219
Atualizar uma integração com o Centro de Identidade do IAM .....	221
Excluir uma conexão do Lake Formation com o Centro de Identidade do IAM .....	223
Conceder permissões a usuários e grupos .....	223
Adicionar uma localização do Amazon S3 ao seu data lake .....	227
Requisitos para funções usadas para registrar locais .....	228
Registrando uma localização do Amazon S3 .....	233
Registrando uma localização criptografada do Amazon S3 .....	237
Registrando uma localização do Amazon S3 em outra conta AWS .....	242
Registrando uma localização criptografada do Amazon S3 em todas as contas AWS .....	244
Cancelar o registro de uma localização do Amazon S3 .....	249
Modo de acesso híbrido .....	249
Casos de uso comuns do modo de acesso híbrido .....	251
Como funciona o modo de acesso híbrido .....	253
Configurando o modo de acesso híbrido - cenários comuns .....	254
Removendo entidades principais e recursos do modo de acesso híbrido .....	272
Visualizando entidades principais e recursos no modo de acesso híbrido .....	273
Recursos adicionais do .....	274
Criando tabelas e bancos de dados do Catálogo de dados .....	274
Criação de um banco de dados .....	275
Criar tabelas .....	276
Trabalhar com visualizações .....	295
Importação de dados usando fluxos de trabalho .....	301
Esquema e fluxos de trabalho .....	302
Criação de um fluxo de trabalho .....	303
Executar um fluxo de trabalho .....	307
Gerenciando permissões do Lake Formation .....	309
Concedendo permissões de localização de dados .....	309

Concessão de permissões de localização de dados (mesma conta) .....	310
Concessão de permissões de localização de dados (conta externa) .....	312
Conceder permissões em um local de dados compartilhado com sua conta .....	316
Conceder e revogar permissões do Catálogo de dados .....	317
Permissões do IAM necessárias para conceder permissões do Lake Formation .....	318
Conceder permissões de data lake usando o método de recurso nomeado .....	321
Controle de acesso com base em tags .....	340
Conceder permissões de data lake usando o método LF-TBAC .....	387
Exemplo de cenário de permissões .....	394
Filtragem de dados e segurança em nível de célula .....	396
Visão geral da filtragem de dados .....	396
Filtros de data .....	398
Suporte partiQL em expressões de filtro de linha .....	402
Notas e restrições para filtragem em nível de coluna .....	404
Permissões necessárias para consultar tabelas com filtragem em nível de célula .....	406
Como gerenciar filtros de dados .....	407
Visualizar permissões de banco de dados e tabelas .....	422
Revogar permissões usando o console .....	426
Compartilhamento de dados entre contas .....	427
Pré-requisitos .....	430
Como atualizar as configurações da versão de compartilhamento de dados entre contas ....	434
Compartilhamento de tabelas e bancos de dados do Catálogo de Dados entre Contas da AWS e entidades principais do IAM a partir de contas externas .....	440
Conceder permissões em um banco de dados ou tabela compartilhada com sua conta .....	443
Como conceder permissões de links de recursos .....	445
Como acessar os dados subjacentes de uma tabela compartilhada .....	447
Registro em várias contas CloudTrail .....	449
Gerenciamento de permissões entre contas usando A e Lake Formation .....	454
Visualizando todas as concessões entre contas usando a operação de GetResourceShares API .....	457
Acessar e visualizar tabelas e bancos de dados compartilhados do Catálogo de Dados .....	458
Aceitar um convite de compartilhamento de recursos do AWS RAM .....	460
Visualizando tabelas e bancos de dados compartilhados do Catálogo de Dados .....	462
Criação de links de recursos .....	464
Como funcionam os links de recursos .....	464
Como criar um link de recurso para uma tabela compartilhada .....	467

Como criar um link de recurso para um banco de dados compartilhado .....	471
Gestão de links de recursos em APIs do AWS Glue. ....	475
Acessar tabelas entre regiões .....	479
Fluxos de trabalho .....	480
Como configurar o acesso à tabela entre regiões .....	484
Compartilhamento de dados no Lake Formation .....	488
Gerenciamento de permissões para dados em uma unidade de compartilhamento de dados do Amazon Redshift. ....	489
Pré-requisitos .....	490
Configuração de permissões para unidades de compartilhamento de dados do Amazon Redshift .....	491
Consultar bancos de dados federados .....	495
Gerenciamento de permissões em conjuntos de dados que usam repositórios de dados externos .....	496
Fluxo de trabalho .....	499
Pré-requisitos .....	500
Conectando o Catálogo de Dados a um repositório externo do Hive .....	502
Recursos adicionais do .....	505
Segurança .....	507
Proteção de dados .....	507
Criptografia em repouso .....	508
Segurança da infraestrutura .....	509
Prevenção contra o ataque do “substituto confuso” em todos os serviços .....	510
Registro em log de eventos de segurança no AWS Lake Formation .....	511
Integração com o Lake Formation .....	512
Como usar a integração de aplicativos Lake Formation .....	512
Como funciona a integração de aplicações do Lake Formation .....	513
Perfis e responsabilidades na integração de aplicativos do Lake Formation .....	515
Fluxo de trabalho do Lake Formation para operações de API de integração de aplicativos ..	516
Como registrar um mecanismo de consulta de terceiros .....	517
Como habilitar permissões para que um mecanismo de consulta de terceiros chame operações de API de integração de aplicativos .....	519
Integração de aplicativos para acesso total à tabela .....	523
Trabalhando com outros AWS serviços .....	526
Amazon Athena .....	526
Suporte a formatos de tabelas transacionais .....	528

Recursos adicionais do .....	531
Amazon Redshift Spectrum .....	531
Suporte para tipos de tabelas transacionais .....	532
Recursos adicionais do .....	534
AWS Glue .....	534
Suporte para tipos de tabelas transacionais .....	535
Recursos adicionais do .....	536
Amazon EMR .....	536
Suporte a formatos de tabelas transacionais .....	537
Recursos adicionais do .....	538
Amazon QuickSight .....	539
Recursos adicionais do .....	539
AWS CloudTrail Lago .....	539
Como fazer log de chamados de API do Lake Formation do AWS usando AWS CloudTrail .....	541
Informações sobre o Lake Formation no CloudTrail .....	541
Como compreender os eventos do Lake Formation .....	542
Práticas recomendadas, considerações e limitações do Lake Formation .....	545
Práticas recomendadas e considerações sobre compartilhamento de dados entre contas .....	545
Limitações de acesso aos dados entre regiões .....	548
Considerações e limitações das visualizações do catálogo de dados .....	548
Limitações de filtragem de dados .....	549
Considerações e limitações do modo de acesso híbrido .....	551
Considerações e limitações do compartilhamento de dados de armazenamento de metadados do Hive .....	552
Limitações do compartilhamento de dados do Amazon Redshift .....	554
Limitações da integração com o Centro de Identidade do IAM .....	556
Considerações e práticas recomendadas de controle de acesso com base em tags do Lake Formation .....	556
Formatos e limitações compatíveis para compactação gerenciada de dados .....	559
Solução de problemas do Lake Formation .....	562
Solução de problemas gerais .....	562
Erro: permissões insuficientes do Lake Formation em <Amazon S3 location> .....	562
Erro: “permissões de chave de criptografia insuficientes para a API Glue” .....	563
Minha consulta do Amazon Athena ou do Amazon Redshift que usa manifestos está falhando .....	563

Erro: "permissão(ões) do Lake Formation insuficiente(s): necessária a criação de tag no catálogo" .....	563
Erro ao excluir administradores de data lake inválidos .....	563
Resolução de problemas de acesso entre contas .....	563
Eu concedi uma permissão para várias contas do Lake Formation, mas o destinatário não consegue ver o recurso .....	564
As entidades principais da conta do destinatário podem ver o recurso do Catálogo de Dados, mas não podem acessar os dados subjacentes .....	564
Erro: “falha na associação porque o chamador não foi autorizado” ao aceitar um convite de compartilhamento de recursos AWS RAM. ....	565
Erro: “não autorizado a conceder permissões para o recurso” .....	565
Erro: “acesso negado para recuperar informações da organização da AWS” .....	566
Erro: “organização <organization-ID> não encontrada” .....	566
Erro: "permissões insuficientes do Lake Formation: combinação ilegal" .....	566
ConcurrentModificationException em solicitações de concessão/revogação para contas externas .....	566
Erro ao usar o Amazon EMR para acessar dados compartilhados por meio de várias contas .....	566
Solução de problemas em esquemas e fluxos de trabalho .....	568
O meu esquema falhou com "usuário: <user-ARN> não está autorizado a realizar: iam:PassRole no recurso: <role-ARN>" .....	568
O meu esquema falhou com "usuário: <user-ARN> não está autorizado a realizar: iam:PassRole no recurso: <role-ARN>" .....	568
Um crawler no meu fluxo de trabalho falhou com “o recurso não existe ou o solicitante não está autorizado a acessar as permissões solicitadas” .....	569
Um crawler no meu fluxo de trabalho falhou com “ocorreu um erro (AccessDeniedException) ao chamar a operação CreateTable...” .....	569
Problemas conhecidos do AWS Lake Formation .....	569
Limitação na filtragem de metadados da tabela .....	570
Problema ao renomear uma coluna excluída .....	571
Problema com a exclusão de colunas em tabelas CSV .....	571
As partições da tabela devem ser adicionadas em um caminho comum .....	571
Problema com a criação de um banco de dados durante a criação do fluxo de trabalho .....	571
Problema com a exclusão e a recriação de um usuário .....	572
As APIs GetTables e SearchTables não atualizam o valor do parâmetro IsRegisteredWithLakeFormation .....	572

As operações da API do Catálogo de dados não atualizam o valor do parâmetro <code>IsRegisteredWithLakeFormation</code> .....	572
As operações do Lake Formation não oferecem suporte ao Registro de esquemas AWS Glue .....	572
Mensagem de erro atualizada .....	573
API Lake Formation .....	574
Permissões .....	575
— operações — .....	575
— tipos de dados — .....	575
Configurações do data lake .....	576
— operações — .....	576
— tipos de dados — .....	576
Integração com o Centro de Identidade do IAM .....	576
— operações — .....	576
— tipos de dados — .....	576
Modo de acesso híbrido .....	576
— operações — .....	577
— tipos de dados — .....	575
Fornecimento de credenciais .....	577
— operações — .....	577
— tipos de dados — .....	578
Tags .....	578
— operações — .....	578
— tipos de dados — .....	578
APIs de filtro de dados .....	579
— operações — .....	579
— tipos de dados — .....	579
Tipos de dados comuns .....	579
ErrorDetail .....	579
Padrões de string .....	580
Regiões compatíveis .....	581
Disponibilidade geral .....	581
AWS GovCloud (US) .....	581
Otimização de transações e armazenamento .....	581
Histórico do documento .....	584
Glossário do AWS .....	597



---

..... dxcviii

# O que é o AWS Lake Formation?

Bem-vindo ao Guia do desenvolvedor do AWS Lake Formation.

O AWS Lake Formation ajuda você a governar, proteger e compartilhar dados de forma centralizada e global para análise e machine learning. Com o Lake Formation, você pode gerenciar um controle de acesso refinado para seus dados de data lake no Amazon Simple Storage Service (Amazon S3) e seus metadados no AWS Glue Data Catalog.

O Lake Formation fornece seu próprio modelo de permissões que amplia o modelo de permissões do IAM. O modelo de permissões do Lake Formation permite acesso refinado aos dados armazenados em lagos de dados por meio de um mecanismo simples de concessão ou revogação, muito parecido com um sistema de gerenciamento de banco de dados relacional (RDBMS). As permissões do Lake Formation são aplicadas usando controles granulares nos níveis de coluna, linha e célula em todos os serviços de análise e machine learning da AWS, incluindo Amazon Athena, Amazon QuickSight, Amazon Redshift Spectrum, Amazon EMR e AWS Glue.

O modo de acesso híbrido do Lake Formation AWS Glue Data Catalog permite que você proteja e acesse os dados catalogados usando as permissões do Lake Formation e as políticas de permissões do IAM para Amazon S3 AWS Glue e ações. Com o modo de acesso híbrido, os administradores de dados podem integrar as permissões do Lake Formation de forma seletiva e incremental, concentrando-se em um caso de uso do data lake por vez.

O Lake Formation também permite que você compartilhe dados interna e externamente em várias Contas da AWS, organizações da AWS ou diretamente com as entidades principais do IAM em outra conta, fornecendo acesso refinado a dados subjacentes e metadados do AWS Glue Data Catalog.

## Tópicos

- [Características do Lake Formation](#)
- [AWS Lake Formation: como funciona](#)
- [Componentes do Lake Formation](#)
- [Terminologia do Lake Formation](#)
- [Integrações de serviços da AWS com o Lake Formation](#)
- [Recursos adicionais do Lake Formation](#)
- [Introdução ao Lake Formation](#)

# Características do Lake Formation

O Lake Formation o ajuda a desfazer monopólios de dados e combinar diferentes tipos de dados estruturados e não estruturados em um repositório centralizado. Primeiro, identifique os armazenamentos de dados existentes no Amazon S3 ou nos bancos de dados relacionais e NoSQL e mova os dados para o seu data lake. Em seguida, rastreie, catalogue e prepare os dados para análise. Em seguida, forneça aos usuários acesso seguro de autoatendimento aos dados por meio de serviços de análise de sua escolha.

## Tópicos

- [Ingestão e gerenciamento de dados](#)
- [Gerenciamento de segurança](#)
- [Compartilhamento de dados](#)

## Ingestão e gerenciamento de dados

### Importar dados de bancos de dados já existentes na AWS

Depois de especificar onde estão seus bancos de dados existentes e fornecer suas credenciais de acesso, o Lake Formation lê os dados e seus metadados (esquema) para entender o conteúdo da fonte de dados. Em seguida, ele importa os dados para seu novo data lake e registra os metadados em um catálogo central. Com o Lake Formation, você pode importar dados de bancos de dados MySQL, PostgreSQL, SQL Server, MariaDB e Oracle executados no Amazon RDS ou hospedados no Amazon EC2. Tanto o carregamento de dados em massa quanto o incremental são suportados.

### Importar dados de outras fontes externas

Você pode usar o Lake Formation para mover dados de bancos de dados on-premises conectando-se ao Java Database Connectivity (JDBC). Identifique suas fontes de destino e forneça credenciais de acesso no console, e o Lake Formation lê e carrega seus dados no data lake. Para importar dados de bancos de dados diferentes dos listados acima, você pode criar trabalhos ETL personalizados com o AWS Glue.

### Catalogue e rotule seus dados

Você pode usar os crawlers do AWS Glue para ler seus dados no Amazon S3 e extrair o esquema do banco de dados e da tabela e armazenar esses dados em um AWS Glue Data Catalog

pesquisável. Em seguida, use o Lake Formation [Controle de acesso baseado em tags do Lake Formation](#) (TBAC) para gerenciar as permissões em bancos de dados, tabelas e colunas. Para obter mais informações sobre como adicionar tabelas ao Catálogo de Dados, consulte [Criando tabelas e bancos de dados do Catálogo de dados](#).

## Gerenciamento de segurança

### Defina e gerencie controles de acesso

O Lake Formation fornece um único local para gerenciar os controles de acesso aos dados em seu data lake. Você pode definir políticas de segurança que restrinjam o acesso aos dados nos níveis de banco de dados, tabela, coluna, linha e célula. Essas políticas se aplicam a usuários e funções do IAM e a usuários e grupos durante a federação por meio de um provedor de identidade externo. Você pode usar controles refinados para acessar dados protegidos pelo Lake Formation no Amazon Redshift Spectrum, Athena, AWS Glue ETL e Amazon EMR para Apache Spark. Sempre que você criar identidades do IAM, siga as práticas recomendadas do IAM. Para obter mais informações, consulte [Práticas recomendadas de segurança](#) no Guia do usuário do IAM.

### Modo de acesso híbrido

O modo de acesso híbrido do Lake Formation oferece a flexibilidade de habilitar seletivamente as permissões do Lake Formation para bancos de dados e tabelas no AWS Glue Data Catalog. Com o modo de acesso híbrido, agora você tem um caminho incremental que permite definir permissões do Lake Formation para um conjunto específico de usuários sem interromper as políticas de permissão de outros usuários ou workload existentes. Para ter mais informações, consulte [Modo de acesso híbrido](#).

### Implementar o registro em log de auditoria

O Lake Formation fornece registros de auditoria abrangentes CloudTrail para monitorar o acesso e mostrar a conformidade com políticas definidas centralmente. Você pode auditar o histórico de acesso aos dados em serviços de análise e machine learning que lêem os dados em seu data lake por meio do Lake Formation. Isso permite que você veja quais usuários ou funções tentaram acessar quais dados, com quais serviços e quando. Você pode acessar os registros de auditoria da mesma forma que acessa qualquer outro CloudTrail registro usando as CloudTrail APIs e o console. Para obter mais informações sobre CloudTrail registros, consulte [Como fazer log de chamados de API do Lake Formation do AWS usando AWS CloudTrail](#).

### Segurança no nível de linha e célula

O Lake Formation fornece filtros de dados que permitem restringir o acesso a uma combinação de colunas e linhas. Use a segurança em nível de linha e célula para proteger dados confidenciais, como Informações Pessoais Identificáveis (IPI). Para obter mais informações sobre segurança em nível de linha, consulte [Visão geral da filtragem de dados](#).

### Controle de acesso com base em tags

Use o [controle de acesso baseado em tags](#) do Lake Formation para gerenciar centenas ou até milhares de permissões de dados criando rótulos personalizados chamados LF-tags. Agora você pode definir tags LF e anexá-las a bancos de dados, tabelas ou colunas. Em seguida, compartilhe o acesso controlado por meio de serviços de análises, de machine learning (ML) e de extração, transformação e carregamento (ETL) para consumo. As tags LF garantem que a governança de dados possa ser escalada facilmente substituindo as definições de políticas de milhares de recursos por algumas tags lógicas. O Lake Formation fornece uma pesquisa baseada em texto sobre esses metadados, para que seus usuários possam encontrar rapidamente os dados que precisam analisar.

### Acesso entre contas

Os recursos de gerenciamento de permissões do Lake Formation simplificam a proteção e o gerenciamento de lagos de dados distribuídos em várias contas AWS por meio de uma abordagem centralizada, fornecendo controle de acesso refinado ao catálogo de dados e aos locais do Amazon S3. Para ter mais informações, consulte [Compartilhamento de dados entre contas no Lake Formation](#).

## Compartilhamento de dados

O recurso de compartilhamento de dados permite configurar permissões em conjuntos de dados armazenados em diferentes fontes de dados, como o Amazon Redshift, sem migrar dados ou metadados para o Amazon S3 ou o AWS Glue Data Catalog. Você pode usar os seguintes métodos para compartilhar dados no Lake Formation:

Para obter mais informações, consulte [Compartilhamento de dados no Lake Formation](#).

- Integração do Lake Formation com o compartilhamento de dados do Amazon Redshift: usar o Lake Formation para gerenciar centralmente as permissões de acesso por banco de dados, tabela, coluna e linha das unidades de compartilhamento de dados do Amazon [Redshift e restringir o acesso dos usuários a objetos](#) em uma unidade de compartilhamento de dados.
- Conexão a repositórios externos do AWS Glue Data Catalog — Conecte-se a repositórios externos do AWS Glue Data Catalog para gerenciar permissões de acesso em conjuntos de dados no

Amazon S3 usando o Lake Formation. Não é necessária nenhuma migração de metadados para o AWS Glue Data Catalog.

Para obter mais informações, consulte [Gerenciamento de permissões em conjuntos de dados que usam repositórios de dados externos](#).

- Como integrar o Lake Formation com o AWS Data Exchange: o Lake Formation oferece suporte ao licenciamento de acesso aos seus dados por meio de AWS Data Exchange. Se você estiver interessado em licenciar seus dados do Lake Formation, consulte [O que está AWS Data Exchange](#) no Guia do Usuário AWS Data Exchange.

## AWS Lake Formation: como funciona

AWS Lake Formation fornece um modelo de permissões do sistema de gerenciamento de banco de dados relacional (RDBMS) para conceder ou revogar o acesso aos recursos do catálogo de dados, como bancos de dados, tabelas e colunas com dados subjacentes no Amazon S3. As permissões fáceis de gerenciar do Lake Formation substituem as políticas complexas de bucket do Amazon S3 e as políticas correspondentes do IAM.

No Lake Formation, você pode implementar permissões em dois níveis:

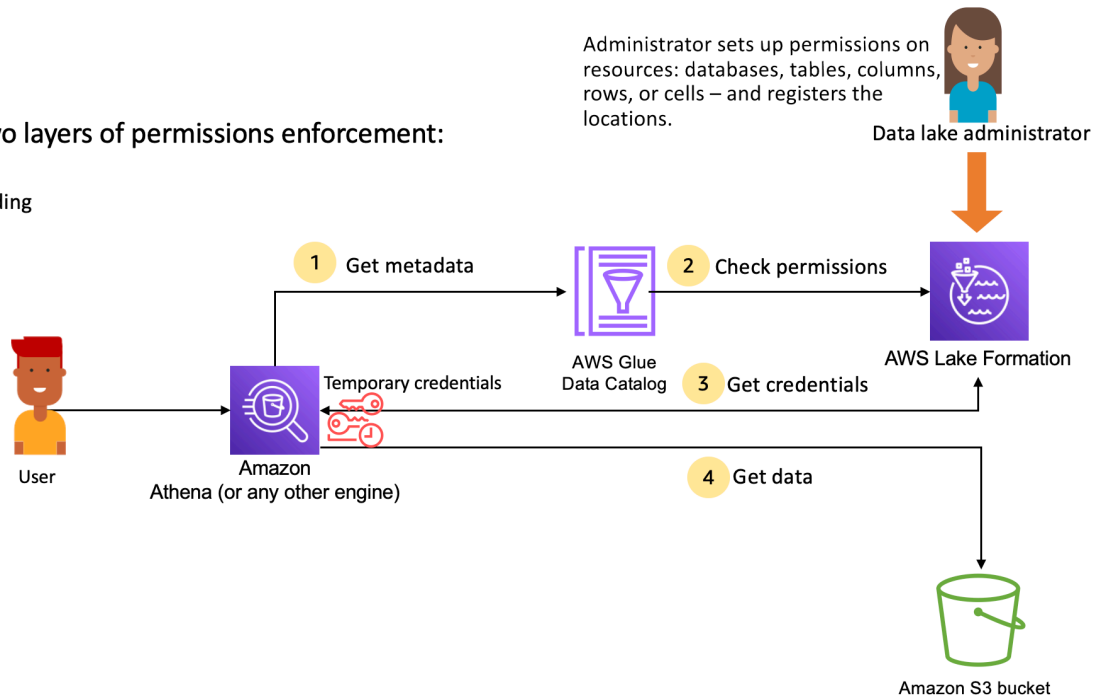
- Aplicação de permissões em nível de metadados nos recursos do Catálogo de dados, como bancos de dados e tabelas
- Gerenciar permissões de acesso ao armazenamento nos dados subjacentes armazenados no Amazon S3 em nome de mecanismos integrados

## Fluxo de trabalho de gerenciamento de permissões do Lake Formation

O Lake Formation se integra aos mecanismos analíticos para consultar armazenamentos de dados e objetos de metadados do Amazon S3 registrados no Lake Formation. O diagrama a seguir mostra como o gerenciamento de permissões funciona no Lake Formation.

## Lake Formation provides two layers of permissions enforcement:

- Metadata layer – Data Catalog
- Storage layer – Credential vending



## Etapas de alto nível do gerenciamento de permissões do Lake Formation

Antes que o Lake Formation possa fornecer controles de acesso aos dados em seu data lake, um [administrador de data lake](#) ou um usuário com permissões administrativas configura políticas de usuário individuais da tabela do Catálogo de dados para permitir ou negar acesso às tabelas do Catálogo de dados usando as permissões do Lake Formation.

Em seguida, o administrador do data lake ou um usuário delegado pelo administrador concede permissões do Lake Formation aos usuários nos bancos de dados e tabelas do Catálogo de dados e registra a localização da tabela no Amazon S3 no Lake Formation.

1. Obter metadados — Uma entidade principal (usuário) envia uma consulta ou um script de ETL para um [mecanismo analítico integrado](#), como Amazon Athena, AWS Glue, Amazon EMR ou Amazon Redshift Spectrum. O mecanismo analítico integrado identifica a tabela que está sendo solicitada e envia uma solicitação de metadados para o Catálogo de dados.
2. Verificar permissões — O Catálogo de dados verifica as permissões do usuário com o Lake Formation e, se o usuário estiver autorizado a acessar a tabela, retorna os metadados que o usuário tem permissão para ver para o mecanismo.
3. Obter credenciais — O Catálogo de dados permite que o mecanismo saiba se a tabela é gerenciada pelo Lake Formation ou não. Se os dados subjacentes forem registrados no Lake Formation, o mecanismo analítico solicitará que o Lake Formation forneça acesso aos dados concedendo acesso temporário.

4. Obter dados — Se o usuário estiver autorizado a acessar a tabela, o Lake Formation fornecerá acesso temporário ao mecanismo analítico integrado. Usando o acesso temporário, o mecanismo analítico busca os dados do Amazon S3 e executa a filtragem necessária, como filtragem por coluna, linha ou célula. Quando o mecanismo termina de executar o trabalho, ele retorna os resultados para o usuário. Esse processo chamado de [fornecimento de credenciais](#).

Se a tabela não for gerenciada pelo Lake Formation, a segunda chamada do mecanismo analítico será feita diretamente para o Amazon S3. A política de bucket do Amazon S3 e a política de usuário do IAM em questão são avaliadas quanto ao acesso aos dados.

Sempre que você usar as políticas do IAM, siga as práticas recomendadas do IAM. Para obter mais informações, consulte [Práticas recomendadas de segurança no IAM](#) no Manual do usuário do IAM.

## Tópicos

- [Permissões de metadados](#)
- [Gerenciar o acesso às do armazenamento](#)
- [Compartilhamento de dados entre contas no Lake Formation](#)

## Permissões de metadados

O Lake Formation fornece autorização e controle de acesso ao Catálogo de dados. Quando um perfil do IAM faz uma chamada à API do Catálogo de dados de qualquer sistema, o Catálogo de dados verifica as permissões de dados do usuário e retorna somente os metadados que o usuário tem permissão para acessar. Por exemplo, se um perfil do IAM tiver acesso a somente uma tabela em um banco de dados e um serviço ou um usuário supondo que a função execute a operação `GetTables`, a resposta conterá somente uma tabela, independentemente do número de tabelas no banco de dados.

### Configurações padrão - permissões de grupo **IAMAllowedPrincipal**

AWS Lake Formation, por padrão, define permissões para todos os bancos de dados e tabelas para um grupo virtual chamado `IAMAllowedPrincipal`. Esse grupo é único e visível somente dentro do Lake Formation. O grupo `IAMAllowedPrincipal` inclui todas as entidades principais do IAM que possuem acesso aos recursos do Catálogo de dados por meio das políticas de entidades principais e de AWS Glue recursos do IAM. Se essas permissões existirem em um banco de dados ou tabela, todas as entidades principais terão acesso ao banco de dados ou à tabela.



Se você quiser fornecer permissões mais granulares em um banco de dados ou tabela, remova a permissão `IAMAllowedPrincipal` e o Lake Formation aplicará todas as outras políticas associadas a esse banco de dados ou tabela. Por exemplo, se houver uma política que permita ao Usuário A acessar o Banco de Dados A com permissões `DESCRIBE` e `IAMAllowedPrincipal` existir com todas as permissões, o Usuário A continuará executando todas as outras ações até que a permissão `IAMAllowedPrincipal` seja revogada.

Além disso, por padrão, o grupo `IAMAllowedPrincipal` tem permissões em todos os novos bancos de dados e tabelas quando eles são criados. Há duas configurações que controlam esse comportamento. A primeira está no nível da conta e da região, que permite isso para bancos de dados recém-criados, e a segunda está no nível do banco de dados. Para modificar a configuração padrão, consulte [Altere o modelo de permissão padrão ou use o modo de acesso híbrido](#).

## Conceder permissões

Os administradores do data lake podem conceder permissões do Catálogo de dados aos diretores para que eles possam criar e gerenciar bancos de dados e tabelas e acessar os dados subjacentes.

### Permissões em nível de banco de dados e tabela

Quando você concede permissões no Lake Formation, o concedente deve especificar o diretor ao qual conceder permissões, os recursos para os quais conceder permissões e as ações que o beneficiário deve ter acesso para realizar. Para a maioria dos recursos do Lake Formation, a lista de entidades principais e os recursos para conceder permissões são semelhantes, mas as ações que um beneficiário pode realizar diferem com base no tipo de recurso. Por exemplo, as permissões `SELECT` estão disponíveis para que as tabelas leiam as tabelas, mas as permissões `SELECT` não são permitidas nos bancos de dados. A permissão `CREATE_TABLE` é permitida em bancos de dados, mas não em tabelas.

Você pode conceder permissões AWS Lake Formation usando dois métodos:

- [Método de recurso nomeado](#) — Permite escolher nomes de banco de dados e tabelas enquanto concede permissões aos usuários.
- [Controle de acesso baseado em LF-Tag \(LF-TBAC\)](#) — Os usuários criam LF-Tags, as associam aos recursos do Catálogo de dados, concedem permissão `Describe` sobre LF-Tags, associam permissões a usuários individuais e escrevem políticas de permissões LF usando LF-Tags para usuários diferentes. Essas políticas baseadas em LF-Tag se aplicam a todos os recursos do Catálogo de dados associados a esses valores de LF-Tags.

**Note**

As LF-Tags são exclusivas do Lake Formation. Elas só são visíveis no Lake Formation e não devem ser confundidas com tags de recursos AWS.

O LF-TBAC é um recurso que permite aos usuários agrupar recursos em categorias definidas pelo usuário de LF-Tags e aplicar permissões a esses grupos de recursos. Portanto, é a melhor maneira de escalar as permissões em um grande número de recursos do Catálogo de dados.

Para obter mais informações, consulte [Controle de acesso baseado em tags do Lake Formation](#).

Quando você concede permissões a uma entidade principal, o Lake Formation avalia as permissões como uma união de todas as políticas desse usuário. Por exemplo, se você tiver duas políticas em uma tabela para uma entidade principal em que uma política concede permissões às colunas col1, col2 e col3 por meio do método de recurso nomeado, e a outra política concede permissões para a mesma tabela e principal para col5 e col6 por meio de LF-Tags, as permissões efetivas serão uma união das permissões que seriam col1, col2, col3, col5 e col6. Isso também inclui filtros de dados e linhas.

### Permissões de localização de dados

As permissões de localização de dados fornecem aos usuários não administrativos a capacidade de criar bancos de dados e tabelas em locais específicos do Amazon S3. Se um usuário tentar criar um banco de dados ou uma tabela em um local que não tenha permissão para criar, a tarefa de criação falhará. Isso evita que os usuários criem tabelas em locais arbitrários dentro do data lake e fornece controle sobre onde esses usuários podem ler e gravar dados. Há uma permissão implícita ao criar tabelas na localização do Amazon S3 dentro do banco de dados em que elas estão sendo criadas. Para obter mais informações, consulte [Concedendo permissões de localização de dados](#).

### Crie permissões de tabela e banco de dados

Usuários não administrativos, por padrão, não têm permissões para criar bancos de dados ou tabelas em um banco de dados. A criação do banco de dados é controlada no nível da conta usando as configurações do Lake Formation para que somente entidades principais autorizadas possam criar bancos de dados. Para obter mais informações, consulte [Criação de um banco de dados](#). Para criar uma tabela, uma entidade principal requer CREATE\_TABLE permissão no banco de dados em que a tabela está sendo criada. Para obter mais informações, consulte [Criar tabelas](#).

## Permissões implícitas e explícitas

O Lake Formation fornece permissões implícitas, dependendo da persona e das ações que a persona realiza. Por exemplo, os administradores do data lake obtêm automaticamente permissões DESCRIBE para todos os recursos no Catálogo de dados, permissões de localização de dados em todos os locais, permissões para criar bancos de dados e tabelas em todos os locais, bem como permissões Grant e Revoke em qualquer recurso. Os criadores de banco de dados obtêm automaticamente todas as permissões de banco de dados nos bancos de dados que eles criam, e os criadores de tabelas obtêm todas as permissões nas tabelas que eles criam. Para obter mais informações, consulte [Permissões implícitas do Lake Formation](#).

## Permissões concedidas

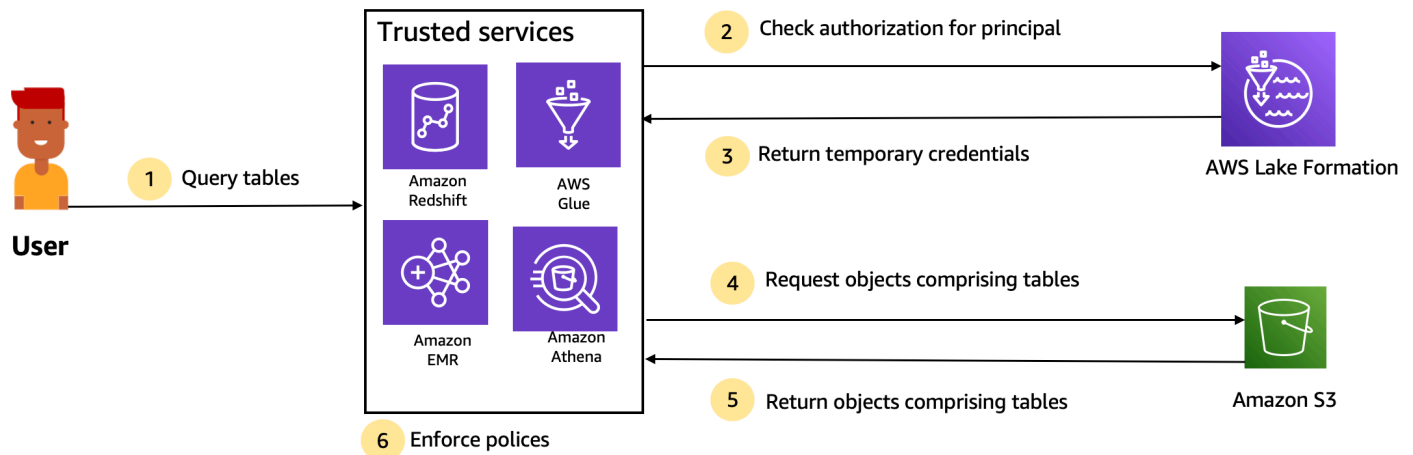
Os administradores do data lake têm a capacidade de delegar o gerenciamento de permissões a usuários não administrativos fornecendo permissões concedidas. Quando uma entidade principal recebe permissões concedidas sobre um recurso e um conjunto de permissões, essa entidade principal ganha a capacidade de conceder permissões a outras entidades principais nesse recurso.

## Gerenciar o acesso às do armazenamento

O Lake Formation usa a funcionalidade de [fornecimento de credenciais](#) para conceder acesso temporário aos dados do Amazon S3. O fornecimento de credenciais, ou fornecimento de tokens, é um padrão comum que disponibiliza credenciais temporárias a usuários, serviços ou alguma outra entidade com o objetivo de conceder acesso de curto prazo a um recurso.

A Lake Formation aproveita esse padrão para fornecer acesso de curto prazo a serviços de análise AWS, como o Athena, para acessar dados em nome da entidade principal responsável pela chamada. Ao conceder permissões, os usuários não precisam atualizar suas políticas de bucket do Amazon S3 ou políticas do IAM e não precisam de acesso direto ao Amazon S3.

O diagrama a seguir mostra como o Lake Formation fornece acesso temporário aos locais registrados:



Trusted services enforce AWS Lake Formation policies (distributed enforcement with fail close).

1. Uma entidade principal (usuário) insere uma consulta ou solicitação de dados para uma tabela por meio de um serviço integrado confiável, como Athena, Amazon EMR, Redshift Spectrum ou AWS Glue.
2. O serviço integrado verifica a autorização da Lake Formation para a tabela e as colunas solicitadas e determina a autorização. Se o usuário não estiver autorizado, o Lake Formation nega o acesso aos dados e a consulta falhará.
3. Depois que a autorização é bem-sucedida e a autorização de armazenamento é ativada para a tabela e o usuário, o serviço integrado recupera as credenciais temporárias do Lake Formation para acessar os dados.
4. O serviço integrado usa as credenciais temporárias do Lake Formation para solicitar objetos do Amazon S3.
5. O Amazon S3 fornece objetos do Amazon S3 para o serviço integrado. Os objetos do Amazon S3 contêm todos os dados da tabela.
6. O serviço integrado executa a aplicação necessária das políticas do Lake Formation, como filtragem em nível de coluna, nível de linha e/ou nível de célula. O serviço integrado processa as consultas e retorna os resultados ao usuário.

Habilite a aplicação de permissões em nível de armazenamento para tabelas do Catálogo de dados

Por padrão, a imposição em nível de armazenamento não está habilitada para tabelas no Catálogo de dados. Para permitir a fiscalização em nível de armazenamento, você deve registrar a localização dos seus dados de origem no Amazon S3 no Lake Formation e fornecer um perfil do IAM. As

permissões em nível de armazenamento serão habilitadas para todas as tabelas com o mesmo caminho de localização da tabela ou prefixo da localização do Amazon S3.

Quando um serviço integrado solicita acesso ao local de dados em nome de um usuário, o serviço Lake Formation assume essa função e retorna as credenciais ao serviço solicitado com permissões reduzidas ao recurso para que o acesso aos dados possa ser feito. O perfil do IAM registrado deve ter todo o acesso necessário às do local do Amazon S3, incluindo chaves AWS KMS.

Para obter mais informações, consulte [Registrando uma localização do Amazon S3](#).

### Serviços da AWS compatíveis

Serviços analíticos AWS, como Athena, Redshift Spectrum, Amazon EMR, AWS Glue, Amazon QuickSight e Amazon SageMaker se integram ao Lake Formation AWS usando as operações da API de fornecimento automático de credenciais do Lake Formation. Para ver uma lista completa dos serviços AWS que se integram ao Lake Formation e o nível de granularidade e os formatos de tabela compatíveis com eles, consulte [Trabalhando com outros AWS serviços](#).

## Compartilhamento de dados entre contas no Lake Formation

Com o Lake Formation, você pode compartilhar recursos do Catálogo de dados (bancos de dados e tabelas) dentro de uma conta AWS e entre contas em uma configuração simples usando o método de recurso nomeado ou LF-Tags. Você pode compartilhar um banco de dados inteiro ou selecionar tabelas de um banco de dados com qualquer entidade principal do IAM (perfis e usuários do IAM) em uma conta, com outras contas AWS no nível da conta ou diretamente com entidades principais do IAM em outra conta.

Você também pode compartilhar tabelas do Catálogo de dados com filtros de dados para restringir o acesso aos detalhes em nível de linha e de célula. O Lake Formation usa AWS Resource Access Manager (AWS RAM) para facilitar a concessão de permissões entre contas. Quando um recurso é compartilhado entre duas contas, AWS RAM envia convites para a conta do destinatário. Quando um usuário aceita um convite de compartilhamento AWS RAM, AWS RAM fornece as permissões necessárias para que o Lake Formation tenha os recursos do Catálogo de dados disponíveis, bem como habilite a imposição do nível de armazenamento. Para obter mais informações, consulte [Compartilhamento de dados entre contas no Lake Formation](#).

Quando o administrador do data lake da conta do destinatário aceita o compartilhamento AWS RAM, os recursos compartilhados ficam disponíveis na conta do destinatário. O administrador do data lake concede mais permissões do Lake Formation no recurso compartilhado a outras entidades

principais do IAM na conta do destinatário, se o administrador tiver permissões GRANTABLE no recurso compartilhado.

No entanto, as entidades principais não podem consultar os recursos compartilhados usando o Athena ou o Redshift Spectrum sem um link de recurso. Um link de recurso é uma entidade no Catálogo de dados e é semelhante ao conceito de Linux-Symlink.

O administrador do data lake da conta do destinatário cria um link no recurso compartilhado. O administrador concede permissões `Describe` no link do recurso com as permissões necessárias no recurso compartilhado original para outros usuários. Um usuário na conta do destinatário pode então usar o link para consultar o recurso compartilhado usando o Athena e o Redshift Spectrum. Para obter mais informações sobre links de recursos, consulte [Criação de links de recursos](#).

## Componentes do Lake Formation

AWS Lake Formation depende da interação de vários componentes para criar e gerenciar seu data lake.

### Lake Formation

Você usa o console do Lake Formation para definir e gerenciar seu data lake e conceder e revogar permissões do Lake Formation. Você pode usar esquemas no console para descobrir, limpar, transformar e ingerir dados. É possível habilitar ou desabilitar o acesso ao console para usuários individuais do Lake Formation.

### API e interface de linha de comando do Lake Formation

Lake Formation fornece operações de API por meio de vários SDKs específicos de linguagem e do AWS Command Line Interface (AWS CLI). O Lake Formation API funciona em conjunto com a API AWS Glue. A API Lake Formation se concentra principalmente no gerenciamento de permissões do Lake Formation, enquanto a API AWS Glue fornece uma API de catálogo de dados e uma infraestrutura gerenciada para definir, programar e executar operações de ETL em seus dados.

Para obter informações sobre a API AWS Glue, consulte o [AWS Glue Guia do desenvolvedor](#). Para obter informações sobre como usar o AWS CLI, consulte a [AWS CLI Referência de comandos](#).

## Outros produtos da AWS

O Lake Formation usa os seguintes serviços:

- [AWS Glue](#) para orquestrar trabalhos e rastreadores para transformar dados usando as transformações AWS Glue.
- O [IAM](#) concederá políticas de permissões às entidades principais do Lake Formation. O modelo de permissão do Lake Formation aumenta o modelo de permissão do IAM para proteger seu data lake.

## Terminologia do Lake Formation

A seguir estão alguns termos importantes que você encontrará neste guia.

### Data lake

O data lake são seus dados persistentes que são armazenados no Amazon S3 e gerenciados pelo Lake Formation usando um Catálogo de dados. Um data lake normalmente armazena o seguinte:

- Dados estruturados e não estruturados
- Dados brutos e dados transformados

Para que um caminho do Amazon S3 esteja dentro de um data lake, ele deve ser registrado com o Lake Formation.

### Acesso aos dados

O Lake Formation fornece acesso seguro e granular aos dados por meio de um novo modelo de concessão/revogação de permissões que amplia as políticas (IAM) AWS Identity and Access Management.

Analistas e cientistas de dados podem usar o portfólio completo de serviços analíticos AWS e de machine learning, como o Amazon Athena, para acessar os dados. As políticas de segurança configuradas do Lake Formation ajudam a garantir que os usuários possam acessar somente os dados que estão autorizados a acessar.

### Modo de acesso híbrido

O modo de acesso híbrido permite proteger e acessar os dados catalogados usando as permissões do Lake Formation e as permissões do IAM e do Amazon S3. O modo de acesso híbrido permite que os administradores de dados integrem as permissões do Lake Formation de forma seletiva e incremental, concentrando-se em um caso de uso do data lake por vez.

## Blueprint

Um esquema é um modelo de gerenciamento de dados que permite a ingestão fácil de dados em um data lake. O Lake Formation fornece vários esquemas, um para cada tipo de fonte predefinido, como um banco de dados relacional ou logs AWS CloudTrail. A partir de um esquema, você pode criar um fluxo de trabalho. Os fluxos de trabalho consistem em crawlers AWS Glue, tarefas e acionadores que são gerados para orquestrar o carregamento e a atualização dos dados. Os esquemas usam a fonte de dados, o destino dos dados e o cronograma como entrada para configurar o fluxo de trabalho.

## Fluxo de trabalho

Um fluxo de trabalho é um contêiner de um conjunto de tarefas relacionadas AWS Glue, crawlers e gatilhos. Você cria o fluxo de trabalho no Lake Formation e ele é executado no serviço AWS Glue. O Lake Formation pode rastrear o status de um fluxo de trabalho como uma entidade única.

Ao definir um fluxo de trabalho, você seleciona o esquema no qual ele se baseia. Em seguida, você pode executar fluxos de trabalho sob demanda ou de acordo com um cronograma.

Os fluxos de trabalho que você cria no Lake Formation são visíveis no console AWS Glue como um gráfico acíclico direcionado (DAG). Usando o DAG, você pode acompanhar o andamento do fluxo de trabalho e solucionar o problema.

## Catálogo de dados

O Catálogo de dados é seu armazenamento de metadados persistente. É um serviço gerenciado que permite armazenar, anotar e compartilhar metadados na Nuvem AWS da mesma maneira que você faria em uma metastore do Apache Hive. Ele fornece um repositório uniforme onde sistemas diferentes podem armazenar e encontrar metadados para rastrear dados em silos de dados e, em seguida, usar esses metadados para consultar e transformar os dados. O Lake Formation usa o Catálogo de dados AWS Glue para armazenar metadados sobre data lakes, fontes de dados, transformações e destinos.

Os metadados sobre fontes e destinos de dados estão na forma de bancos de dados e tabelas. As tabelas armazenam informações de esquemas, localização e muito mais. Bancos de dados são coleções de tabelas. O Lake Formation fornece uma hierarquia de permissões para controlar o acesso a bancos de dados e tabelas no Catálogo de dados.

Cada conta AWS tem um Catálogo de dados por região AWS.



## Dados subjacentes

Os dados subjacentes se referem aos dados de origem ou aos dados dentro dos data lakes para os quais as tabelas do Catálogo de dados apontam.

## Entidade principal

Uma entidade principal é um usuário ou um perfil (IAM) AWS Identity and Access Management ou um usuário do Diretório Ativo.

## Administrador do data lake

Um administrador de data lake é uma entidade principal que pode conceder a qualquer entidade principal (inclusive a si mesmo) qualquer permissão em qualquer recurso ou local de dados do Catálogo de dados. Designe um administrador de data lake como o primeiro usuário do Catálogo de dados. Esse usuário pode, então, conceder permissões mais granulares de recursos a outras entidades principais.

### Note

Os usuários administrativos do IAM – usuários com a política `AdministratorAccess` gerenciada por AWS – não são automaticamente administradores de data lake. Por exemplo, eles não podem conceder permissões do Lake Formation em objetos do catálogo, a menos que tenham recebido permissão para fazer isso. No entanto, eles podem usar o console ou a API do Lake Formation para se designarem como administradores do data lake.

Para obter informações sobre os recursos de um administrador de data lake, consulte [Permissões implícitas do Lake Formation](#). Para obter informações sobre como designar um usuário como administrador de data lake, consulte [Crie um administrador de data lake](#).

## Integrações de serviços da AWS com o Lake Formation

Você pode usar o Lake Formation para gerenciar as permissões de acesso no nível do banco de dados, da tabela e da coluna sobre os dados armazenados no Amazon S3. Depois que os dados forem registrados no Lake Formation, você poderá usar serviços analíticos da AWS como AWS Glue, Amazon Athena, Amazon Redshift Spectrum e Amazon EMR para consultar os dados. Os

serviços da AWS a seguir se integram com o AWS Lake Formation e honram as permissões do Lake Formation.

Serviço da AWS	Os detalhes da integração
<a href="#">AWS Glue</a>	<p>Tópico de referência: <a href="#">Usando AWS Lake Formation com AWS Glue</a></p> <p>O AWS Glue e o Lake Formation compartilham o mesmo Catálogo de Dados. Para operações de console (como visualizar uma lista de tabelas) e todas as operações de API, os usuários do AWS Glue podem acessar somente os bancos de dados e tabelas nos quais têm permissões do Lake Formation.</p>
<a href="#">Amazon Athena</a>	<p>Tópico de referência: <a href="#">Usando AWS Lake Formation com o Amazon Athena</a></p> <p>Usar o Lake Formation para permitir ou negar permissões para ler dados no Amazon S3. Quando os usuários do Amazon Athena selecionam o catálogo do AWS Glue no editor de consultas, eles podem consultar somente os bancos de dados, tabelas e colunas nos quais têm permissões do Lake Formation. Consultas usando manifestos não são aceitas.</p> <p>Atualmente, o Lake Formation não oferece suporte ao gerenciamento de permissões em operações de gravação como VACUUM, MERGE, UPDATE e OPTIMIZE em tabelas em formatos de tabela aberta.</p> <p>Além das entidades principais que se autenticam com o Athena pelo AWS Identity and Access Management (IAM), o Lake Formation oferece suporte aos usuários do Athena que se conectam pelo driver JDBC ou ODBC e se autenticam por SAML. Os provedores de SAML aceitos incluem o Okta e o Microsoft Active Directory Federation Service (AD FS).</p>
<a href="#">Amazon Redshift Spectrum</a>	<p>Tópico de referência: <a href="#">Usando AWS Lake Formation com o Amazon Redshift Spectrum</a></p> <p>Quando os usuários do Amazon Redshift criam um esquema externo em um banco de dados no AWS Glue Data Catalog, eles podem</p>

Serviço da AWS	Os detalhes da integração
	consultar somente as tabelas e colunas desse esquema nas quais tenham permissões do Lake Formation.
<a href="#">Edição Amazon QuickSight Enterprise</a>	Referência: <a href="#">Usando AWS Lake Formation com a Amazon QuickSight</a>  Quando um usuário do Amazon QuickSight Enterprise Edition consulta um conjunto de dados em um local do Amazon S3, o usuário deve ter a permissão Lake SELECT Formation nos dados.
<a href="#">Amazon EMR</a>	Referência: <a href="#">Usando AWS Lake Formation com o Amazon EMR</a>  Você pode integrar as permissões do Lake Formation ao criar um cluster do Amazon EMR com uma função de runtime.  Uma função de runtime é um perfil do IAM que você associa a trabalhos ou consultas do Amazon EMR e, em seguida, o Amazon EMR usa essa função para acessar recursos da AWS.

O Lake Formation também trabalha com o [AWS Key Management Service](#)(AWS KMS) para permitir que você configure com mais facilidade esses serviços integrados para criptografar e descriptografar dados em locais do Amazon Simple Storage Service (Amazon S3).

## Recursos adicionais do Lake Formation

### Tópicos

- [Blogs](#)
- [Palestras técnicas e webinars](#)
- [Arquitetura moderna](#)
- [Recursos de data mesh](#)
- [Guias práticas recomendadas](#)

### Blogs

- [AWS Lake Formation — Análise do ano de 2022](#)

- [Arquitetura de dados moderna e multirregional altamente resiliente](#)
- [Compartilhamento entre contas usando LF-tags para direcionar entidades principais do IAM](#)
- [Painel de inventário de permissões do Lake Formation](#)
- [Data mesh orientada por eventos](#)

## Palestras técnicas e webinars

- re:Invent 2020 — [Data lakes: Crie, proteja e compartilhe com facilidade com AWS Lake Formation](#)
- re:Invent 2022 — [Construindo e operando um datalake no Amazon S3](#)
- AWS Summit SF 2022 — [Entendendo e alcançando uma arquitetura de dados moderna](#)
- AWS Summit ATL 2022 — [Data lakes modernos com AWS Lake Formation, Amazon Redshift e AWS Glue](#)
- AWS Summit ANZ 2022 — [Data lakes, lake houses e data mesh: o quê, por quê e como?](#)
- AWS Palestras técnicas online — [Simplificando as permissões e a governança em seu data lake](#)

## Arquitetura moderna

- [Padrões de arquitetura moderna](#)

## Recursos de data mesh

- [Crie uma arquitetura de dados moderna e um padrão de data mesh em grande escala usando controle de acesso AWS Lake Formation baseado em tags](#)
- [Como o JPMorgan Chase construiu uma arquitetura de data mesh para gerar valor significativo e aprimorar sua plataforma de dados corporativos](#)
- [Crie uma data mesh no AWS](#)

## Guias práticas recomendadas

- [Guias práticas recomendadas do AWS Lake Formation](#)

# Introdução ao Lake Formation

Recomendamos que você inicie por estas seções:

- [AWS Lake Formation: como funciona](#): aprenda sobre a terminologia essencial e como os vários componentes interagem.
- [Introdução ao Lake Formation](#): obtenha informações sobre os pré-requisitos e conclua tarefas importantes de configuração.
- [Tutoriais](#)— Siga os step-by-step tutoriais para aprender a usar o Lake Formation.
- [Segurança em AWS Lake Formation](#): entenda como você pode ajudar a proteger o acesso a dados no Lake Formation.

# Introdução ao Lake Formation

Se você ainda não se cadastrou na AWS ou precisa de ajuda para começar, conclua as tarefas a seguir.

## Tópicos

- [Concluir as tarefas iniciais de configuração do AWS](#)
- [Configurar o AWS Lake Formation](#)
- [Atualizar as permissões AWS Glue de dados para o modelo AWS Lake Formation](#)
- [AWS Lake Formation e endpoint da VPC de interface \(AWS PrivateLink\)](#)

## Concluir as tarefas iniciais de configuração do AWS

Para usar o AWS Lake Formation, antes é necessário concluir as seguintes tarefas:

## Tópicos

- [Inscrever-se para uma Conta da AWS](#)
- [Criar um usuário administrativo](#)
- [Conceder acesso programático](#)

## Inscrever-se para uma Conta da AWS

Se você ainda não tem uma Conta da AWS, siga as etapas a seguir para criar uma.

### Como cadastrar-se para uma Conta da AWS

1. Abra <https://portal.aws.amazon.com/billing/signup>.
2. Siga as instruções on-line.

Parte do procedimento de cadastramento envolve o recebimento de uma chamada telefônica e a inserção de um código de verificação no teclado do telefone.

Quando você se cadastra para uma Conta da AWS, um Usuário raiz da conta da AWS é criado. O usuário raiz tem acesso a todos os Serviços da AWS e atributos na conta. Como prática recomendada de segurança, [atribua acesso administrativo a um usuário administrativo](#) e utilize somente o usuário raiz para executar as [tarefas que exigem acesso do usuário raiz](#).

A AWS envia um e-mail de confirmação depois que o processo de cadastramento é concluído. A qualquer momento, é possível visualizar as atividades da conta atual e gerenciar sua conta acessando <https://aws.amazon.com/> e selecionando Minha conta.

## Criar um usuário administrativo

Depois de se inscrever em uma Conta da AWS, proteja seu Usuário raiz da conta da AWS, habilite o AWS IAM Identity Center e crie um usuário administrativo para não usar o usuário raiz em tarefas cotidianas.

### Proteção da Usuário raiz da conta da AWS

1. Faça login no [AWS Management Console](#) como o proprietário da conta ao escolher a opção Usuário raiz e inserir o endereço de e-mail da Conta da AWS. Na próxima página, insira a senha.

Para obter ajuda ao fazer login usando o usuário raiz, consulte [Fazer login como usuário raiz](#) no Guia do usuário do Início de Sessão da AWS.

2. Ative a autenticação multifator (MFA) para o usuário raiz.c

Para obter instruções, consulte [Ativar um dispositivo MFA virtual para o usuário raiz da Conta da AWS \(console\)](#) no Guia do usuário do IAM.

### Criar um usuário administrativo

1. Habilitar o IAM Identity Center.

Para obter instruções, consulte [Enabling AWS IAM Identity Center](#) no Manual do Usuário do AWS IAM Identity Center.

2. No Centro de Identidade do IAM, conceda acesso administrativo a um usuário administrativo.

Para ver um tutorial sobre como usar o Diretório do Centro de Identidade do IAM como fonte de identidade, consulte [Configure user access with the default Diretório do Centro de Identidade do IAM](#) no Manual do Usuário do AWS IAM Identity Center.

### Login como usuário administrativo

- Para fazer login com o usuário do Centro de Identidade do IAM, utilize o URL de login enviado ao seu endereço de e-mail quando você criou o usuário do Centro do Usuário do IAM.

Para obter ajuda com o login utilizando um usuário do Centro de Identidade do IAM, consulte [Fazer login no portal de acesso da AWS](#), no Guia do usuário do Início de Sessão da AWS.

## Conceder acesso programático

Os usuários precisam de acesso programático se quiserem interagir com a AWS de fora do AWS Management Console. A forma de conceder acesso programático depende do tipo de usuário que está acessando a AWS.

Para conceder acesso programático aos usuários, selecione uma das seguintes opções:

Qual usuário precisa de acesso programático?	Para	Por
Identificação da força de trabalho  (Usuários gerenciados no Centro de Identidade do IAM)	Use credenciais temporárias para assinar solicitações programáticas para a AWS CLI, os SDKs da AWS ou as APIs da AWS.	Siga as instruções da interface que deseja utilizar. <ul style="list-style-type: none"> <li>• Para a AWS CLI, consulte <a href="#">Configuração da AWS CLI para usar o AWS IAM Identity Center</a> no AWS Command Line Interface Guia do usuário da .</li> <li>• Para os SDKs da AWS, ferramentas e APIs da AWS, consulte <a href="#">Autenticação do Centro de Identidade do IAM</a> no Guia de referência de ferramentas e SDKs da AWS.</li> </ul>
IAM	Use credenciais temporárias para assinar solicitações programáticas para a AWS CLI, os SDKs da AWS ou as APIs da AWS.	Siga as instruções em <a href="#">Como usar credenciais temporárias com recursos da AWS</a> no Guia do usuário do IAM.



Qual usuário precisa de acesso programático?	Para	Por
IAM	(Não recomendado) Use credenciais de longo prazo para assinar solicitações programáticas para a AWS CLI, os SDKs da AWS ou as APIs da AWS.	Siga as instruções da interface que deseja utilizar. <ul style="list-style-type: none"> <li>• Para a AWS CLI, consulte <a href="#">Autenticação usando as credenciais de usuário do IAM</a> no Guia do usuário da AWS Command Line Interface.</li> <li>• Para as ferramentas e SDKs da AWS, consulte <a href="#">Autenticação usando as credenciais de longo prazo</a> no Guia de referência de ferramentas e SDKs da AWS.</li> <li>• Para as APIs da AWS, consulte <a href="#">Gerenciamento de chaves de acesso de usuários do IAM</a> no Guia do usuário do IAM.</li> </ul>

## Configurar o AWS Lake Formation

As seções a seguir fornecem informações sobre a configuração do Lake Formation pela primeira vez. Nem todos os tópicos desta seção são necessários para começar a usar o Lake Formation. Você pode usar as instruções para configurar o modelo de permissões do Lake Formation para gerenciar seus objetos AWS Glue Data Catalog e local de dados existentes no Amazon Simple Storage Service (Amazon S3).

1. [Crie um administrador de data lake](#)
2. [Altere o modelo de permissão padrão ou use o modo de acesso híbrido](#)
3. [the section called “Como configurar um local no Amazon S3 para o data lake”](#)

4. [the section called “Atribuir permissões aos usuários do Lake Formation”](#)
5. [the section called “Integrar o Centro de Identidade do IAM”](#)
6. [the section called “\(Opcional\) Configurações externas de filtragem de dados”](#)
7. [the section called “\(Opcional\) Conceder acesso à chave de criptografia do Catálogo de Dados”](#)
8. [\(Opcional\) Crie uma função do IAM para fluxos de trabalho](#)

Esta seção mostra como configurar os recursos do Lake Formation de duas maneiras diferentes:

- Usando um modelo do AWS CloudFormation
- Usando o console Lake Formation

Para configurar o Lake Formation usando o console AWS acesse [Crie um administrador de data lake](#).

## Configurar recursos do Lake Formation usando o modelo AWS CloudFormation

### Note

A AWS CloudFormation pilha executa as etapas 1 a 6 acima, exceto as etapas 2 e 5. Execute [Altere o modelo de permissão padrão ou use o modo de acesso híbrido](#) e [the section called “Integrar o Centro de Identidade do IAM”](#) manualmente a partir do console Lake Formation.

1. Faça login no console do AWS CloudFormation em <https://console.aws.amazon.com/cloudformation> como administrador do IAM na região Leste dos EUA (Norte da Virgínia).
2. Selecione [Iniciar Pilha](#).
3. Na página Criar pilha, selecione Avançar.
4. Digite um Nome de pilha.
5. Para DatalakeAdminName e DatalakeAdminPassword, insira seu nome de usuário e senha para o usuário administrador do data lake.
6. Para DatalakeUser1Name e DatalakeUser1Password, digite seu nome de usuário e senha para o usuário do data lake analyst.

7. Para `DataLakeBucketName`, insira o nome do novo bucket que será criado.
8. Selecione Avançar.
9. Na página seguinte, selecione Avançar.
10. Revise os detalhes na página final e selecione Eu reconheço que AWS CloudFormation pode criar recursos do IAM.
11. Selecione Criar.

A criação da pilha pode levar até dois minutos.

## Limpar recursos

Se você quiser limpar os recursos da pilha AWS CloudFormation:

1. Cancele o registro do bucket do Amazon S3 que sua pilha criou e registrou como um local de data lake.
2. Excluir a pilha AWS CloudFormation. Isso excluirá todos os recursos criados pela pilha.

## Crie um administrador de data lake

Inicialmente, os administradores do Data Lake são os únicos usuários ou perfis AWS Identity and Access Management (IAM) que podem conceder permissões do Lake Formation sobre local de dados e recursos do Catálogo de Dados a qualquer entidade principal (inclusive a si mesmo). Para obter mais informações sobre os recursos do administrador de data lake, consulte [Permissões implícitas do Lake Formation](#). Por padrão, o Lake Formation permite criar até 30 administradores de data lake.

Você pode criar um administrador de data lake usando o console do Lake Formation ou a operação `PutDataLakeSettings` da API do Lake Formation.

As permissões a seguir são necessárias para criar um administrador de data lake. O usuário `Administrator` tem essas permissões implicitamente.

- `lakeformation:PutDataLakeSettings`
- `lakeformation:GetDataLakeSettings`

Se você conceder a política `AWSLakeFormationDataAdmin` a um usuário, esse usuário não poderá criar usuários administradores adicionais do Lake Formation.

## Como criar um administrador de data lake (console)

1. Se o usuário que será administrador do data lake ainda não existir, use o console do IAM para criá-lo. Caso contrário, selecione um usuário existente que será o administrador do data lake.

### Note

Recomendamos que você não selecione um usuário administrativo do IAM (usuário com a política `AdministratorAccess` AWS gerenciada) para ser o administrador do data lake.

Anexe as seguintes políticas AWS gerenciadas ao usuário:

Políticas	Obrigatório?	Observações
<code>AWSLakeFormationDataAdmin</code>	Obrigatório	Permissões básicas do administrador do data lake. Essa política gerenciada AWS contém uma restrição explícita para a operação da API Lake Formation, <code>PutDataLakeSetting</code> que impede que os usuários criem novos administradores de data lake.
<code>AWSGlueConsoleFullAccess</code> , <code>CloudWatchLogsReadOnlyAccess</code>	Opcional	Anexe essas políticas se o administrador do data lake estiver solucionando problemas de fluxos de trabalho criados a partir dos esquemas do Lake Formation. Essas políticas permitem que o administrador do data lake visualize as informações de solução de problemas no console AWS Glue e no console Amazon CloudWatch Logs. Para obter informações sobre fluxos de trabalho, consulte <a href="#">the section called “Importação de dados usando fluxos de trabalho”</a> .

Políticas	Obrigatório?	Observações
AWSLakeFormationCrossAccountManager	Opcional	Anexe esta política para permitir que o administrador do data lake conceda e revogue permissões entre contas nos recursos do Catálogo de Dados. Para ter mais informações, consulte <a href="#">Compartilhamento de dados entre contas no Lake Formation</a> .
AmazonAthenaFullAccess	Opcional	Anexe essa política se o administrador do data lake estiver executando consultas em Amazon Athena.

2. Anexe a seguinte política em linha, que concede ao administrador do data lake permissão para criar o perfil vinculado ao serviço do Lake Formation. O nome sugerido para a política é LakeFormationSLR.

O perfil vinculado ao serviço permite que o administrador do data lake registre com mais facilidade o local do Amazon S3 no Lake Formation. Para obter mais informações sobre o perfil vinculado ao serviço Lake Formation, consulte [the section called “Usar perfis vinculados ao serviço”](#).

#### Important

Em toda a política a seguir, substitua *<account-id>* por um número de conta AWS válido.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": "iam:CreateServiceLinkedRole",
      "Resource": "*",
      "Condition": {
        "StringEquals": {
          "iam:AWSServiceName": "lakeformation.amazonaws.com"
        }
      }
    }
  ]
}
```

```

    }
  },
  {
    "Effect": "Allow",
    "Action": [
      "iam:PutRolePolicy"
    ],
    "Resource": "arn:aws:iam::<account-id>:role/aws-service-role/
lakeformation.amazonaws.com/AWSServiceRoleForLakeFormationDataAccess"
  }
]
}

```

- (Opcional) Anexe a seguinte política em linha ao usuário PassRole. Essa política permite que o administrador do data lake crie e execute fluxos de trabalho. A permissão `iam:PassRole` permite que o fluxo de trabalho assumo o perfil `LakeFormationWorkflowRole` para criar crawlers e trabalhos, e anexe o perfil aos rastreadores e trabalhos criados. O nome sugerido para a política é `UserPassRole`.

#### Important

Substitua `<account-id>` por um número de conta AWS válido.

```

{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "PassRolePermissions",
      "Effect": "Allow",
      "Action": [
        "iam:PassRole"
      ],
      "Resource": [
        "arn:aws:iam::<account-id>:role/LakeFormationWorkflowRole"
      ]
    }
  ]
}

```

- (Opcional) Anexe essa política adicional em linha se sua conta estiver concedendo ou recebendo permissões entre contas do Lake Formation. Essa política permite que o administrador do data lake visualize e aceite AWS Resource Access Manager (AWS RAM) convites de compartilhamento de recursos. Além disso, para administradores de data lake na conta de gerenciamento AWS Organizations, a política inclui uma autorização para permitir concessões entre contas para organizações. Para ter mais informações, consulte [Compartilhamento de dados entre contas no Lake Formation](#).

O nome sugerido para a política é RAMAccess.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "ram:AcceptResourceShareInvitation",
        "ram:RejectResourceShareInvitation",
        "ec2:DescribeAvailabilityZones",
        "ram:EnableSharingWithAwsOrganization"
      ],
      "Resource": "*"
    }
  ]
}
```

- Abra o console AWS Lake Formation em <https://console.aws.amazon.com/lakeformation/> e faça login como o usuário administrador que você criou em [Criar um usuário administrativo](#) ou como um usuário com política AWS gerenciada pelo usuário AdministratorAccess.
- Se a janela de Boas-vindas ao Lake Formation for exibida, escolha o usuário do IAM que você criou ou selecionou na Etapa 1 e, em seguida, escolha Começa.
- Se você não vir a janela de Boas-vindas ao Lake Formation, execute as etapas a seguir para configurar um administrador do Lake Formation.
  - No painel de navegação, em Administradores, selecione Perfis e tarefas administrativas. Na seção Administradores do Data Lake da página do console, selecione Adicionar.
  - Na caixa de diálogo Adicionar administradores, em Tipo de Acesso, selecione Administrador do Data Lake.

- c. Para Usuários e perfis do IAM, selecione o usuário do IAM que você criou ou selecionou na Etapa 1 e, em seguida, selecione Salvar.

## Altere o modelo de permissão padrão ou use o modo de acesso híbrido

O Lake Formation começa com as configurações "Usar somente controle de acesso IAM" ativadas para compatibilidade com o comportamento existente em AWS Glue Data Catalog. Essas configurações permitem que você gerencie o acesso aos seus dados no data lake e seus metadados por meio de políticas do IAM e políticas de bucket do Amazon S3.

Para facilitar a transição das permissões do data lake de um modelo do IAM e do Amazon S3 para as permissões do Lake Formation, recomendamos que você use o modo de acesso híbrido para o Catálogo de Dados. Com o modo de acesso híbrido, você tem um caminho incremental em que pode habilitar as permissões do Lake Formation para um conjunto específico de usuários sem interromper outros usuários ou workloads existentes.

Para ter mais informações, consulte [Modo de acesso híbrido](#).

Desative as configurações padrão para mover todos os usuários existentes de uma tabela para o Lake Formation em uma única etapa.

### Important

Se você tiver bancos de dados AWS Glue Data Catalog e tabelas existentes, não siga as instruções descritas nesta seção. Em vez disso, siga as instruções em [the section called "Atualizar as permissões de dados AWS Glue para o modelo do Lake Formation"](#).

### Warning

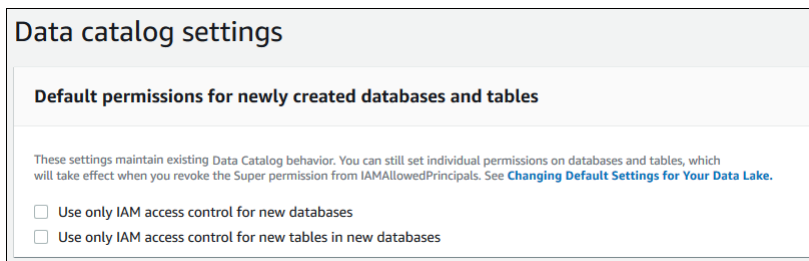
Se você tiver uma automação que cria bancos de dados e tabelas no Catálogo de Dados, as etapas a seguir podem fazer com que as tarefas de automação e extração, transformação e carregamento (ETL) posteriores falhem. prossiga somente depois de modificar seus processos existentes ou conceder permissões explícitas do Lake Formation às entidades principais necessárias. Para obter informações sobre as permissões do Lake Formation, consulte [the section called "Referência de permissões do Lake Formation"](#).



## Como alterar as configurações padrão do Catálogo de Dados

1. Continue no console do Lake Formation em <https://console.aws.amazon.com/lakeformation/>. Certifique-se de estar conectado como o usuário administrador que você criou em [Criar um usuário administrativo](#) ou como um usuário com a política AdministratorAccess gerenciada por AWS.
2. Modificar as configurações do catálogo de dados:

- a. No painel de navegação, em Administração, selecione Configurações do Catálogo de Dados.
- b. Desmarque as duas caixas de seleção e selecione Salvar.



3. Revogar a permissão IAMAllowedPrincipals para criadores de banco de dados.
  - a. No painel de navegação, em Administração, selecione Perfis e tarefas administrativas.
  - b. Na página do console de Perfis e tarefas administrativas, na seção Criadores de banco de dados, selecione o grupo IAMAllowedPrincipals, e selecione Revogar.

A caixa de diálogo Revogar permissões é exibida, mostrando que IAMAllowedPrincipals tem a permissão para Criar banco de dados.

- c. Selecione Revogar.

## Atribuir permissões aos usuários do Lake Formation

Crie um usuário para ter acesso ao data lake em AWS Lake Formation. Esse usuário tem as permissões de privilégio mínimo para consultar o data lake.

Para obter mais informações sobre a criação de usuários ou grupos, consulte [Identidades do IAM](#) no Guia do Usuário do IAM.

## Como anexar permissões a um usuário não administrador para acessar os dados do Lake Formation

1. Abra o console do IAM em <https://console.aws.amazon.com/iam> e faça login como um usuário administrador que você criou em [Criar um usuário administrativo](#) ou como usuário com a política AdministratorAccess gerenciada por AWS.
2. Selecione Usuários ou Grupos de usuários.
3. Na lista, selecione o nome do usuário ou do grupo ao qual deseja incorporar uma política.

Selecione Permissões.

4. Selecione Adicionar permissões e selecione Anexar políticas diretamente. Digite Athena no campo de texto Políticas de filtro. Na lista de resultados, marque a caixa de seleção AmazonAthenaFullAccess.
5. Selecione o botão Criar política. Na página Criar política, selecione a guia JSON. Copie e cole o código a seguir no editor de políticas.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "lakeformation:GetDataAccess",
        "glue:GetTable",
        "glue:GetTables",
        "glue:SearchTables",
        "glue:GetDatabase",
        "glue:GetDatabases",
        "glue:GetPartitions",
        "lakeformation:GetResourceLFTags",
        "lakeformation:ListLFTags",
        "lakeformation:GetLFTag",
        "lakeformation:SearchTablesByLFTags",
        "lakeformation:SearchDatabasesByLFTags"
      ],
      "Resource": "*"
    }
  ]
}
```

6. Selecione o botão Avançar na parte inferior até ver a página Revisar política. Digite um nome para a política, por exemplo, `DataLakeUserBasic`. Selecione Criar política e feche a guia Políticas ou a janela do navegador.

## Como configurar um local no Amazon S3 para o data lake

Para usar o Lake Formation para gerenciar e proteger os dados em seu data lake, você deve primeiro registrar um local no Amazon S3. Quando você registra um local, esse caminho do Amazon S3 e todas as pastas sob esse caminho são registradas, o que permite que o Lake Formation aplique permissões de nível de armazenamento. Quando o usuário solicita dados de um mecanismo integrado como o Amazon Athena, o Lake Formation fornece acesso aos dados em vez de usar as permissões do usuário.

Ao registrar um local, você especifica um perfil do IAM que concede permissões de leitura/gravação neste local. O Lake Formation assume esse perfil ao fornecer credenciais temporárias para serviços integrados do AWS que solicitam acesso aos dados no local registrado no Amazon S3. É possível especificar o perfil vinculado ao serviço (SLR) do Lake Formation ou criar seu próprio perfil.

Use um perfil personalizado nas seguintes situações:

- Você planeja publicar métricas no Amazon CloudWatch Logs. A função definida pelo usuário deve incluir uma política para adicionar registros em CloudWatch registros e publicar métricas, além das permissões de SLR. Para obter um exemplo de política em linha que concede CloudWatch as permissões necessárias, consulte [Requisitos para funções usadas para registrar locais](#).
- O local do Amazon S3 existe em uma conta diferente. Para obter detalhes, consulte [the section called “Registrando uma localização do Amazon S3 em outra conta AWS”](#).
- O local do Amazon S3 contém dados criptografados com um Chave gerenciada pela AWS. Para obter mais detalhes, consulte [Registrando uma localização criptografada do Amazon S3](#) e [Registrando uma localização criptografada do Amazon S3 em todas as contas AWS](#).
- Você planeja acessar o local do Amazon S3 usando o Amazon EMR. Para obter mais informações sobre os requisitos de perfil, consulte [Perfis do IAM para o Lake Formation](#) no Guia de Gerenciamento do Amazon EMR.

O perfil que você escolher deve ter as permissões necessárias, conforme descrito em [Requisitos para funções usadas para registrar locais](#). Para obter instruções sobre como registrar um local no Amazon S3, consulte [Adicionar uma localização do Amazon S3 ao seu data lake](#).

## (Opcional) Configurações externas de filtragem de dados

Se você pretende analisar e processar dados em seu data lake usando mecanismos de consulta de terceiros, você deve optar por permitir que mecanismos externos acessem dados gerenciados pelo Lake Formation. Se você não optar por participar, mecanismos externos não poderão acessar dados em locais do Amazon S3 registrados no Lake Formation.

O Lake Formation oferece suporte a permissões em nível de coluna para restringir o acesso a colunas específicas em uma tabela. Serviços analíticos integrados Amazon Athena, como Amazon Redshift Spectrum e Amazon EMR, recuperam metadados de tabela não filtrados do AWS Glue Data Catalog. A filtragem real das colunas nas respostas da consulta é de responsabilidade do serviço integrado. É responsabilidade dos administradores terceirizados lidar adequadamente com as permissões para evitar o acesso não autorizado aos dados.

Para optar por permitir que mecanismos de terceiros acessem e filtrem dados (console)

1. Continue no console do Lake Formation em <https://console.aws.amazon.com/lakeformation/>. Certifique-se de estar conectado como entidade principal com a permissão do IAM na operação da API PutDataLakeSettings do Lake Formation. O usuário administrador do IAM que você criou em [Inscrever-se para uma Conta da AWS](#) tem essa permissão.
2. No painel de navegação, em Administração, selecione Configurações de integração de aplicativos.
3. Na página Configurações de integração de aplicativos, faça o seguinte:
  - a. Marque a caixa Permitir que mecanismos externos filtrem dados em locais do Amazon S3 registrados no Lake Formation.
  - b. Digite os Valores da tag de sessão definidos para mecanismos de terceiros.
  - c. Para IDs de conta AWS, digite os IDs de conta de onde mecanismos de terceiros podem acessar locais registrados no Lake Formation. Pressione Enter após cada ID da conta.
  - d. Selecione Salvar.

Para permitir que mecanismos externos acessem dados sem a validação da tag de sessão, consulte [Integração de aplicativos para acesso total à tabela](#)

## (Opcional) Conceder acesso à chave de criptografia do Catálogo de Dados

Se AWS Glue Data Catalog for criptografado, conceda permissões AWS Identity and Access Management (IAM) na chave AWS KMS a qualquer entidade principal que precise conceder permissões do Lake Formation nos bancos de dados e tabelas do Data Catalog.

Para obter mais informações, consulte o Guia do desenvolvedor do AWS Key Management Service.

## (Opcional) Crie uma função do IAM para fluxos de trabalho

Com AWS Lake Formation, você pode importar seus dados usando fluxos de trabalho executados por AWS Glue crawlers. Um fluxo de trabalho define a fonte de dados e o cronograma para importar dados para o seu data lake. Você pode definir facilmente fluxos de trabalho usando os esquemas ou modelos fornecidos pelo Lake Formation.

Ao criar um fluxo de trabalho, você deve atribuir a ele um perfil do AWS Identity and Access Management (IAM) que conceda ao Lake Formation as permissões necessárias para ingerir os dados.

O procedimento a seguir requer familiaridade com o IAM.

Para criar um perfil do IAM para fluxos de trabalho

1. Abra o console do IAM em <https://console.aws.amazon.com/iam> e faça login como o usuário administrador que você criou em [Criar um usuário administrativo](#) ou como usuário com a política AdministratorAccess gerenciada por AWS.
2. No painel de navegação, selecione Perfis e depois Criar perfil.
3. Na página Criar perfil, selecione Serviço AWS e, em seguida, selecione Glue. Selecione Avançar.
4. Na página Adicionar permissões, pesquise a política AWSGlueServiceRolegerenciada e marque a caixa de seleção ao lado do nome da política na lista. Em seguida, conclua o assistente de Criação de perfil, nomeando o perfilLakeFormationWorkflowRole. Para finalizar, selecione Criar perfil.
5. Na página Perfis, procure seu perfil por LakeFormationWorkflowRole e selecione o perfil.
6. Na página Resumo do perfil, na guia Permissões, selecione Adicionar política em linha. Na tela Criar política, navegue até a guia JSON e adicione a seguinte política em linha. O nome sugerido para a política é LakeFormationWorkflow.

**⚠ Important**

Na política a seguir, substitua `<account-id>` por um número Conta da AWS válido.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "lakeformation:GetDataAccess",
        "lakeformation:GrantPermissions"
      ],
      "Resource": "*"
    },
    {
      "Effect": "Allow",
      "Action": ["iam:PassRole"],
      "Resource": [
        "arn:aws:iam::<account-id>:role/LakeFormationWorkflowRole"
      ]
    }
  ]
}
```

Seguem breves descrições das permissões nesta política:

- `lakeformation:GetDataAccess` permite que trabalhos criados pelo fluxo de trabalho sejam gravados no local de destino.
  - `lakeformation:GrantPermissions` permite que o fluxo de trabalho conceda a `SELECT` permissão nas tabelas de destino.
  - `iam:PassRole` permite que o serviço assuma o perfil de `LakeFormationWorkflowRole` para criar crawlers e trabalhos (instâncias de fluxos de trabalho) e anexe o perfil aos crawlers e trabalhos criados.
7. Verifique se o perfil `LakeFormationWorkflowRole` tem duas políticas anexadas.
  8. Se você estiver ingerindo dados fora do local do data lake, adicione uma política em linha que conceda permissões para ler os dados de origem.

# Atualizar as permissões AWS Glue de dados para o modelo AWS Lake Formation

Permissões AWS Lake Formation garantem controle de acesso para dados em seu data lake. Você pode usar o modelo de permissões do Lake Formation para gerenciar seus locais de dados e objetos do AWS Glue Data Catalog existentes no Amazon Simple Storage Service (Amazon S3).

O modelo de permissões do Lake Formation usa permissões granulares do AWS Identity and Access Management (IAM) para acesso ao serviço de API. Ele restringe os dados que seus usuários e esses serviços podem acessar pela funcionalidade do Lake Formation. Em comparação, o modelo do AWS Glue concede acesso aos dados por meio de permissões IAM de [controle de acesso refinadas](#). Para fazer a troca, siga as etapas deste guia.

Para obter mais informações, consulte [Visão geral das permissões do Lake Formation](#).

## Tópicos

- [Atualizar as permissões de dados para o modelo Lake Formation](#)
- [Etapa 1: listar as permissões existentes dos usuários e das funções](#)
- [Etapa 2: configurar permissões equivalentes do Lake Formation](#)
- [Etapa 3: conceder aos usuários permissões do IAM para usar o Lake Formation](#)
- [Etapa 4: mude seus armazenamentos de dados para o modelo de permissões do Lake Formation](#)
- [Etapa 5: proteja os novos recursos do Catálogo de Dados](#)
- [Etapa 6: fornecer aos usuários uma nova política do IAM para acesso futuro ao data lake](#)
- [Etapa 7: limpar políticas do IAM existentes](#)

## Atualizar as permissões de dados para o modelo Lake Formation

Para manter a compatibilidade com versões anteriores do AWS Glue, por padrão, o AWS Lake Formation concede a permissão `Super` ao grupo `IAMAllowedPrincipals` em todos os recursos do AWS Glue existentes do Catálogo de Dados e concede a permissão `Super` em novos recursos do Catálogo de Dados se as configurações de controle de acesso usando apenas IAM estiverem ativadas. Isso efetivamente faz com que o acesso aos recursos do catálogo de dados e aos locais do Amazon S3 seja controlado exclusivamente por políticas do AWS Identity and Access Management (IAM). O grupo `IAMAllowedPrincipals` inclui todos os usuários e funções do IAM que têm permissão para acessar seus objetos do Catálogo de Dados por meio de suas políticas do IAM. A

permissão Super possibilita que uma entidade principal execute todas as operações suportadas do Lake Formation no banco de dados ou na tabela em que ela foi concedida.

Você pode começar a usar o Lake Formation para gerenciar o acesso aos seus dados registrando os locais dos recursos existentes do Catálogo de Dados no Lake Formation ou usando o modo de acesso híbrido. Ao registrar a localização do Amazon S3 no modo de acesso híbrido, você pode habilitar as permissões do Lake Formation optando por entidades principais para bancos de dados e tabelas nesse local.

Para facilitar a transição das permissões do data lake de um modelo do IAM e do Amazon S3 para as permissões do Lake Formation, recomendamos que você use o modo de acesso híbrido para o Catálogo de Dados. Com o modo de acesso híbrido, você tem um caminho incremental em que pode habilitar as permissões do Lake Formation para um conjunto específico de usuários sem interromper outros usuários ou workloads existentes.

Para obter mais informações, consulte [Modo de acesso híbrido](#).

Desative as configurações padrão do Catálogo de Dados para transferir todos os usuários existentes de uma tabela para o Lake Formation em uma única etapa.

Para começar a usar as permissões do Lake Formation com bancos de dados e tabelas do Catálogo de Dados do AWS Glue existentes, você deve fazer o seguinte:

1. Determine as permissões do IAM existentes dos seus usuários para cada banco de dados e tabela.
2. Replique essas permissões no Lake Formation.
3. Para cada local do Amazon S3 que contém dados:
  - a. Revogue a permissão Super do grupo `IAMAllowedPrincipals` em cada recurso do Catálogo de Dados que faça referência a esse local.
  - b. Registre o local com o Lake Formation.
4. Limpe as políticas IAM de controle de acesso finas existentes.

#### Important

Para adicionar novos usuários durante o processo de transição do seu catálogo de dados, você deve configurar permissões do AWS Glue granulares no IAM como antes. Você também deve replicar essas permissões no Lake Formation conforme descrito nesta seção. Se os novos usuários tiverem as políticas gerais do IAM descritas neste guia, eles poderão



listar quaisquer bancos de dados ou tabelas que tenham a permissão `Super` concedida para `IAMAllowedPrincipals`. Eles também podem visualizar os metadados desses recursos.

Siga as etapas desta seção para atualizar para o modelo de permissões do Lake Formation. Comece com [the section called “Etapa 1: listar permissões existentes”](#).

## Etapa 1: listar as permissões existentes dos usuários e das funções

Para começar a usar permissões do AWS Lake Formation com tabelas e bancos de dados do AWS Glue existentes, você deve primeiro determinar as permissões existentes dos seus usuários.

### Important

Antes de começar, realize as seguintes tarefas em [Introdução](#).

### Tópicos

- [Usando a operação da API](#)
- [Usar a AWS Management Console](#)
- [Usar o AWS CloudTrail](#)

## Usando a operação da API

Use a operação da API [ListPoliciesGrantingServiceAccess](#) do AWS Identity and Access Management (IAM) para determinar as políticas do IAM anexadas a cada entidade principal (usuário ou função). A partir das políticas retornadas nos resultados, você pode determinar as permissões do IAM que são concedidas à entidade principal. Você deve invocar a API para cada entidade principal separadamente.

### Example

O exemplo da AWS CLI a seguir retorna a política anexada a um usuário `glue_user1`.

```
aws iam list-policies-granting-service-access --arn arn:aws:iam::111122223333:user/glue_user1 --service-namespaces glue
```

O comando retorna resultados semelhantes ao seguinte.

```
{
  "PoliciesGrantingServiceAccess": [
    {
      "ServiceNamespace": "glue",
      "Policies": [
        {
          "PolicyType": "INLINE",
          "PolicyName": "GlueUserBasic",
          "EntityName": "glue_user1",
          "EntityType": "USER"
        },
        {
          "PolicyType": "MANAGED",
          "PolicyArn": "arn:aws:iam::aws:policy/AmazonAthenaFullAccess",
          "PolicyName": "AmazonAthenaFullAccess"
        }
      ]
    }
  ],
  "IsTruncated": false
}
```

## Usar a AWS Management Console

Você também pode ver essas informações no console do AWS Identity and Access Management (IAM), na guia [Acessar o advisor](#) na página [Resumo do usuário](#) ou da função:

1. Abra o console do IAM em <https://console.aws.amazon.com/iam/>.
2. No painel de navegação, selecione **Usuários** ou **Funções**.
3. Escolha um nome na lista para abrir a página **Resumo** e escolha a guia **Supervisor de Acesso**.
4. Inspeção cada uma das políticas para determinar a combinação de bancos de dados, tabelas e ações para as quais cada usuário tem permissões.

Lembre-se de inspecionar as funções, além dos usuários, durante esse processo, pois seus trabalhos de processamento de dados podem estar assumindo funções para acessar os dados.

## Usar o AWS CloudTrail

Outra forma de determinar suas permissões existentes é procurar no AWS CloudTrail por chamadas de API do AWS Glue, em que o campo `additionalEventData` dos logs contém uma entrada

`insufficientLakeFormationPermissions`. Essa entrada lista o banco de dados e a tabela nos quais o usuário precisa das permissões do Lake Formation para realizar a mesma ação.

Esses são logs de acesso a dados, portanto, não é garantido que produzam uma lista abrangente de usuários e suas permissões. Recomendamos escolher um intervalo de tempo amplo para capturar a maioria dos padrões de acesso aos dados de seus usuários, por exemplo, várias semanas ou meses.

Para obter mais informações, consulte [Visualizar eventos com o histórico de eventos CloudTrail](#) no Guia do Usuário do AWS CloudTrail.

Em seguida, você pode configurar as permissões do Lake Formation de acordo com as permissões do AWS Glue. Consulte [Etapa 2: configurar permissões equivalentes do Lake Formation](#).

## Etapa 2: configurar permissões equivalentes do Lake Formation

Usando as informações coletadas em [Etapa 1: listar as permissões existentes dos usuários e das funções](#), conceda permissões do AWS Lake Formation que correspondam às permissões do AWS Glue. Use qualquer um dos métodos a seguir para realizar as concessões:

- Você pode usar o console do Lake Formation ou a AWS CLI.

Consulte [the section called “Conceder e revogar permissões do Catálogo de dados”](#).

- Use as operações da API `GrantPermissions` ou `BatchGrantPermissions`.

Consulte [APIs de permissões](#).

Para obter mais informações, consulte [Visão geral das permissões do Lake Formation](#).

Depois de configurar as permissões do Lake Formation, siga para a [Etapa 3: conceder aos usuários permissões do IAM para usar o Lake Formation](#).

## Etapa 3: conceder aos usuários permissões do IAM para usar o Lake Formation

Para usar o modelo de permissões do AWS Lake Formation, as entidades principais devem ter permissões do AWS Identity and Access Management (IAM) nas APIs do Lake Formation.

Crie a política a seguir no IAM e a anexe a todos os usuários que precisam acessar seu data lake. Atribua o nome `LakeFormationDataAccess` à política.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "LakeFormationDataAccess",
      "Effect": "Allow",
      "Action": [
        "lakeformation:GetDataAccess"
      ],
      "Resource": "*"
    }
  ]
}
```

Em seguida, atualize para as permissões do Lake Formation, um local de dados por vez. Consulte [Etapa 4: mude seus armazenamentos de dados para o modelo de permissões do Lake Formation](#).

## Etapa 4: mude seus armazenamentos de dados para o modelo de permissões do Lake Formation

Atualize para as permissões do Lake Formation, um local de dados por vez. Para isso, repita essa seção inteira até ter registrado todos os caminhos do Amazon Simple Storage Service (Amazon S3) referenciados pelo catálogo de dados.

### Tópicos

- [Verifique permissões do Lake Formation](#)
- [Proteja os recursos existentes do Catálogo de Dados](#)
- [Ative as permissões do Lake Formation para sua localização no Amazon S3](#)

## Verifique permissões do Lake Formation

Antes de registrar um local, execute uma etapa de verificação para garantir que as entidades principais corretas tenham as permissões necessárias para o Lake Formation e que nenhuma permissão para o Lake Formation seja concedida às entidades principais que não deveriam tê-las. Usando a operação da API `GetEffectivePermissionsForPath` do Lake Formation, identifique os recursos do catálogo de dados que fazem referência à localização do Amazon S3, junto com as entidades principais que têm permissões sobre esses recursos.

A AWS CLI de exemplo a seguir retorna os bancos de dados e tabelas do catálogo de dados que fazem referência ao bucket do Amazon S3 products.

```
aws lakeformation get-effective-permissions-for-path --resource-arn
arn:aws:s3:::products --profile datalake_admin
```

Observe a opção `profile`. Recomendamos que você execute o comando como administrador do data lake.

A seguir está um trecho dos resultados retornados.

```
{
  "PermissionsWithGrantOption": [
    "SELECT"
  ],
  "Resource": {
    "TableWithColumns": {
      "Name": "inventory_product",
      "ColumnWildcard": {},
      "DatabaseName": "inventory"
    }
  },
  "Permissions": [
    "SELECT"
  ],
  "Principal": {
    "DataLakePrincipalIdentifier": "arn:aws:iam::111122223333:user/
datalake_user1",
    "DataLakePrincipalType": "IAM_USER"
  }
},...
```

#### Important

Se o catálogo de dados do AWS Glue estiver criptografado, `GetEffectivePermissionsForPath` retornará somente bancos de dados e tabelas que foram criados ou modificados após a disponibilidade geral do Lake Formation.

## Proteja os recursos existentes do Catálogo de Dados

Em seguida, revogue a permissão `Super` de cada tabela e banco de dados `IAMAllowedPrincipals` que você identificou para o local.

### Warning

Se você tiver uma automação que cria bancos de dados e tabelas no Catálogo de Dados, as etapas a seguir podem fazer com que as tarefas de automação e extração, transformação e carregamento (ETL) posteriores falhem. prossiga somente depois de modificar seus processos existentes ou conceder permissões explícitas do Lake Formation às entidades principais necessárias. Para obter informações sobre as permissões do Lake Formation, consulte [the section called “Referência de permissões do Lake Formation”](#).

Para revogar **Super** de **IAMAllowedPrincipals** em uma tabela

1. Abra o console do AWS Lake Formation em <https://console.aws.amazon.com/lakeformation/>. Faça login como administrador de data lake.
2. No painel de navegação, selecione Tabelas.
3. Na página Tabelas, selecione o botão de opção ao lado da tabela desejada.
4. No menu Ações, selecione Revogar.
5. Na caixa de diálogo Revogar permissões, na lista de usuários e funções do IAM, role para baixo até o título Grupo e escolha `IamAllowedPrincipals`.
6. Em Permissões da tabela, verifique se a opção `Super` está selecionada e escolha Revogar.

Para revogar **Super** de **IAMAllowedPrincipals** em um banco de dados

1. Abra o console do AWS Lake Formation em <https://console.aws.amazon.com/lakeformation/>. Faça login como administrador de data lake.
2. No painel de navegação, escolha Bancos de dados.
3. Na página Banco de dados, selecione o botão de opção ao lado do banco de dados desejado.
4. No menu Ações, escolha Editar.
5. Na página Editar banco de dados, desmarque Usar somente o controle de acesso do IAM para novas tabelas nesse banco de dados e escolha Salvar.

6. De volta à página Banco de dados, verifique se o banco de dados ainda está selecionado e, no menu Ações, escolha Revogar.
7. Na caixa de diálogo Revogar permissões, na lista de usuários e funções do IAM, role para baixo até o título Grupo e escolha lamAllowedPrincipals.
8. Em Permissões de banco de dados, verifique se a opção Super está selecionada e escolha Revogar.

## Ative as permissões do Lake Formation para sua localização no Amazon S3

Em seguida, registre o local do Amazon S3 com o Lake Formation. Para isso, você pode usar o processo descrito em [Adicionar uma localização do Amazon S3 ao seu data lake](#). Ou utilize a operação da API RegisterResource, conforme descrito em [APIs de fornecimento de credenciais](#).

### Note

Se o local pai estiver registrado, você não precisa registrar locais filho.

Depois de concluir essas etapas e testar se seus usuários podem acessar seus dados, você atualizou com sucesso para as permissões do Lake Formation. Continue na próxima etapa, [Etapa 5: proteja os novos recursos do Catálogo de Dados](#).

## Etapa 5: proteja os novos recursos do Catálogo de Dados

Em seguida, assegure todos os novos recursos do Catálogo de Dados alterando as configurações padrão do Catálogo de Dados. Desative as opções de usar somente o controle de acesso do AWS Identity and Access Management (IAM) para novos bancos de dados e tabelas.

### Warning

Se você tiver uma automação que cria bancos de dados e tabelas no Catálogo de Dados, as etapas a seguir podem fazer com que as tarefas de automação e extração, transformação e carregamento (ETL) posteriores falhem. Prossiga somente depois de modificar seus processos existentes ou conceder permissões explícitas do Lake Formation às entidades principais necessárias. Para obter informações sobre as permissões do Lake Formation, consulte [the section called “Referência de permissões do Lake Formation”](#).

## Como alterar as configurações do Catálogo de Dados padrão

1. Abra o console do AWS Lake Formation em <https://console.aws.amazon.com/lakeformation/>. Faça login como usuário administrativo do IAM (o usuário Administrator ou outro usuário com a política AdministratorAccess AWS gerenciada).
2. No painel de navegação, selecione Configurações.
3. Na página Configurações do catálogo de dados, desmarque as duas caixas de seleção e escolha Salvar.

A próxima etapa é conceder aos usuários acesso a bancos de dados ou tabelas adicionais no futuro. Consulte [Etapa 6: fornecer aos usuários uma nova política do IAM para acesso futuro ao data lake](#).

## Etapa 6: fornecer aos usuários uma nova política do IAM para acesso futuro ao data lake

Para conceder aos seus usuários acesso a bancos de dados ou tabelas adicionais do Data Catalog no futuro, você deve fornecer a eles a política em linha de granularidade geral do AWS Identity and Access Management (IAM) que segue. Atribua o nome `GlueFullReadAccess` à política.

### Important

Se você anexar essa política a um usuário antes da revogação de `Super` de `IAMAllowedPrincipals` em cada banco de dados e tabela em seu Catálogo de Dados, esse usuário poderá visualizar todos os metadados de qualquer recurso ao qual `Super` tenha sido concedido a `IAMAllowedPrincipals`.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "GlueFullReadAccess",
      "Effect": "Allow",
      "Action": [
        "lakeformation:GetDataAccess",
        "glue:GetTable",
        "glue:GetTables",
        "glue:SearchTables",

```



```
        "glue:GetDatabase",
        "glue:GetDatabases",
        "glue:GetPartitions"
    ],
    "Resource": "*"
}
]
```

### Note

As políticas em linha designadas nesta etapa e nas etapas anteriores contêm permissões mínimas do IAM. Para políticas sugeridas para administradores de data lake, analistas de dados e outras pessoas, consulte [the section called “Referência de personas e permissões do IAM do Lake Formation”](#).

Em seguida, prossiga para a [Etapa 7: limpar políticas do IAM existentes](#).

## Etapa 7: limpar políticas do IAM existentes

Depois de configurar as permissões do AWS Lake Formation e criar e anexar as políticas gerais de controle de acesso do AWS Identity and Access Management (IAM), conclua a seguinte etapa final:

- Remova dos usuários, grupos e funções as antigas [políticas refinadas do IAM de controle de acesso](#) que você replicou no Lake Formation.

Ao fazer isso, você garante que essas entidades principais não tenham mais acesso direto aos dados no Amazon Simple Storage Service (Amazon S3). Em seguida, você pode gerenciar o acesso dessas entidades principais ao data lake inteiramente por meio do Lake Formation.

## AWS Lake Formation e endpoint da VPC de interface (AWS PrivateLink)

A Amazon VPC é um produto da AWS que pode ser utilizado para iniciar os recursos da AWS em uma rede virtual definida por você. Com a VPC, você tem controle sobre as configurações de rede, como o intervalo de endereços IP, sub-redes, tabelas de rotas e gateways de rede.

Se você usar o Amazon Virtual Private Cloud (Amazon VPC) para hospedar seus recursos AWS, poderá estabelecer uma conexão privada entre sua VPC e o Lake Formation. Você usa essa conexão para que o Lake Formation possa se comunicar com os recursos em sua VPC sem passar pela Internet pública.

É possível estabelecer uma conexão privada entre a VPC e o AWS Lake Formation criando um endpoint da VPC de interface. Os endpoints de interface são alimentados por [AWS PrivateLink](#), uma tecnologia que permite acessar APIs do Lake Formation de forma privada sem um gateway da internet, dispositivo NAT, conexão VPN ou conexão AWS Direct Connect. As instâncias na sua VPC não precisam de endereços IP públicos para se comunicar com as APIs do Lake Formation. O tráfego entre sua VPC e Lake Formation não sai da rede Amazon.

Cada endpoint de interface é representado por uma ou mais [Interfaces de Rede Elástica](#) nas sub-redes.

Para obter mais informações, consulte [Endpoints da VPC da interface \(AWS PrivateLink\)](#) no Manual do Usuário do Amazon VPC.

## Considerações sobre endpoints da VPC do Lake Formation

Antes de configurar um endpoint da VPC de interface para o Lake Formation, revise [Propriedades e limitações do endpoint de interface](#) no Guia do usuário da Amazon VPC.

Lake Formation oferece suporte para fazer chamadas para todas as ações de API de sua VPC. Você pode usar o Lake Formation com endpoints da VPC em todos os Regiões da AWS que oferecem suporte ao Lake Formation e aos endpoints da VPC da Amazon.

## Criando um endpoint da VPC de interface para o Lake Formation

Você pode criar um endpoint da VPC para o serviço Lake Formation usando o console da Amazon VPC ou o AWS Command Line Interface (AWS CLI). Para obter mais informações, consulte [Criar um endpoint da interface](#) no Guia do usuário da Amazon VPC.

Crie um endpoint da VPC para o Lake Formation usando o seguinte nome de serviço:

- `com.amazonaws.region.lakeformation`

Se você habilitar o DNS privado para o endpoint, poderá fazer solicitações de API para Lake Formation usando seu nome DNS padrão para a região, por exemplo, `lakeformation.us-east-1.amazonaws.com`.

Para obter mais informações, consulte [Acessar um serviço por um endpoint de interface](#) no Guia do usuário da Amazon VPC.

## Criando uma política de endpoint da VPC para o Lake Formation

O Lake Formation oferece suporte a políticas de endpoint da VPC. Uma política de endpoint da VPC é uma política de recursos (IAM) AWS Identity and Access Management que você anexa a um endpoint ao criar ou modificar o endpoint.

Você pode anexar uma política de endpoint ao endpoint da VPC que controla o acesso ao Lake Formation. Essa política especifica as seguintes informações:

- A entidade principal que pode executar ações.
- As ações que podem ser executadas.
- Os recursos sobre os quais as ações podem ser realizadas.

Para obter mais informações, consulte [Controlar o acesso a serviços com endpoints da VPC](#) no Guia do usuário da Amazon VPC.

### Exemplo: política de endpoint da VPC para ações do Lake Formation

O exemplo a seguir da política de endpoints da VPC para o Lake Formation permite o fornecimento de credenciais usando as permissões do Lake Formation. Você pode usar essa política para executar consultas usando as permissões do Lake Formation de um cluster do Amazon Redshift ou de um cluster Amazon EMR localizado em uma sub-rede privada.

```
{
  "Statement": [
    {
      "Effect": "Allow",
      "Action": "lakeformation:GetDataAccess",
      "Resource": "*",
      "Principal": "*"
    }
  ]
}
```

**Note**

Se você não anexar uma política ao criar um endpoint, uma política padrão que permite acesso total ao serviço será anexada.

Para obter mais informações, consulte estes tópicos na documentação da Amazon VPC:

- [O que é Amazon VPC?](#)
- [Crie um terminal de interface](#)
- [Usar políticas de endpoint VPC](#)

# Tutoriais

Os tutoriais a seguir estão organizados em três faixas e fornecem instruções passo a passo sobre como criar um data lake, ingerir dados, compartilhar e proteger data lakes usando AWS Lake Formation:

1. Criar um data lake e ingerir dados: aprenda a criar um data lake e use esquemas para mover, armazenar, catalogar, limpar e organizar seus dados. Você também aprenderá a configurar tabelas controladas. Uma tabela controlada é um novo tipo de tabela do Amazon S3 que suporta transações atômicas, consistentes, isoladas e duráveis (ACID).

Antes de começar, certifique-se de ter concluído as etapas em [Introdução ao Lake Formation](#).

- [Criação de um data lake a partir de uma AWS CloudTrail fonte](#)

Crie e carregue seu primeiro data lake usando seus próprios logs do CloudTrail como fonte de dados.

- [Criação de um data lake a partir de uma fonte JDBC no Lake Formation](#)

Crie um data lake usando um dos seus armazenamentos de dados acessíveis por JDBC, como um banco de dados relacional, como fonte de dados.

2. Proteger data lakes: aprenda a usar controles de acesso baseados em tags e em nível de linha para proteger e gerenciar com eficácia o acesso aos seus data lakes.

- [Como configurar permissões para formatos de armazenamento de tabelas abertas no Lake Formation](#)

Este tutorial demonstra como configurar permissões para formatos de tabela transacional de código aberto (tabelas do Apache Iceberg, Apache Hudi e Linux Foundation Delta Lake) no Lake Formation.

- [Como gerenciar um data lake usando o controle de acesso baseado em tags do Lake Formation](#)

Aprenda a gerenciar o acesso aos dados em um data lake usando o controle de acesso baseado em tags no Lake Formation.

- [Como proteger os data lakes com controle de acesso em nível de linha](#)

Aprenda a configurar permissões em nível de linha que permitem restringir o acesso a linhas específicas com base nas políticas de conformidade e governança de dados no Lake Formation.

3. Compartilhar dados: aprenda a compartilhar seus dados com segurança Contas da AWS usando o controle de acesso baseado em tags (TBAC) e gerencie permissões granulares em conjuntos de dados compartilhados entre eles. Contas da AWS

- [Compartilhamento de um Data Lake usando controle de acesso baseado em tags do Lake Formation e recursos nomeados](#)

Neste tutorial, você aprenderá como compartilhar seus dados com segurança Contas da AWS usando o Lake Formation.

- [Como compartilhar um data lake usando o controle de acesso refinado do Lake Formation](#)

Neste tutorial, você aprende como compartilhar conjuntos de dados de forma rápida e fácil usando o Lake Formation ao gerenciar vários Contas da AWS com AWS Organizations.

## Tópicos

- [Criação de um data lake a partir de uma AWS CloudTrail fonte](#)
- [Criação de um data lake a partir de uma fonte JDBC no Lake Formation](#)
- [Como configurar permissões para formatos de armazenamento de tabelas abertas no Lake Formation](#)
- [Como gerenciar um data lake usando o controle de acesso baseado em tags do Lake Formation](#)
- [Como proteger os data lakes com controle de acesso em nível de linha](#)
- [Compartilhamento de um Data Lake usando controle de acesso baseado em tags do Lake Formation e recursos nomeados](#)
- [Como compartilhar um data lake usando o controle de acesso refinado do Lake Formation](#)

## Criação de um data lake a partir de uma AWS CloudTrail fonte

Este tutorial orienta você pelas ações a serem tomadas no console do Lake Formation para criar e carregar seu primeiro data lake a partir de uma AWS CloudTrail fonte.

### Etapas de alto nível para criar um data lake

1. Registre um caminho do Amazon Simple Storage Service (Amazon S3) como um data lake.
2. Conceda permissões do Lake Formation para gravar no Catálogo de Dados e nos locais do Amazon S3 no data lake.
3. Crie um banco de dados para organizar as tabelas de metadados no Catálogo de Dados.

4. Use um esquema para criar um fluxo de trabalho. Execute o fluxo de trabalho para ingerir dados de uma fonte de dados.
5. Configure suas permissões do Lake Formation para permitir que outras pessoas gerenciem dados no Catálogo de Dados e no Data Lake.
6. Configure o Amazon Athena para consultar os dados que você importou no seu data lake do Amazon S3.
7. Para alguns tipos de armazenamento de dados, configure o Amazon Redshift Spectrum para consultar os dados que você importou para o seu data lake do Amazon S3.

## Tópicos

- [Público-alvo](#)
- [Pré-requisitos](#)
- [Etapa 1: Criar um usuário de analista de dados](#)
- [Etapa 2: adicionar permissões para ler AWS CloudTrail registros à função do fluxo de trabalho](#)
- [Etapa 3: Criar um bucket do Amazon S3 para o data lake](#)
- [Etapa 4: Registrar um caminho do Amazon S3](#)
- [Etapa 5: Conceder permissões de local de dados](#)
- [Etapa 6: Criar um banco de dados no Catálogo de Dados](#)
- [Etapa 7: Conceder permissões de dados](#)
- [Etapa 8: usar um esquema para criar um fluxo de trabalho](#)
- [Etapa 9: executar o fluxo de trabalho](#)
- [Etapa 10: Conceder SELECT nas tabelas](#)
- [Etapa 11: Consultar o data lake usando Amazon Athena](#)

## Público-alvo

A tabela a seguir lista as perfis usadas neste tutorial para criar um data lake.

## Público-alvo

Função	Descrição
Administrador do IAM	Tem a política AWS gerenciada: <code>AdministratorAccess</code> . Criar perfis do IAM e buckets do Amazon S3.
Administrador do data lake	Usuário que pode acessar o Catálogo de Dados, criar bancos de dados e conceder permissões do Lake Formation a outros usuários. Tem menos permissões do IAM do que o administrador do IAM, mas o suficiente para administrar o data lake.
Analista de dados	Usuário que pode executar consultas no data lake. Tem permissões suficientes apenas para executar consultas.
Perfil de fluxo de trabalho	Perfil com as políticas de IAM necessárias para executar um fluxo de trabalho. Para ter mais informações, consulte <a href="#">(Opcional) Crie uma função do IAM para fluxos de trabalho</a> .

## Pré-requisitos

### Antes de começar

- Verifique se você concluiu as tarefas no [Configurar o AWS Lake Formation](#).
- Conheça a localização dos seus CloudTrail registros.
- O Athena exige que a pessoa do analista de dados crie um bucket do Amazon S3 para armazenar os resultados da consulta antes de usar o Athena.

Presume-se familiaridade com AWS Identity and Access Management (IAM). Para obter informações sobre o IAM, consulte o [Guia do Usuário do IAM](#).



## Etapa 1: Criar um usuário de analista de dados

Esse usuário tem as permissões de privilégio mínimo para consultar o data lake.

1. Abra o console do IAM em <https://console.aws.amazon.com/iam>. Entre como o usuário administrador que você criou [Criar um usuário administrativo](#) ou como um usuário com a política AdministratorAccess AWS gerenciada.
2. Crie um usuário nomeado `dataLake_user` com as seguintes configurações:
  - Habilite AWS Management Console o acesso.
  - Defina uma senha e não solicite redefinição de senha.
  - Anexe a política AmazonAthenaFullAccess AWS gerenciada.
  - Anexe a seguinte política embutida. Atribua o nome `DataLakeUserBasic` à política.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "lakeformation:GetDataAccess",
        "glue:GetTable",
        "glue:GetTables",
        "glue:SearchTables",
        "glue:GetDatabase",
        "glue:GetDatabases",
        "glue:GetPartitions",
        "lakeformation:GetResourceLFTags",
        "lakeformation:ListLFTags",
        "lakeformation:GetLFTag",
        "lakeformation:SearchTablesByLFTags",
        "lakeformation:SearchDatabasesByLFTags"
      ],
      "Resource": "*"
    }
  ]
}
```

## Etapa 2: adicionar permissões para ler AWS CloudTrail registros à função do fluxo de trabalho

1. Anexe a política em linha a seguir ao perfil LakeFormationWorkflowRole. A política concede permissão para ler seus AWS CloudTrail registros. Atribua o nome DataLakeGetCloudTrail à política.

Para criar a função LakeFormationWorkflowRole, consulte [\(Opcional\) Crie uma função do IAM para fluxos de trabalho](#).

### Important

<your-s3-cloudtrail-bucket>Substitua pela localização dos seus CloudTrail dados no Amazon S3.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": "s3:GetObject",
      "Resource": ["arn:aws:s3:::<your-s3-cloudtrail-bucket>/*"]
    }
  ]
}
```

2. Verifique se há três políticas vinculadas ao perfil.

## Etapa 3: Criar um bucket do Amazon S3 para o data lake

Crie o bucket do Amazon S3 que será o local raiz do seu data lake.

1. Abra o console do Amazon S3 em <https://console.aws.amazon.com/s3/> e entre como o usuário administrador que você criou no [Criar um usuário administrativo](#).
2. Escolha Criar bucket e acesse o assistente para criar um bucket nomeado como <yourName>-datalake-cloudtrail, onde <yourName> está seu primeiro nome e sobrenome. Por exemplo: jdoe-datalake-cloudtrail.

Para obter instruções detalhadas sobre como criar um bucket do Amazon S3, consulte [Como criar um bucket](#).

## Etapa 4: Registrar um caminho do Amazon S3

Registre um caminho do Amazon S3 como o local raiz do seu data lake.

1. Abra o console do Lake Formation em <https://console.aws.amazon.com/lakeformation/>. Faça login como administrador de data lake.
2. No painel de navegação, em Registrar e ingerir, selecione Locais do Data Lake.
3. Selecione Registrar local e, em seguida, Navegar.
4. Selecione o bucket do `<yourName>-datalake-cloudtrail` que você criou anteriormente, aceite o perfil padrão do IAM `AWSServiceRoleForLakeFormationDataAccess` e selecione Registrar local.

Para obter mais informações sobre o registro de locais, consulte [Adicionar uma localização do Amazon S3 ao seu data lake](#).

## Etapa 5: Conceder permissões de local de dados

As entidades principais devem ter permissões de local de dados em um local de data lake para criar tabelas ou bancos de dados do Catálogo de Dados que apontem para esse local. Você deve conceder permissões de local de dados ao perfil do IAM para fluxos de trabalho para que o fluxo de trabalho possa gravar no destino de ingestão de dados.

1. No painel de navegação, em Permissões, selecione Locais de dados.
2. Selecione Conceder e, na caixa de diálogo Conceder permissões, faça estas seleções:
  - a. Para Usuários e perfis do IAM, selecione `LakeFormationWorkflowRole`.
  - b. Para locais de armazenamento, escolha seu bucket `<yourName>-datalake-cloudtrail`.
3. Selecione Conceder.

Para obter mais informações sobre permissões de localização de dados, consulte [Underlying data access control](#).

## Etapa 6: Criar um banco de dados no Catálogo de Dados

As tabelas de metadados no Catálogo de Dados do Lake Formation são armazenadas em um banco de dados.

1. No painel de navegação, em Catálogo de dados, selecione Bancos de dados.
2. selecione Criar banco de dados e, em Informações do banco de dados, digite o nome `lakeformation_cloudtrail`.
3. Deixe os outros campos em branco e escolha Criar banco de dados.

## Etapa 7: Conceder permissões de dados

Você deve conceder permissões para criar tabelas de metadados no Catálogo de Dados. Como o fluxo de trabalho será executado com o perfil de `LakeFormationWorkflowRole`, você deve conceder essas permissões ao perfil.

1. No console do Lake Formation, no painel de navegação, em Catálogo de Dados, selecione Bancos de dados.
2. Selecione o banco de dados `lakeformation_cloudtrail`, e na lista suspensa Ações, selecione Conceder sob o título Permissões.
3. Na caixa de diálogo Conceder permissões de dados, faça estas seleções:
  - a. Em Entidade principais, para Usuário e perfis do IAM, selecione `LakeFormationWorkflowRole`.
  - b. Em LF-tags ou recursos de catálogo, selecione Recursos nomeados de Catálogo de Dados.
  - c. Para Bancos de dados, você deve ver que o banco de dados `lakeformation_cloudtrail` já foi adicionado.
  - d. Em Permissões do banco de dados, selecione Criar tabela, Alterar e Eliminar e desmarque Super se estiver selecionado.

Sua caixa de diálogo Conceder permissões de dados agora deve ter a aparência desta captura de tela.

## Grant data permissions

### Principals

**IAM users and roles**

Users or roles from this AWS account.

**SAML users and groups**

SAML users and group or QuickSight ARNs.

**External accounts**

AWS accounts or AWS organizations outside of this account.

#### IAM users and roles

Add one or more IAM users or roles.

Choose IAM principals to add

LakeFormationWorkflowRole ✕  
Role

### LF-Tags or catalog resources

**Resources matched by LF-Tags (recommended)**

Manage permissions indirectly for resources or data matched by a specific set of LF-Tags.

**Named data catalog resources**

Manage permissions for specific databases or tables, in addition to fine-grained data access.

#### Databases

Select one or more databases.

Choose databases

Load more

lakeformation-cloudtrail ✕  
007436865787

#### Tables - optional

Select one or more tables.

Choose tables

Load more

### Database permissions

#### Database permissions

Choose specific access permissions to grant.

- Create table  Alter  Drop  
 Describe

**Super**

This permission is the union of all the individual permissions to the left, and supersedes them.

#### Grantable permissions

Choose the permission that may be granted to others.

- Create table  Alter  Drop  
 Describe

**Super**

This permission allows the principal to grant any of the permissions to the left, and supersedes those grantable permissions.

## 4. Selecione Conceder.

Para obter mais informações sobre como conceder permissões ao Lake Formation, consulte [Gerenciando permissões do Lake Formation](#).

## Etapa 8: usar um esquema para criar um fluxo de trabalho

Para ler os CloudTrail registros, entender sua estrutura e criar as tabelas apropriadas no Catálogo de Dados, precisamos configurar um fluxo de trabalho que consiste em AWS Glue rastreadores, trabalhos, acionadores e fluxos de trabalho. Os esquemas do Lake Formation simplificam esse processo.

O fluxo de trabalho gera trabalhos, crawlers e gatilhos que descobrem e ingerem dados em seu data lake. Crie um fluxo de trabalho com base em um dos esquemas predefinidos do Lake Formation.

1. No console do Lake Formation, no painel de navegação, selecione Esquemas e, em seguida, selecione Usar um esquema.
2. Na página Usar um blueprint, em Tipo de blueprint, escolha AWS CloudTrail
3. Em Importar fonte, escolha uma CloudTrail fonte e uma data de início.
4. Em Destino da importação, especifique estes parâmetros:

Bancos de dados de destino	lakeformation_cloudtrail
Local de armazenamento de destino	s3://<yourName> -datalake-cloudtrail
Formato de dados	Parquet

5. Para frequência de importação, selecione Executar sob demanda.
6. Em Opções de importação, especifique estes parâmetros:

Nome do fluxo de trabalho	lakeformationcloudtrailtest
Perfil do IAM	LakeFormationWorkflowRole
Prefixo da tabela	cloudtrailtest

 Note

Deve estar em letras minúsculas.

7. Escolha Criar e aguarde até que o console informe que o fluxo de trabalho foi criado com sucesso.

 Tip

Você recebeu a seguinte mensagem de erro?

```
User: arn:aws:iam::<account-id>:user/<datalake_administrator_user> is not authorized to perform: iam:PassRole on resource:arn:aws:iam::<account-id>:role/LakeFormationWorkflowRole...
```

Nesse caso, verifique se você substituiu <account-id>a política em linha do usuário administrador do data lake por um número de AWS conta válido.

## Etapa 9: executar o fluxo de trabalho

Como você especificou que o fluxo de trabalho é run-on-demand, você deve iniciar manualmente o fluxo de trabalho.

- Na página Esquemas, selecione o fluxo de trabalho lakeformationcloudtrailtest e, no menu Ações, selecione Iniciar.

À medida que o fluxo de trabalho é executado, você pode ver seu progresso na coluna de Status da última execução. Selecione o botão de atualizar ocasionalmente.

O status vai de EM EXECUÇÃO, para Descoberta, para Importação, e CONCLUÍDO.

Quando o fluxo de trabalho for concluído:

- O Catálogo de Dados terá novas tabelas de metadados.
- Seus CloudTrail registros serão ingeridos no data lake.

Se o fluxo de trabalho falhar, faça o seguinte:

- a. Selecione o fluxo de trabalho e, no menu Ações, selecione Exibir gráfico.  
  
O fluxo de trabalho é aberto no console do AWS Glue.
- b. Certifique-se de que o fluxo de trabalho esteja selecionado e escolha a guia Histórico.
- c. Em Histórico, selecione a execução mais recente e selecione Exibir informações da execução.
- d. Selecione um trabalho ou crawler com falha no gráfico dinâmico (runtime) e revise a mensagem de erro. Os nós com falha são vermelhos ou amarelos.

## Etapa 10: Conceder SELECT nas tabelas

Você deve conceder a permissão SELECT nas novas tabelas do Catálogo de Dados para que o analista de dados possa consultar os dados para os quais as tabelas apontam.

### Note

Um fluxo de trabalho concede automaticamente a permissão SELECT nas tabelas que ele cria ao usuário que o executou. Como o administrador do data lake executou esse fluxo de trabalho, você deve conceder SELECT ao analista de dados.

1. No console do Lake Formation, no painel de navegação, em Catálogo de Dados, selecione Bancos de dados.
2. Selecione o banco de dados `lakeformation_cloudtrail`, e na lista suspensa Ações, selecione Conceder sob o título Permissões.
3. Na caixa de diálogo Conceder permissões de dados, faça estas seleções:
  - a. Em Entidade principais, para Usuário e perfis do IAM, selecione `datalake_user`.
  - b. Em LF-tags ou recursos de catálogo, selecione Recursos nomeados de Catálogo de Dados.
  - c. Para Bancos de dados, o banco de dados `lakeformation_cloudtrail` já deve estar selecionado.
  - d. Para Tabelas, selecione `cloudtrailtest-cloudtrail`.
  - e. Em Permissões de tabela e coluna, clique em Selecionar.
4. Selecione Conceder.



A próxima etapa é executada como analista de dados.

## Etapa 11: Consultar o data lake usando Amazon Athena

Use o Amazon Athena console para consultar os CloudTrail dados em seu data lake.

1. Abra o console do Athena em <https://console.aws.amazon.com/athena/> e faça login como analista de dados, usuário `dataLake_user`.
2. Se necessário, selecione Começar para continuar com o editor de consultas do Athena.
3. Em Data source (Fonte de dados), selecione `AwsDataCatalog`.
4. Para o Banco de dados, selecione `lakeformation_cloudtrail`.

A lista de Tabelas é preenchida.

5. No menu flutuante (3 pontos horizontais) ao lado da tabela `cloudtrailtest-cloudtrail`, selecione Exibir tabela e selecione Executar.

A consulta é executada e exibe 10 linhas de dados.

Se você nunca usou o Athena antes, primeiro deve configurar um local do Amazon S3 no console do Athena para armazenar os resultados da consulta. O `dataLake_user` deve ter as permissões exigidas para acessar o bucket do Amazon S3 escolhido.

### Note

Agora que você concluiu o tutorial, conceda permissões de dados e permissões de local de dados às entidades principais da sua organização.

## Criação de um data lake a partir de uma fonte JDBC no Lake Formation

Este tutorial orienta você pelas etapas a serem seguidas no console AWS Lake Formation para criar e carregar seu primeiro data lake a partir de uma fonte JDBC usando o Lake Formation.

Tópicos

- [Público-alvo](#)
- [Pré-requisitos do tutorial JDBC](#)

- [Etapa 1: criar um usuário de analista de dados](#)
- [Etapa 2: criar uma conexão em AWS Glue](#)
- [Etapa 3: criar um bucket do Amazon S3 para o data lake](#)
- [Etapa 4: registrar um caminho do Amazon S3](#)
- [Etapa 5: conceder permissões de localização de dados](#)
- [Etapa 6: criar um banco de dados no Catálogo de dados](#)
- [Etapa 7: conceder permissões de dados](#)
- [Etapa 8: usar um esquema para criar um fluxo de trabalho](#)
- [Etapa 9: executar o fluxo de trabalho](#)
- [Etapa 10: conceder SELECIONAR nas tabelas](#)
- [Etapa 11: consultar o data lake usando Amazon Athena](#)
- [Etapa 12: consulte os dados no data lake usando o Amazon Redshift Spectrum](#)
- [Etapa 13: conceder ou revogar permissões do Lake Formation usando o Amazon Redshift Spectrum](#)

## Público-alvo

A tabela a seguir lista as funções que são usadas neste tutorial do [JDBC AWS Lake Formation](#).

Função	Descrição
Administrador do IAM	Um usuário que pode criar usuários e perfis do AWS Identity and Access Management (IAM) e buckets do Amazon Simple Storage Service (Amazon S3). Possui a política <code>AdministratorAccess</code> gerenciada por AWS.
Administrador do data lake	Um usuário que pode acessar o Catálogo de dados, criar bancos de dados e conceder permissões do Lake Formation a outros usuários. Possui menos permissões do IAM do que o administrador do IAM, mas o suficiente para administrar o data lake.

Função	Descrição
Analista de dados	Um usuário que pode executar consultas no data lake. Tem permissões suficientes apenas para executar consultas.
Função do fluxo de trabalho	Um perfil com as políticas de IAM necessárias para executar um fluxo de trabalho.

Para obter informações sobre os pré-requisitos para concluir o tutorial, consulte [Pré-requisitos do tutorial JDBC](#).

## Pré-requisitos do tutorial JDBC

Antes de começar o [tutorial do AWS Lake Formation JDBC](#), certifique-se de ter feito o seguinte:

- Conclua as tarefas em [Introdução ao Lake Formation](#).
- Escolha um armazenamento de dados acessível por JDBC que você deseja usar para o tutorial.
- Reúna as informações necessárias para criar uma conexão AWS Glue do tipo JDBC. Esse objeto do Catálogo de dados inclui o URL do armazenamento de dados, as credenciais de login e, se o armazenamento de dados foi criado em uma Amazon Virtual Private Cloud (Amazon VPC), informações adicionais de configuração específicas da VPC. Para obter mais informações, consulte [Definindo conexões no catálogo de dados AWS Glue](#) no Guia do desenvolvedor AWS Glue.

O tutorial pressupõe que você esteja familiarizado com AWS Identity and Access Management (IAM). Para obter informações sobre o IAM, consulte o [Guia do usuário do IAM](#).

Para começar, vá para [the section called “Etapa 1: criar um usuário de analista de dados”](#).

## Etapa 1: criar um usuário de analista de dados

Nesta etapa, você cria um usuário AWS Identity and Access Management (IAM) para ser o analista de dados do seu data lake em AWS Lake Formation.

Esse usuário tem o conjunto mínimo de permissões para consultar o data lake.

1. Abra o console do IAM em <https://console.aws.amazon.com/iam>. Entre como o usuário administrador que você criou [Criar um usuário administrativo](#) ou como um usuário com a política AdministratorAccess gerenciada por AWS.
2. Crie um usuário nomeado `dataLake_user` com as seguintes configurações:
  - Habilitar acesso a AWS Management Console.
  - Defina uma senha e não solicite redefinição de senha.
  - Anexe a política AmazonAthenaFullAccess gerenciada por AWS.
  - Anexe a seguinte política embutida. Atribua o nome `DataLakeUserBasic` à política.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "lakeformation:GetDataAccess",
        "glue:GetTable",
        "glue:GetTables",
        "glue:SearchTables",
        "glue:GetDatabase",
        "glue:GetDatabases",
        "glue:GetPartitions",
        "lakeformation:GetResourceLFTags",
        "lakeformation:ListLFTags",
        "lakeformation:GetLFTag",
        "lakeformation:SearchTablesByLFTags",
        "lakeformation:SearchDatabasesByLFTags"
      ],
      "Resource": "*"
    }
  ]
}
```

## Etapa 2: criar uma conexão em AWS Glue

### Note

Ignore esta etapa se você já tiver uma conexão AWS Glue com sua fonte de dados JDBC.

AWS Lake Formation acessa fontes de dados JDBC por meio de uma conexão AWS Glue. Uma conexão é um objeto do Catálogo de dados que contém todas as informações necessárias para se conectar à fonte de dados. Você pode criar uma conexão usando o console AWS Glue.

### Como criar uma conexão

1. Abra o console AWS Glue em <https://console.aws.amazon.com/glue/> e entre como o usuário administrador que você criou em [Criar um usuário administrativo](#).
2. No painel de navegação, em Data catalog (Catálogo de dados), selecione Connections (Conexões).
3. Na página Connectors (Conectores) escolha Create custom connector (Criar um conector personalizado).
4. Na página Propriedades do conector, insira **datalake-tutorial** o nome da conexão e escolha JDBC como o tipo de conexão. Em seguida, escolha Next (Próximo).
5. Continue através do assistente de conexão e salve a conexão.

Para obter informações sobre como criar uma conexão, consulte [Propriedades da conexão JDBC AWS Glue](#) no Guia do desenvolvedor AWS Glue.

## Etapa 3: criar um bucket do Amazon S3 para o data lake

Nesta etapa, você criará o bucket do Amazon Simple Storage Service (Amazon S3) que será a localização raiz do data lake.

1. Abra o console do Amazon S3 em <https://console.aws.amazon.com/s3/> e faça login como o usuário administrador que você criou em [Criar um usuário administrativo](#).
2. Escolha Criar bucket e acesse o assistente para criar um bucket nomeado como **<yourName>-datalake-tutorial**, onde **<yourName>** está seu primeiro nome e sobrenome. Por exemplo: **jdoe-datalake-tutorial**.

Para obter instruções detalhadas sobre como criar um bucket do Amazon S3, consulte [Como criar um bucket do S3?](#) no Guia do usuário do Amazon Simple Storage Service.

## Etapa 4: registrar um caminho do Amazon S3

Nesta etapa, você registra um caminho do Amazon Simple Storage Service (Amazon S3) como a localização raiz do data lake.

1. Abra o console do Lake Formation em <https://console.aws.amazon.com/lakeformation/>. Faça login como administrador de data lake.
2. No painel de navegação, em Registrar e ingerir, escolha as localizações do Data lake.
3. Escolha Registrar localização e, em seguida, escolha Procurar.
4. Selecione o bucket `<yourName>-datalake-tutorial` que você criou anteriormente, aceite o perfil padrão do IAM `AWSserviceRoleForLakeFormationDataAccess` e escolha Registrar localização.

Para obter mais informações sobre o registro de locais, consulte [Adicionar uma localização do Amazon S3 ao seu data lake](#).

## Etapa 5: conceder permissões de localização de dados

Os diretores devem ter permissões de localização de dados em uma localização de data lake para criar tabelas ou bancos de dados do Catálogo de dados que apontem para essa localização. Você deve conceder permissões de localização de dados ao perfil do IAM para fluxos de trabalho para que o fluxo de trabalho possa gravar no destino da ingestão de dados.

1. No console do Lake Formation, no painel de navegação, em Permissões, escolha Localizações de dados.
2. Escolha Conceder e, na caixa de diálogo Conceder permissões, faça o seguinte:
  - a. Para usuários e perfis do IAM, escolha `LakeFormationWorkflowRole`.
  - b. Para locais de armazenamento, escolha seu bucket `<yourName>-datalake-tutorial`.
3. Escolha Grant (Conceder).

Para obter mais informações sobre permissões de localização de dados, consulte [Underlying data access control](#).

## Etapa 6: criar um banco de dados no Catálogo de dados

As tabelas de metadados no Catálogo de dados do Lake Formation são armazenadas em um banco de dados.

1. No console do Lake Formation, no painel de navegação, em Catálogo de dados, escolha Bancos de dados.
2. Escolha Criar banco de dados e, em Detalhes do banco de dados, insira o nome `lakeformation_tutorial`.
3. Deixe os outros campos em branco e escolha Criar banco de dados.

## Etapa 7: conceder permissões de dados

Você deve conceder permissões para criar tabelas de metadados no Catálogo de dados. Como o fluxo de trabalho é executado com a função `LakeFormationWorkflowRole`, você deve conceder essas permissões à função.

1. No console do Lake Formation, no painel de navegação, em Permissões, escolha Permissões do data lake.
2. Escolha Conceder e, na caixa de diálogo Conceder permissões de dados, faça o seguinte:
  - a. Em Entidades principais, para Usuário e perfis do IAM, escolha `LakeFormationWorkflowRole`.
  - b. Em LF-Tags ou recursos de catálogo, escolha Recursos de catálogo de dados nomeados.
  - c. Em Bancos de dados, escolha o banco de dados que você criou anteriormente, `lakeformation_tutorial`.
  - d. Em Permissões do banco de dados, selecione Criar tabela, Alterar e Eliminar, e desmarque Super se estiver selecionado.
3. Escolha Grant (Conceder).

Para obter mais informações sobre como conceder permissões ao Lake Formation, consulte [Visão geral das permissões do Lake Formation](#).

## Etapa 8: usar um esquema para criar um fluxo de trabalho

O fluxo de trabalho AWS Lake Formation gera AWS Glue trabalhos, rastreadores e gatilhos que descobrem e ingerem dados em seu data lake. Você cria um fluxo de trabalho com base em um dos esquemas predefinidos do Lake Formation.

1. No console do Lake Formation, no painel de navegação, escolha Esquemas e, em seguida, escolha Usar esquema.
2. Na página Usar um esquema, em Tipo de esquema, escolha Snapshot de banco de dados.
3. Em Importar fonte, para Conexão de banco de dados, escolha a conexão que você acabou de criar, `dataLake-tutorial` ou escolha uma conexão existente para sua fonte de dados.
4. Em Caminho de dados de origem, insira o caminho do qual ingerir dados, no formulário `<database>/<schema>/<table>`.

Você pode substituir o curinga percentual (%) por esquema ou tabela. Para bancos de dados que suportam esquemas, insira `<database>/<schema>/%` para corresponder todas as tabelas em `<schema>` dentro de `<database>`. Banco de dados Oracle e MySQL não suportam esquema no caminho; em vez disso, insira `<database>/%`. Para o Banco de dados Oracle, `<database>` é o identificador do sistema (SID).

Por exemplo, se um banco de dados Oracle tiver `orc1` como SID, insira `orc1/%` para corresponder a todas as tabelas às quais o usuário especificado na conexão JDBC tem acesso.

### Important

Este campo diferencia letras maiúsculas de minúsculas.


5. Em Destino da importação, especifique estes parâmetros:

Bancos de dados de destino	lakeformation_tutorial
Local de armazenamento de destino	s3://<yourName> -dataLake-tutorial
Formato de dados	(Escolha Parquet ou CSV)

6. Para frequência de importação, escolha Executar sob demanda.
7. Em Opções de importação, especifique estes parâmetros:



Nome do fluxo de trabalho	lakeformationjdbctest
Perfil do IAM	LakeFormationWorkflowRole
Prefixo da tabela	jdbctest

 Note

Deve estar em letras minúsculas.

- Escolha Criar e aguarde até que o console informe que o fluxo de trabalho foi criado com sucesso.

 Tip

Você recebeu a seguinte mensagem de erro?

```
User: arn:aws:iam::<account-id>:user/<dataLake_administrator_user> is not authorized to perform: iam:PassRole on resource:arn:aws:iam::<account-id>:role/LakeFormationWorkflowRole...
```

Nesse caso, verifique se você substituiu *<account-id>* na política embutida do usuário administrador do data lake por um número de conta válido AWS.

## Etapa 9: executar o fluxo de trabalho

Como você especificou que o fluxo de trabalho é executado sob demanda, você deve iniciá-lo manualmente em AWS Lake Formation.

- No console do Lake Formation, na página Esquemas, selecione o fluxo de trabalho lakeformationjdbctest.
- Escolha Ações e, em seguida, escolha Iniciar.
- À medida que o fluxo de trabalho é executado, visualize seu progresso na coluna Status da última execução. Escolha o botão de atualização ocasionalmente.

O status vai de EM EXECUÇÃO, para Descobrimdo, depois Importando e, em seguida, CONCLUÍDO.

Quando o fluxo de trabalho estiver concluído:

- O Catálogo de dados tem novas tabelas de metadados.
- Seus dados são ingeridos no data lake.

Se o fluxo de trabalho falhar, faça o seguinte:

- a. Selecione o fluxo de trabalho. Escolha Ações e, em seguida, escolha Exibir gráfico.

O fluxo de trabalho é aberto no console AWS Glue.

- b. Selecione o fluxo de trabalho e escolha a guia Histórico.
- c. Selecione a execução mais recente e escolha Exibir detalhes da execução.
- d. Selecione um trabalho ou crawler com falha no gráfico dinâmico (runtime) e revise a mensagem de erro. Os nós com falha são vermelhos ou amarelos.

## Etapa 10: conceder SELECIONAR nas tabelas

Você deve conceder a permissão SELECT nas novas tabelas do Catálogo de dados em AWS Lake Formation para que o analista de dados possa consultar os dados para os quais as tabelas apontam.

### Note

Um fluxo de trabalho concede automaticamente a permissão SELECT nas tabelas que ele cria ao usuário que o executou. Como o administrador do data lake executou esse fluxo de trabalho, você deve conceder SELECT ao analista de dados.

1. No console do Lake Formation, no painel de navegação, em Permissões, escolha Permissões do data lake.
2. Escolha Conceder e, na caixa de diálogo Conceder permissões de dados, faça o seguinte:
  - a. Em Entidades principais, para usuário e perfis do IAM, escolha `dataLake_user`.
  - b. Em LF-Tags ou recursos de catálogo, escolha Recursos de catálogo de dados nomeados.

- c. Para Bancos de dados, escolha `lakeformation_tutorial`.  
A lista de Tabelas é preenchida.
  - d. Para Tabelas, selecione uma ou mais tabelas da fonte de dados.
  - e. Em Permissões de tabela e coluna, escolha Selecionar.
3. Escolha Grant (Conceder).

A próxima etapa é executada como analista de dados.

## Etapa 11: consultar o data lake usando Amazon Athena

Use o console Amazon Athena para consultar os dados em seu data lake.

1. Abra o console do Athena em <https://console.aws.amazon.com/athena/> e faça login como analista de dados, usuário `dataLake_user`.
2. Se necessário, escolha Começar para continuar com o editor de consultas do Athena.
3. Para Fonte de dados, escolha `AWSDataCatalog`.
4. Para o Database (Banco de dados), escolha `lakeformation_tutorial`.

A lista Tabelas é preenchida.

5. No menu pop-up ao lado de uma das tabelas, escolha Visualizar tabela.

A consulta é executada e exibe 10 linhas de dados.

## Etapa 12: consulte os dados no data lake usando o Amazon Redshift Spectrum

Você pode configurar o Amazon Redshift Spectrum para consultar os dados importados para o data lake do Amazon Simple Storage Service (Amazon S3). Primeiro, crie um perfil (IAM) AWS Identity and Access Management que seja usado para iniciar o cluster do Amazon Redshift e consultar os dados do Amazon S3. Em seguida, conceda a essa função as permissões `Select` nas tabelas que você deseja consultar. Em seguida, conceder permissões do usuário para usar o Editor de consultas do Amazon Redshift. Por fim, crie um cluster do Amazon Redshift e execute consultas.

Você cria o cluster como administrador e consulta o cluster como analista de dados.

Para obter mais informações sobre o Amazon Redshift Spectrum, consulte [Usando o Amazon Redshift Spectrum para consultar dados externos](#) no Guia do desenvolvedor de banco de dados do Amazon Redshift.

Para configurar permissões para executar consultas do Amazon Redshift

1. Abra o console do IAM em <https://console.aws.amazon.com/iam/>. Entre como o usuário administrador que você criou em [Criar um usuário administrativo](#) (nome de usuário Administrator) ou como um usuário com a política AdministratorAccess gerenciada por AWS.
2. No painel de navegação, escolha Políticas (Políticas).

Se essa for a primeira vez que escolhe Políticas, a página Bem-vindo às políticas gerenciadas será exibida. Escolha Get Started.

3. Escolha Create policy (Criar política).
4. Escolha a guia JSON.
5. Cole no seguinte documento de política JSON.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "lakeformation:GetDataAccess",
        "glue:GetTable",
        "glue:GetTables",
        "glue:SearchTables",
        "glue:GetDatabase",
        "glue:GetDatabases",
        "glue:GetPartitions",
        "lakeformation:GetResourceLFTags",
        "lakeformation:ListLFTags",
        "lakeformation:GetLFTag",
        "lakeformation:SearchTablesByLFTags",
        "lakeformation:SearchDatabasesByLFTags"
      ],
      "Resource": "*"
    }
  ]
}
```

```
}
```

6. Ao terminar, escolha Review (Revisar) para revisar a política. O validador de política indica se há qualquer erro de sintaxe.
7. Na página Revisar política, insira o Nome da política que você está criando. Insira uma Description (Descrição) (opcional). Revise o Resumo da política para ver as permissões que são concedidas pela política. Em seguida, escolha Criar política para salvar seu trabalho.
8. No painel de navegação do console do IAM, escolha Roles (Perfis) e, em seguida, Create role (Criar perfil).
9. Em Select trusted entity (Selecionar entidade confiável), escolha AWS Service (Serviço).
10. Escolha o serviço do Amazon Redshift para assumir essa função.
11. Escolha o caso de uso Redshift Customizable (Personalizável pelo Redshift) para o serviço. Então, escolha Próximo: Permissões.
12. Procure a política de permissões que você criou, RedshiftLakeFormationPolicy e marque a caixa de seleção ao lado do nome da política na lista.
13. Escolha Next: Tags (Próximo: etiquetas).
14. Escolha Next: Review (Próximo: revisar).
15. Em Role name (Nome da função), insira o nome **RedshiftLakeFormationRole**.
16. (Opcional) Em Role description (Descrição da função), digite uma descrição para a nova função.
17. Reveja a função e escolha Create role (Criar função).

Para conceder permissões **Select** na tabela a ser consultada no banco de dados do Lake Formation

1. Abra o console do Lake Formation em <https://console.aws.amazon.com/lakeformation/>. Faça login como administrador de data lake.
2. No painel de navegação, em Permissões, escolha Permissões do Data Lake e, em seguida, escolha Conceder.
3. Forneça as informações a seguir:
  - Para usuários e funções do IAM, selecione o perfil do IAM criado, RedshiftLakeFormationRole. Ao executar o editor de consulta do Amazon Redshift, ele usa esse perfil do IAM para conceder permissão aos dados.
  - Para o Database (Banco de dados), escolha lakeformation\_tutorial.

A lista de tabelas é preenchida.

- Em Tabela, escolha uma tabela na fonte de dados a ser consultada.
- Escolha a permissão Selecionar tabela.

4. Escolha Grant (Conceder).

Para configurar o Amazon Redshift Spectrum e executar consultas

1. Abra o console do Amazon Redshift em <https://console.aws.amazon.com/redshift>. Faça login como usuário Administrator.
2. Selecione Create cluster (Criar cluster).
3. Na página Criar cluster, insira `redshift-lakeformation-demo` o Identificador do cluster.
4. Para o Tipo de nó, selecione `dc2.large`.
5. Role para baixo e, em Configurações do banco de dados, insira ou aceite estes parâmetros:
  - Nome de usuário administrador: `awsuser`
  - Senha do usuário administrador: (*Choose a password*)
6. Expanda as permissões do cluster e, para os perfis do IAM disponíveis, escolha `RedshiftLakeFormationRole`. Depois, escolha Add IAM role (Adicionar perfil do IAM).
7. Se você precisar usar uma porta diferente do valor padrão de 5439, ao lado de Configurações adicionais, desative a opção Usar padrões. Expanda a seção Configurações do banco de dados e, em seguida, insira um novo número de Porta do banco de dados.
8. Selecione Create cluster (Criar cluster).

A página Clusters é carregada.

9. Espere até que o status do cluster se torne Disponível. Escolha o ícone de atualização periodicamente.
10. Conceda permissão ao analista de dados para executar consultas no cluster. Para fazer isso, execute as etapas a seguir:
  - a. Abra o console do IAM em <https://console.aws.amazon.com/iam/> e faça login como usuário Administrator.
  - b. No painel de navegação, escolha Usuários e anexe as seguintes políticas gerenciadas ao usuário `dataLake_user`.

- AmazonRedshiftQueryEditor
- AmazonRedshiftReadOnlyAccess

11. Saia do console do Amazon Redshift e entre novamente como usuário `dataLake_user`.
12. Na barra de ferramentas vertical esquerda, escolha o ícone EDITOR para abrir o editor de consultas e conectar-se ao cluster. Se a caixa de diálogo Conectar ao banco de dados for exibida, escolha o nome do cluster `redshift-lakeformation-demo` e insira o nome do banco de dados `dev`, o nome do usuário `awsuser` e a senha que você criou. Então escolha Connect to database (Conectar-se ao banco de dados).

#### Note

Se os parâmetros de conexão não forem solicitados e outro cluster já estiver selecionado no editor de consultas, escolha Alterar conexão para abrir a caixa de diálogo Conectar ao banco de dados.

13. Na caixa de texto Nova consulta 1, insira e execute a seguinte instrução para mapear o banco de dados `lakeformation_tutorial` no Lake Formation para o nome do esquema Amazon Redshift `redshift_jdbc`:

#### Important

Substitua `<account-id>` por um número de conta válido AWS e `<region>` por um nome de região válido AWS (por exemplo, `us-east-1`).

```
create external schema if not exists redshift_jdbc from DATA CATALOG
  database 'lakeformation_tutorial' iam_role 'arn:aws:iam::<account-id>:role/
  RedshiftLakeFormationRole' region '<region>;'
```

14. Na lista de esquemas, em Selecionar esquema, escolha `redshift_jdbc`.

A lista de tabelas é preenchida. O editor de consultas mostra somente as tabelas nas quais você recebeu permissões de data lake do Lake Formation.

15. No menu pop-up ao lado do nome da tabela, escolha Visualizar dados.

O Amazon Redshift retorna as 10 primeiras linhas.

Agora você pode executar consultas nas tabelas e colunas para as quais você tem permissões.

## Etapa 13: conceder ou revogar permissões do Lake Formation usando o Amazon Redshift Spectrum

O Amazon Redshift oferece suporte à capacidade de conceder e revogar permissões do Lake Formation em bancos de dados e tabelas usando instruções SQL modificadas. Essas declarações são semelhantes às declarações existentes do Amazon Redshift. Para obter mais informações, consulte [CONCEDER](#) e [REVOGAR](#) no Guia do desenvolvedor de banco de dados do Amazon Redshift.

## Como configurar permissões para formatos de armazenamento de tabelas abertas no Lake Formation

O AWS Lake Formation suporta o gerenciamento de permissões de acesso para Open Table Formats (OTFs), como [Apache Iceberg](#), [Apache Hudi](#) e [Linux Foundation Delta Lake](#). Neste tutorial, você aprenderá como criar Iceberg, Hudi e Delta Lake com tabelas de [manifesto](#) de links simbólicos no AWS Glue Data Catalog usando AWS Glue, configurar permissões refinadas usando o Lake Formation e consultar dados usando o Amazon Athena.

### Note

Os serviços de análise da AWS não oferecem suporte a todos os formatos de tabela transacional. Para ter mais informações, consulte [Trabalhando com outros AWS serviços](#). Este tutorial aborda manualmente a criação de um novo banco de dados e uma tabela no Catálogo de Dados usando somente AWS Glue trabalhos.

Este tutorial inclui um AWS CloudFormation modelo para configuração rápida. É possível revisá-lo e personalizá-lo para atender às suas necessidades.

### Tópicos

- [Público-alvo](#)
- [Pré-requisitos](#)
- [Etapa 1: Provisionar os recursos](#)



- [Etapa 2: Configurar permissões para uma tabela do Iceberg](#)
- [Etapa 3: Configurar permissões para uma tabela do Hudi](#)
- [Etapa 4: Configurar permissões para uma tabela do Delta Lake](#)
- [Etapa 5: Limpar os recursos do AWS](#)

## Público-alvo

Este tutorial é destinado a administradores de IAM, administradores de data lake e analistas de negócios. A tabela a seguir lista os perfis usados neste tutorial para criar uma tabela controlada usando o Lake Formation.

Função	Descrição
Administrador do IAM	Um usuário que pode criar Usuários e perfis de IAM e buckets do Amazon S3. Usa a política AdministratorAccess gerenciada pelo AWS.
Administrador do data lake	Um usuário que pode acessar o Catálogo de Dados, criar bancos de dados e conceder permissões do Lake Formation a outros usuários. Tem menos permissões do IAM do que o administrador do IAM, mas o suficiente para administrar o data lake.
Analista de negócios	Um usuário que pode executar consultas no data lake. Tem permissões para executar consultas.

## Pré-requisitos

Antes de começar este tutorial, você deve ter um Conta da AWS que possa fazer login como usuário com as permissões corretas. Para obter mais informações, consulte [Inscrever-se para uma Conta da AWS](#) e [Criar um usuário administrativo](#).

O tutorial pressupõe que você esteja familiarizado com os perfis e políticas do IAM. Para obter informações sobre o IAM, consulte o [Guia do Usuário do IAM](#).

Você precisa configurar os seguintes recursos do AWS para concluir este tutorial:

- Usuário administrador do data lake
- Configurações do data lake Formation
- Versão 3 do mecanismo Amazon Athena

### Como criar um administrador de data lake

1. Faça login no console do Lake Formation em <https://console.aws.amazon.com/lakeformation/> como usuário administrador. Você criará recursos na região Leste dos EUA (Norte da Virgínia) para este tutorial.
2. No console do Lake Formation, no painel de navegação, em Permissões, selecione Perfis e tarefas administrativas..
3. Selecione Escolher administradores em Administradores de data lake.
4. Na janela Gerenciar administradores do data lake, em Usuários e perfis do IAM, selecione Usuário Administrador do IAM.
5. Escolha Salvar.

### Como ativar as configurações do data lake

1. Abra o console do Lake Formation em <https://console.aws.amazon.com/lakeformation/>. No painel de navegação, em Catálogo de dados, selecione Configurações. Desmarque o seguinte:
  - Use somente o controle de acesso do IAM para novos bancos de dados.
  - Use somente o controle de acesso do IAM para novas tabelas em novos bancos de dados.
2. Em Configurações da versão entre contas, selecione Versão 3 como a versão entre contas.
3. Escolha Salvar.

### Como atualizar o mecanismo Amazon Athena para a versão 3

1. Abra o console do Athena em <https://console.aws.amazon.com/athena/>.
2. Selecione o Grupo de trabalho e selecione o grupo de trabalho principal.

3. Certifique-se de que o grupo de trabalho tenha pelo menos a versão 3. Se não estiver, edite o grupo de trabalho, selecione Manual for Atualizar o mecanismo de consulta e selecione a versão 3.
4. Selecione Salvar alterações.

## Etapa 1: Provisionar os recursos

Esta seção mostra como configurar os recursos AWS usando um modelo AWS CloudFormation.

Como criar seus recursos usando o modelo AWS CloudFormation

1. Faça login no console do AWS CloudFormation em <https://console.aws.amazon.com/cloudformation> como administrador do IAM na região Leste dos EUA (Norte da Virgínia).
2. Selecione [Iniciar Pilha](#).
3. Na página Criar pilha, selecione Avançar.
4. Digite um Nome de pilha.
5. Selecione Avançar.
6. Na página seguinte, selecione Avançar.
7. Revise os detalhes na página final e selecione Eu reconheço que AWS CloudFormation pode criar recursos do IAM.
8. Selecione Criar.

A criação da pilha pode levar até dois minutos.

Ao iniciar a pilha de Cloud Formation, os seguintes recursos são criados:

- If-otf-datalake-123456789012 — Bucket Amazon S3 para armazenar dados

### Note

O ID da conta anexado ao nome do bucket do Amazon S3 é substituído pelo ID da sua conta.

- If-otf-tutorial-123456789012 — Bucket Amazon S3 para armazenar resultados de consultas e scripts de trabalho AWS Glue
- Ificebergdb — Banco de dados Iceberg do AWS Glue

- `lfhudidb` — Banco de dados Hudi do AWS Glue
- `lfdeltadb` — Banco de dados Delta do AWS Glue
- `native-iceberg-create` — AWS Glue trabalho que cria uma tabela Iceberg no Catálogo de Dados
- `native-hudi-create` — AWS Glue trabalho que cria uma tabela Hudi no Catálogo de Dados
- `native-delta-create` — AWS Glue trabalho que cria uma tabela Delta no catálogo de dados
- `LF-OTF- GlueServiceRole` — Função do IAM para a qual você passa AWS Glue para executar os trabalhos. Essa perfil tem as políticas necessárias anexadas para acessar os recursos, como o Catálogo de Dados, o bucket do Amazon S3, etc.
- `LF-OTF- RegisterRole` — Função do IAM para registrar a localização do Amazon S3 no Lake Formation. Esse perfil tem o `LF-Data-Lake-Storage-Policy` anexado ao perfil.
- `lf-consumer-analystuser` — Usuário do IAM para consultar os dados usando o Athena
- `lf-consumer-analystuser-credentials` — Senha para o usuário analista de dados armazenada em AWS Secrets Manager

Depois que as criações da pilha forem concluídas, navegue até a guia de saída e anote os valores para:

- `AthenaQueryResultLocation` — Localização do Amazon S3 para saída de consulta do Athena
- `BusinessAnalystUserCredentials` — Senha para o usuário analista de dados

Como recuperar o valor da senha:

1. Selecione o valor `lf-consumer-analystuser-credentials` navegando até o console do Secrets Manager.
2. Na seção Valor secreto, selecione Recuperar valor secreto.
3. Anote o valor secreto da senha.

## Etapa 2: Configurar permissões para uma tabela do Iceberg

Nesta seção, você aprenderá como criar uma tabela do Iceberg no AWS Glue Data Catalog, configurar permissões de dados no AWS Lake Formation e consultar dados usando o Amazon Athena.

## Como criar uma tabela no Iceberg

Nesta etapa, você executará um trabalho AWS Glue que cria uma tabela transacional do Iceberg no Catálogo de Dados.

1. Abra o console do AWS Glue em <https://console.aws.amazon.com/glue/> na região Leste dos EUA (Norte da Virgínia) como usuário administrador do data lake.
2. Selecione trabalhos no painel de navegação esquerdo.
3. Selecione `native-iceberg-create`.

The screenshot shows the AWS Glue console interface. The top section is titled 'Create job' and includes a 'Create' button. Below this, there are six options for creating a job, each with a radio button and a description:

- Visual with a source and target**: Start with a source, ApplyMapping transform, and target.
- Visual with a blank canvas**: Author using an interactive visual interface.
- Spark script editor**: Write or upload your own Spark code.
- Python Shell script editor**: Write or upload your own Python shell script.
- Jupyter Notebook**: Write your own code in a Jupyter Notebook for interactive development.
- Ray script editor** (New): Write your own code to run on Ray.

Below the options, there are two dropdown menus for 'Source' and 'Target', both set to 'Amazon S3'. The Source description is 'JSON, CSV, or Parquet files stored in S3.' and the Target description is 'S3 bucket by specifying a bucket path as the data target.'

The bottom section is titled 'Your jobs (24)' and includes a search bar and a table of jobs. The table has columns for 'Job name', 'Type', and 'Last modified'. The job 'native-iceberg-create' is selected, and its 'Actions' menu is open, showing options like 'Edit job', 'Clone job', 'Schedule job', 'Delete job(s)', and 'Reset job bookmark'.

Job name	Type	Last modified
<input type="checkbox"/> native-delta-create	Glue ETL	2/24/2023, 9:22:31 AM
<input checked="" type="checkbox"/> native-iceberg-create	Glue ETL	2/24/2023, 9:22:31 AM
<input type="checkbox"/> native-hudi-create	Glue ETL	2/24/2023, 9:22:30 AM

4. Em Ações, selecione Editar trabalho.
5. Em Informações do trabalho, expanda Propriedades avançadas, e marque a caixa ao lado de Usar AWS Glue Data Catalog como repositório do Hive para adicionar os metadados da tabela no AWS Glue Data Catalog. Isso especifica o AWS Glue Data Catalog como o repositório para os recursos do Catálogo de Dados usados no trabalho e habilita as permissões do Lake Formation a serem aplicadas posteriormente nos recursos do catálogo.
6. Escolha Salvar.
7. Escolha Executar. Você pode exibir o status da tarefa durante sua execução.

Para obter mais informações sobre os trabalhos AWS Glue, consulte [Trabalhos no console AWS Glue](#) no Guia do desenvolvedor AWS Glue.

Esse trabalho cria uma tabela no Iceberg nomeada como product banco de dados lfacebergdb. Verifique a tabela de produtos no console do Lake Formation.

## Como registrar o local dos dados com Lake Formation

Em seguida, registre o caminho do Amazon S3 como o local do seu data lake.

1. Abra o console do Lake Formation em <https://console.aws.amazon.com/lakeformation/> com o usuário administrador do data lake.
2. No painel de navegação, em Registrar e ingerir, selecione Local dos dados.
3. No canto superior direito do console, selecione Registrar local.
4. Na página Registrar local, digite o seguinte:
  - Caminho do Amazon S3 — selecione Navegar e selecione lf-otf-dataLake-123456789012. Clique na seta direita (>) ao lado do local raiz do Amazon S3 para navegar até o local s3/buckets/lf-otf-dataLake-123456789012/transactionaldata/native-iceberg.
  - Perfil do IAM — selecione LF-OTF-RegisterRole como perfil do IAM.
  - Selecione Registrar local.

## Register location

### Amazon S3 location

Register an Amazon S3 path as the storage location for your data lake.

#### Amazon S3 path

Choose an Amazon S3 path for your data lake.

 /transactionaldata/native-iceberg"/>

#### Review location permissions - strongly recommended

Registering the selected location may result in your users gaining access to data already at that location. Before registering a location, we recommend that you review existing location permissions on resources in that location.

#### IAM role

To add or update data, Lake Formation needs read/write access to the chosen Amazon S3 path. Choose a role that you know has permission to do this, or choose the **AWSServiceRoleForLakeFormationDataAccess** service-linked role. When you register the first Amazon S3 path, the service-linked role and a new inline policy are created on your behalf. Lake Formation adds the first path to the inline policy and attaches it to the service-linked role. When you register subsequent paths, Lake Formation adds the path to the existing policy.

 Enable Catalog Federation

Lake Formation will only assume a role to access a registered location when accessing a table under a federated database

Para obter mais informações sobre como registrar um local de dados no Lake Formation, consulte [Adicionar uma localização do Amazon S3 ao seu data lake](#).

### Como conceder permissões do Lake Formation na tabela do Iceberg

Nesta etapa, concederemos permissões de data lake ao usuário do analista de negócios.

1. Em Permissões do Data Lake, selecione Conceder.
2. Na tela Conceder permissões de dados, selecione Usuários e perfis do IAM.
3. Selecione `lf-consumer-analystuser` no menu suspenso.

## Principals

**IAM users and roles**  
Users or roles from this AWS account.

**SAML users and groups**  
SAML users and group or QuickSight ARNs.

**External accounts**  
AWS account, AWS organization or IAM principal outside of this account

**IAM users and roles**  
Add one or more IAM users or roles.

Choose IAM principals to add ▼

**lf-consumer-analystuser** ✕  
User

4. Selecione Recurso nomeado de Catálogo de Dados.
5. Para o Banco de dados, selecione lficebergdb.
6. Para Tabelas, selecione product.



### LF-Tags or catalog resources

Resources matched by LF-Tags (recommended)  
Manage permissions indirectly for resources or data matched by a specific set of LF-Tags.

Named data catalog resources  
Manager permissions for specific databases or tables, in addition to fine-grained data access.

**Databases**  
Select one or more databases.

Choose databases ▼

Load more

lforcebergdb ✕

**Tables - optional**  
Select one or more tables.

Choose tables ▼

Load more

product ✕

**Data filters - optional**  
Select one or more data filters.

Choose data filters ▼

Load more

Create new

[Manage data filters](#) ↗

7. Em seguida, você pode conceder acesso baseado em colunas especificando colunas.
  - a. Em Permissões de tabela, clique em Selecionar.
  - b. Em Permissões de dados, selecione Acesso baseado em colunas, selecione Incluir colunas.
  - c. Selecione as colunas `product_name`, `price` e `category`.
  - d. Selecione Conceder.

### Table permissions

**Table permissions**  
Choose specific access permissions to grant.

Select     Insert     Delete  
 Describe     Alter     Drop

**Grantable permissions**  
Choose the permission that may be granted to others.

Select     Insert     Delete  
 Describe     Alter     Drop

Super  
This permission is the union of all the individual permissions to the left, and supersedes them.

Super  
This permission allows the principal to grant any of the permissions to the left, and supersedes those grantable permissions.

---

### Data permissions

All data access  
Grant access to all data without any restrictions.

Column-based access  
Grant data access to specific columns only.

**Choose permission filter**  
Choose whether to include or exclude columns.

Include columns  
Grant permissions to access specific columns.

Exclude columns  
Grant permissions to access all but specific columns.

**Select columns**

Choose one or more columns ▼

product\_name ×    price ×    category ×  
string    bigint    string

**Cancel**    **Grant**

Para consultar a tabela do Iceberg usando o Athena

Agora você pode começar a consultar a tabela do Iceberg que você criou usando o Athena. Se for a primeira vez que você executa consultas no Athena, você precisará configurar a localização do resultado da consulta. Para obter mais informações, acesse [Especificação de um local de resultado de consulta](#).

1. Saia como usuário administrador do data lake e faça login como `lf-consumer-analystuser` na região Leste dos EUA (Norte da Virgínia) usando a senha anotada anteriormente na AWS CloudFormation saída.
2. Abra o console do Athena em <https://console.aws.amazon.com/athena/>.
3. Selecione Configurações e selecione Gerenciar.
4. Na caixa Local do resultado da consulta, insira o caminho para o bucket que você criou nas saídas AWS CloudFormation. Copie o valor de **AthenaQueryResultLocation** (`s3://lf-otf-tutorial-123456789012/athena-results/`) e escolha Salvar.
5. Execute a consulta a seguir para exibir 10 registros armazenados na tabela do Iceberg:

```
select * from lficebergdb.product limit 10;
```

Para obter mais informações sobre como consultar tabelas do Iceberg usando o Athena, consulte [Como consultar tabelas do Iceberg](#) no Guia do usuário do Amazon Athena.

## Etapa 3: Configurar permissões para uma tabela do Hudi

Nesta seção, você aprenderá como criar uma tabela do Hudi no AWS Glue Data Catalog, configurar permissões de dados no AWS Lake Formation e consultar dados usando o Amazon Athena.

### Como criar uma tabela Hudi

Nesta etapa, você executará um trabalho AWS Glue que cria uma tabela transacional do Hudi no Catálogo de Dados.

1. Faça login no AWS Glue console em <https://console.aws.amazon.com/glue/> na região Leste dos EUA (Norte da Virgínia)  
  
como o usuário administrador do data lake.
2. Selecione trabalhos no painel de navegação esquerdo.
3. Selecione `native-hudi-create`.
4. Em Ações, selecione Editar trabalho.
5. Em Informações do trabalho, expanda Propriedades avançadas, e marque a caixa ao lado de Usar AWS Glue Data Catalog como repositório do Hive para adicionar os metadados da tabela no AWS Glue Data Catalog. Isso especifica o AWS Glue Data Catalog como o repositório

para os recursos do Catálogo de Dados usados no trabalho e habilita as permissões do Lake Formation a serem aplicadas posteriormente nos recursos do catálogo.

6. Escolha Salvar.
7. Escolha Executar. Você pode exibir o status da tarefa durante sua execução.

Para obter mais informações sobre os trabalhos AWS Glue, consulte [Trabalhos no console AWS Glue](#) no Guia do desenvolvedor AWS Glue.

Esse trabalho cria uma tabela do Hudi (COW) no banco de dados: lfhudidb. Verifique a tabela no console product do Lake Formation.

### Como registrar o local dos dados com Lake Formation

Em seguida, registre um caminho do Amazon S3 como o local raiz do data lake.

1. Faça login no console do Lake Formation em <https://console.aws.amazon.com/lakeformation/> como usuário administrador do data lake.
2. No painel de navegação, em Registrar e ingerir, selecione Local dos dados.
3. No canto superior direito do console, selecione Registrar local.
4. Na página Registrar local, digite o seguinte:
  - Caminho do Amazon S3 — selecione Navegar e selecione lf-otf-dataLake-123456789012. Clique na seta direita (>) ao lado do local raiz do Amazon S3 para navegar até o local s3/buckets/lf-otf-dataLake-123456789012/transactionaldata/native-hudi.
  - Perfil do IAM — selecione LF-OTF-RegisterRole como perfil do IAM.
  - Selecione Registrar local.

### Como conceder permissões de data lake na tabela do Hudi

Nesta etapa, concederemos permissões de data lake ao usuário do analista de negócios.

1. Em Permissões do Data Lake, selecione Conceder.
2. Na tela Conceder permissões de dados, selecione Usuários e perfis do IAM.
3. lf-consumer-analystuser do menu suspenso.
4. Selecione Recurso nomeado de Catálogo de Dados.

5. Para o Banco de dados, selecione `lfhudidb`.
6. Para Tabelas, selecione `product`.
7. Em seguida, você pode conceder acesso baseado em colunas especificando colunas.
  - a. Em Permissões de tabela, clique em Selecionar.
  - b. Em Permissões de dados, selecione Acesso baseado em colunas, selecione Incluir colunas.
  - c. Selecione as colunas `product_name`, `price` e `category`.
  - d. Selecione Conceder.

### Como consultar a tabela do Hudi usando o Athena

Agora comece a consultar a tabela do Hudi que você criou usando o Athena. Se for a primeira vez que você executa consultas no Athena, você precisará configurar a localização do resultado da consulta. Para obter mais informações, acesse [Especificação de um local de resultado de consulta](#).

1. Saia como usuário administrador do data lake e faça login como `lf-consumer-analystuser` na região Leste dos EUA (Norte da Virgínia) usando a senha anotada anteriormente na AWS CloudFormation saída.
2. Abra o console do Athena em <https://console.aws.amazon.com/athena/>.
3. Selecione Configurações e selecione Gerenciar.
4. Na caixa Local do resultado da consulta, insira o caminho para o bucket que você criou nas saídas AWS CloudFormation. Copie o valor de **AthenaQueryResultLocation** (`s3://lf-otf-tutorial-123456789012/athena-results/`) e salve.
5. Execute a consulta a seguir para exibir 10 registros armazenados na tabela do Hudi:

```
select * from lfhudidb.product limit 10;
```

Para obter mais informações sobre como consultar tabelas do Hudi usando o Athena, consulte [Como consultar tabelas do Hudi](#) no Guia do usuário do Amazon Athena.

## Etapa 4: Configurar permissões para uma tabela do Delta Lake

Nesta seção, você aprenderá como criar uma tabela do Delta Lake com arquivo de manifesto de link simbólico no AWS Glue Data Catalog, configurar permissões de dados no AWS Lake Formation e consultar dados usando o Amazon Athena.

## Como criar uma tabela do Delta Lake

Nesta etapa, você executará um trabalho AWS Glue que cria uma tabela transacional do Delta Lake no Catálogo de Dados.

1. Faça login no AWS Glue console em <https://console.aws.amazon.com/glue/> na região Leste dos EUA (Norte da Virgínia)  
  
como o usuário administrador do data lake.
2. Selecione trabalhos no painel de navegação esquerdo.
3. Selecione `native-delta-create`.
4. Em Ações, selecione Editar trabalho.
5. Em Informações do trabalho, expanda Propriedades avançadas, e marque a caixa ao lado de Usar AWS Glue Data Catalog como repositório do Hive para adicionar os metadados da tabela no AWS Glue Data Catalog. Isso especifica o AWS Glue Data Catalog como o repositório para os recursos do Catálogo de Dados usados no trabalho e habilita as permissões do Lake Formation a serem aplicadas posteriormente nos recursos do catálogo.
6. Escolha Salvar.
7. Selecione Executar em Ações.

Esse trabalho cria uma tabela no Delta Lake nomeada como `product` banco de dados `1fdeltadb`. Verifique a tabela no console `product` do Lake Formation.

## Como registrar o local dos dados com Lake Formation

Em seguida, registre o caminho do Amazon S3 como o local raiz do data lake.

1. Abra o console do Lake Formation em <https://console.aws.amazon.com/lakeformation/> com o usuário administrador do data lake.
2. No painel de navegação, em Registrar e ingerir, selecione Local dos dados.
3. No canto superior direito do console, selecione Registrar local.
4. Na página Registrar local, digite o seguinte:
  - Caminho do Amazon S3 — selecione Navegar e selecione `1f-otf-datalake-123456789012`. Clique na seta direita (>) ao lado do local raiz do Amazon S3 para navegar até o local `s3/buckets/1f-otf-datalake-123456789012/transactionaldata/native-delta`.

- Perfil do IAM — selecione `LF-OTF-RegisterRole` como perfil do IAM.
- Selecione Registrar local.

## Como conceder permissões de data lake na tabela do Delta Lake

Nesta etapa, concederemos permissões de data lake ao usuário do analista de negócios.

1. Em Permissões do Data Lake, selecione Conceder.
2. Na tela Conceder permissões de dados, selecione Usuários e perfis do IAM.
3. `lf-consumer-analystuser` do menu suspenso.
4. Selecione Recurso nomeado de Catálogo de Dados.
5. Para o Banco de dados, selecione `lfdeltadb`.
6. Para Tabelas, selecione `product`.
7. Em seguida, você pode conceder acesso baseado em colunas especificando colunas.
  - a. Em Permissões de tabela, clique em Selecionar.
  - b. Em Permissões de dados, selecione Acesso baseado em colunas, selecione Incluir colunas.
  - c. Selecione as colunas `product_name`, `price` e `category`.
  - d. Selecione Conceder.

## Como consultar a tabela do Delta Lake usando o Athena

Agora comece a consultar a tabela do Delta Lake que você criou usando o Athena. Se for a primeira vez que você executa consultas no Athena, você precisará configurar a localização do resultado da consulta. Para obter mais informações, acesse [Especificação de um local de resultado de consulta](#).

1. Saia como usuário administrador do data lake e faça login como `BusinessAnalystUser` na região Leste dos EUA (Norte da Virgínia) usando a senha anotada anteriormente na saída AWS CloudFormation.
2. Abra o console do Athena em <https://console.aws.amazon.com/athena/>.
3. Selecione Configurações e selecione Gerenciar.
4. Na caixa Local do resultado da consulta, insira o caminho para o bucket que você criou nas saídas AWS CloudFormation. Copie o valor de **AthenaQueryResultLocation** (`s3://lf-otf-tutorial-123456789012/athena-results/`) e salve.
5. Execute a consulta a seguir para exibir 10 registros armazenados na tabela do Delta Lake:

```
select * from lfdeltadb.product limit 10;
```

Para obter mais informações sobre como consultar tabelas do Delta Lake usando o Athena, consulte [Como consultar tabelas do Delta Lake](#) no Guia do usuário do Amazon Athena.

## Etapa 5: Limpar os recursos do AWS

### Como limpar recursos

Para evitar cobranças indesejadasConta da AWS, exclua AWS os recursos que você usou neste tutorial.

1. Faça login no AWS CloudFormation console em <https://console.aws.amazon.com/cloudformation> como administrador do IAM.
2. [Exclua a pilha do Cloud Formation](#). As tabelas que você criou são excluídas automaticamente com a pilha.

## Como gerenciar um data lake usando o controle de acesso baseado em tags do Lake Formation

Milhares de clientes estão criando data lakes em escala de petabytes no AWS. Muitos desses clientes usam AWS Lake Formation para criar e compartilhar facilmente seus data lakes em toda a organização. À medida que o número de tabelas e usuários aumenta, administradores de dados procuram maneiras de gerenciar facilmente as permissões em data lakes em grande escala. O controle de acesso baseado em Lake Formation tags (LF-TBAC) resolve essa questão, permitindo que administradores de dados criem LF-tags (com base em sua classificação e ontologia de dados) que podem ser anexadas aos recursos.

O LF-TBAC é uma estratégia de autorização que define permissões com base em atributos. No Lake Formation, esses atributos são chamados de LF-tags. Você pode anexar LF-tags aos recursos do Catálogo de Dados e às entidades principais do Lake Formation. Administradores do Data Lake podem atribuir e revogar permissões nos recursos do Lake Formation usando LF-tags. Para obter mais informações, consulte [Controle de acesso baseado em tags do Lake Formation](#).

Este tutorial demonstra como criar uma política de controle de acesso baseada em tags do Lake Formation usando um conjunto de dados público do AWS. Além disso, mostra como consultar



tabelas, bancos de dados e colunas que têm políticas de acesso baseadas em tags do Lake Formation associadas a eles.

Você pode usar o LF-TBAC para os seguintes casos:

- Você tem um grande número de tabelas e princípios que o administrador do data lake precisa conceder acesso.
- Você deseja classificar seus dados com base em uma ontologia e conceder permissões com base na classificação
- O administrador do data lake deseja atribuir permissões dinamicamente, com acoplamento fraco

A seguir estão as etapas de alto nível para configurar as permissões usando o LF-TBAC:

1. O administrador de dados define a ontologia da tag com duas LF-tags: `Confidential` e `Sensitive`. Os dados com `Confidential=True` têm controles de acesso mais rígidos. Os dados com `Sensitive=True` requerem uma análise específica do analista.
2. O administrador de dados atribui diferentes níveis de permissão ao engenheiro de dados para criar tabelas com diferentes LF-tags.
3. O engenheiro de dados cria dois bancos de dados: `tag_database` e `col_tag_database`. Todas as tabelas em `tag_database` são configuradas com `Confidential=True`. Todas as tabelas do `col_tag_database` são configuradas com `Confidential=False`. Algumas colunas da tabela `col_tag_database` estão marcadas com `Sensitive=True` para necessidades específicas de análise.
4. O engenheiro de dados concede permissão de leitura ao analista para tabelas com condições de expressão específicas `Confidential=True` e `Confidential=False, Sensitive=True`.
5. Com essa configuração, o analista de dados pode se concentrar em realizar análises com os dados certos.

## Tópicos

- [Público-alvo](#)
- [Pré-requisitos](#)
- [Etapa 1: Provisionar os recursos](#)
- [Etapa 2: registre sua localização de dados, crie uma ontologia LF-Tag e conceda permissões](#)
- [Etapa 3: Criar bancos de dados do Lake Formation](#)

- [Etapa 4: Conceder permissões de dados](#)
- [Etapa 5: Executar uma consulta no Amazon Athena para verificar as permissões](#)
- [Etapa 6: Limpar os recursos do AWS](#)

## Público-alvo

Este tutorial é destinado a administradores de dados, engenheiros de dados e analistas de dados. Quando se trata de gerenciar AWS Glue Data Catalog e administrar permissões no Lake Formation, os administradores de dados nas contas produtoras têm propriedade funcional com base nos perfis que suportam e podem conceder acesso a vários consumidores, organizações externas e contas.

A tabela a seguir lista os perfis usados neste tutorial:

Perfil	Descrição
Administrador de dados (administrador)	<p>O usuário <code>lf-data-steward</code> tem o seguinte acesso:</p> <ul style="list-style-type: none"><li>• Acesso de leitura a todos os recursos do Catálogo de Dados</li><li>• Pode criar LF-tags e associar-se ao perfil de engenheiro de dados para conceder permissão a outras entidades principais</li></ul>
Engenheiro de dados	<p>O usuário <code>lf-data-engineer</code> tem o seguinte acesso:</p> <ul style="list-style-type: none"><li>• Acesso completo de leitura, gravação e atualização a todos os recursos do Catálogo de Dados</li><li>• Permissões de local de dados no data lake</li><li>• Pode associar LF-tags e associar-se ao Catálogo de Dados</li><li>• Pode anexar LF-tags aos recursos, fornecendo acesso às entidades principais</li></ul>

Perfil	Descrição
	com base em quaisquer políticas criadas por administradores de dados
Analista de dados	O usuário <code>lf-data-analyst</code> tem o seguinte acesso: <ul style="list-style-type: none"><li>• Acesso refinado aos recursos compartilhados pelas políticas de acesso baseadas em tags do Lake Formation</li></ul>

## Pré-requisitos

Antes de iniciar este tutorial, você deve ter uma conta Conta da AWS que possa ser usada para fazer login como um usuário administrativo com as permissões corretas. Para ter mais informações, consulte [Concluir as tarefas iniciais de configuração do AWS](#).

O tutorial pressupõe que você esteja familiarizado com o IAM. Para obter informações sobre o IAM, consulte o [Guia do Usuário do IAM](#).

## Etapa 1: Provisionar os recursos

Este tutorial inclui um modelo AWS CloudFormation para uma configuração rápida. É possível revisá-lo e personalizá-lo para atender às suas necessidades. O modelo cria três funções diferentes (listadas em [Público-alvo](#)) para realizar esse exercício e copia o `nyc-taxi-data` conjunto de dados para seu bucket local do Amazon S3.

- Um bucket do Amazon S3
- Configurações apropriadas do Lake Formation
- Recursos apropriados do Amazon EC2
- Três perfis do IAM com credenciais

### Criar seus recursos

1. Fazer login no console do AWS CloudFormation em <https://console.aws.amazon.com/cloudformation> na região Leste dos EUA (Norte da Virgínia).

2. Selecione [Iniciar Pilha](#).
3. Selecione Avançar.
4. Na seção Configuração do usuário, digite a senha para três perfis: `DataStewardUserPassword`, `DataEngineerUserPassword` e `DataAnalystUserPassword`.
5. Revise os detalhes na página final e selecione Eu reconheço que AWS CloudFormation pode criar recursos do IAM.
6. Selecione Criar.

A criação da pilha pode levar até cinco minutos.

#### Note

Após concluir o tutorial, talvez você queira excluir a pilha AWS CloudFormation para evitar que continuem incorrendo em cobranças. Verifique se os recursos foram excluídos com sucesso no status de eventos da pilha.

## Etapa 2: registre sua localização de dados, crie uma ontologia LF-Tag e conceda permissões

Nesta etapa, o usuário administrador de dados define a ontologia de tags com duas tags LF: `Confidential` e `Sensitive`, e fornece aos diretores específicos do IAM a capacidade de anexar tags LF recém-criadas aos recursos.

Registre uma localização de dados e defina a ontologia LF-Tag

1. Execute a primeira etapa como usuário administrador de dados (`lf-data-steward`) para verificar os dados no Amazon S3 e no Catálogo de Dados no Lake Formation.
  - a. Faça login no console do Lake Formation em <https://console.aws.amazon.com/lakeformation/> `lf-data-steward` com a senha usada ao implantar a AWS CloudFormation pilha.
  - b. No painel de navegação, em Permissões, selecione Perfis e tarefas administrativas.
  - c. Escolha Adicionar na seção Administradores do Data Lake.

- d. Na página Adicionar administrador, para usuários e funções do IAM, escolha o usuário `lf-data-steward`.
  - e. selecione Salvar para adicionar `lf-data-steward` como administrador do Lake Formation.
2. Em seguida, atualize as configurações do Catálogo de Dados para usar a permissão do Lake Formation para controlar os recursos do catálogo em vez do controle de acesso baseado no IAM.
  - a. No painel de navegação, em Administração, selecione Configurações do Catálogo de Dados.
  - b. Desmarque Usar somente o controle de acesso do IAM para novos bancos de dados.
  - c. Desmarque Usar somente o controle de acesso do IAM para novas tabelas em novos bancos de dados.
  - d. Clique em Salvar.
3. Em seguida, registre o local dos dados para o data lake.
  - a. No painel de navegação, em Administração em Locais de data lake.
  - b. Selecione Registrar local.
  - c. Na página Registrar localização, para o caminho do Amazon S3, insira. `s3://lf-tagbased-demo-Account-ID`
  - d. Para o Perfil do IAM, deixe o valor padrão `AWSServiceRoleForLakeFormationDataAccess` como está.
  - e. Escolha Lake Formation como modo de permissão.
  - f. Selecione Registrar local.
4. Em seguida, crie a ontologia definindo uma LF-tag.
  - a. Em Permissões no painel de navegação, escolha LF-Tags e permissões. .
  - b. Selecione Adicionar LF-tag.
  - c. Em Chave, digite `Confidential`.
  - d. Para Valores, adicione `True` e `False`.
  - e. Selecione Adicionar LF-tag.
  - f. Repita as etapas para criar a etiqueta LF Sensitive com o valor. `True`

Você criou todas as etiquetas LF necessárias para este exercício.

## Conceda permissões a usuários do IAM

1. Em seguida, forneça às entidades principais específicas do IAM a capacidade de anexar LF-tags recém-criadas aos recursos.
  - a. Em Permissões no painel de navegação, escolha LF-Tags e permissões.
  - b. Na seção Permissões do LF-Tag, escolha Conceder permissões.
  - c. Em Tipo de permissão, escolha permissões do par chave-valor da tag LF.
  - d. Selecione Usuários e perfis do IAM.
  - e. Para Usuários e perfis do IAM, pesquise e selecione o perfil `lf-data-engineer`.
  - f. Na seção LF-Tags, adicione a chave `Confidential` com valores `True` e `False`, e a key `Sensitive` com valor `True`.
  - g. Em Permissões, selecione Descrever e associar para permissões e permissões concedidas.
  - h. Selecione Conceder.
2. Em seguida, conceda permissões de `lf-data-engineer` para criar bancos de dados em nosso Catálogo de Dados e no bucket subjacente do Amazon S3 criado por AWS CloudFormation.
  - a. Em Administração no painel de navegação, escolha Funções e tarefas administrativas.
  - b. Na seção Criadores de banco de dados, selecione Conceder.
  - c. Para Usuários e perfis do IAM, selecione o perfil `lf-data-engineer`.
  - d. Para Permissões de catálogo, selecione Criar banco de dados.
  - e. Selecione Conceder.
3. Em seguida, conceda permissões no bucket do Amazon S3 do (`s3://lf-tagbased-demo-Account-ID`) ao usuário `lf-data-engineer`.
  - a. No painel de navegação, em Permissões, selecione Locais de dados.
  - b. Selecione Conceder.
  - c. Selecione Minha conta.
  - d. Para Usuários e perfis do IAM, selecione o perfil `lf-data-engineer`.

- e. Para Locais de armazenamento, digite o bucket do Amazon S3 criado pelo AWS CloudFormation modelo (`s3://lf-tagbased-demo-Account-ID`).
  - f. Selecione Conceder.
4. Em seguida, `lf-data-engineer` conceda permissões concedidas aos recursos associados à expressão LF-tag. `Confidential=True`
- a. No painel de navegação, em Permissões, escolha Permissões do data lake.
  - b. Selecione Conceder.
  - c. Selecione Usuários e perfis do IAM.
  - d. Selecione o perfil de `lf-data-engineer`.
  - e. Na seção LF-Tags ou recursos do catálogo, selecione Recursos correspondentes às LF-Tags.
  - f. Escolha Adicionar par de valores-chave da etiqueta LF.
  - g. Adicione a chave `Confidential` com os valores `True`.
  - h. Na seção Permissões do banco de dados, selecione Descrever para Permissões de banco de dados e Permissões concedíveis.
  - i. Na seção Permissões de tabela, selecione Descrever, Selecionar e Alterar para permissões de tabela e permissões concedidas.
  - j. Selecione Conceder.
5. Em seguida, `lf-data-engineer` conceda permissões concedidas aos recursos associados à expressão LF-tag. `Confidential=False`
- a. No painel de navegação, em Permissões, escolha Permissões do data lake.
  - b. Selecione Conceder.
  - c. Selecione Usuários e perfis do IAM.
  - d. Selecione o perfil de `lf-data-engineer`.
  - e. Selecione Recursos correspondentes às LF-tags.
  - f. Selecione Adicionar LF-tag.
  - g. Adicione a chave `Confidential` com o valor `False`.
  - h. Na seção Permissões do banco de dados, selecione Descrever para Permissões de banco de dados e Permissões concedíveis.
  - i. Na seção Permissões de tabela e coluna, não selecione nada.
  - j. Selecione Conceder.

6. Em seguida, `lf-data-engineer` concedemos permissões concedidas aos recursos associados aos pares de valores-chave da tag LF e. `Confidential=False Sensitive=True`
  - a. No painel de navegação, em Permissões, escolha Permissões de dados.
  - b. Selecione Conceder.
  - c. Selecione Usuários e perfis do IAM.
  - d. Selecione o perfil de `lf-data-engineer`.
  - e. Na seção LF-Tags ou recursos do catálogo, selecione Recursos correspondentes às LF-Tags.
  - f. Selecione Adicionar LF-tag.
  - g. Adicione a chave `Confidential` com o valor `False`.
  - h. Escolha Adicionar par de valores-chave da etiqueta LF.
  - i. Adicione a chave `Sensitive` com o valor `True`.
  - j. Na seção Permissões do banco de dados, selecione Descrever para Permissões de banco de dados e Permissões concedíveis.
  - k. Na seção Permissões de tabela, selecione Descrever, Selecionar e Alterar para permissões de tabela e permissões concedidas.
  - l. Selecione Conceder.

## Etapa 3: Criar bancos de dados do Lake Formation

Nesta etapa, você cria dois bancos de dados e anexa tags LF aos bancos de dados e colunas específicas para fins de teste.

Crie seus bancos de dados e sua tabela para acesso em nível de banco de dados

1. Primeiro, crie o banco de dados `tag_database`, a tabela `source_data` e anexe as tags LF apropriadas.
  - a. No console do Lake Formation (<https://console.aws.amazon.com/lakeformation/>), em Catálogo de dados, escolha Bancos de dados.
  - b. Selecione Criar banco de dados.
  - c. Em Nome, digite `tag_database`.
  - d. Em Local, digite o local do Amazon S3 criado pelo AWS CloudFormation modelo (`s3://lf-tagbased-demo-Account-ID/tag_database/`).



- e. Desmarque Usar somente controle de acesso do IAM para novas tabelas nesse banco de dados.
  - f. Selecione Criar banco de dados.
2. Em seguida, crie uma nova tabela dentro dela com tag\_database.
- a. Na página Bancos de dados, selecione o banco de dados tag\_database.
  - b. Selecione Exibir tabelas e clique em Criar tabela.
  - c. Em Nome, digite source\_data.
  - d. Em Banco de dados, selecione o banco de dados tag\_database.
  - e. Em Formato de tabela, escolha AWS GlueTabela padrão.
  - f. Em Dados localizados em, selecione Caminho especificado em minha conta.
  - g. Em Incluir caminho, digite o caminho para tag\_database criado pelo modelo AWS CloudFormation (s3://lf-tagbased-demo*Account-ID*/tag\_database/).
  - h. Em Formato de dados, selecione CSV.
  - i. Em Esquema de upload, digite a seguinte matriz JSON da estrutura da coluna para criar um esquema:

```
[
  {
    "Name": "vendorid",
    "Type": "string"
  },
  {
    "Name": "lpep_pickup_datetime",
    "Type": "string"
  },
  {
    "Name": "lpep_dropoff_datetime",
    "Type": "string"
  },
  {
    "Name": "store_and_fwd_flag",
    "Type": "string"
  },
  {
    "Name": "ratecodeid",
    "Type": "string"
  }
]
```

```
    },
    {
      "Name": "pulocationid",
      "Type": "string"
    },
    {
      "Name": "dolocationid",
      "Type": "string"
    },
    {
      "Name": "passenger_count",
      "Type": "string"
    },
    {
      "Name": "trip_distance",
      "Type": "string"
    },
    {
      "Name": "fare_amount",
      "Type": "string"
    },
    {
      "Name": "extra",
      "Type": "string"
    },
    {
      "Name": "mta_tax",
      "Type": "string"
    },
    {
      "Name": "tip_amount",
      "Type": "string"
    },
    {
      "Name": "tolls_amount",
```

```
        "Type": "string"
    },
    {
        "Name": "ehail_fee",
        "Type": "string"
    },
    {
        "Name": "improvement_surcharge",
        "Type": "string"
    },
    {
        "Name": "total_amount",
        "Type": "string"
    },
    {
        "Name": "payment_type",
        "Type": "string"
    }
]
```

- j. Selecione Carregar. Após fazer o upload do esquema, o esquema da tabela deve ter a aparência da seguinte captura de tela:

#	Column Name	▼	Data type
1	vendorid		string
2	lpep_pickup_datetime		string
3	lpep_dropoff_datetime		string
4	store_and_fwd_flag		string
5	ratecodeid		string
6	pulocationid		string
7	dolocationid		string
8	passenger_count		string
9	trip_distance		string
10	fare_amount		string
11	extra		string
12	mta_tax		string
13	tip_amount		string
14	tolls_amount		string
15	ehail_fee		string
16	improvement_surcharge		string
17	total_amount		string
18	payment_type		string

- k. Selecione Enviar.
3. Em seguida, anexe tags LF no nível do banco de dados.
    - a. Na página Bancos de dados, localize e selecione `tag_database`.
    - b. No menu Ações, escolha Editar tags LF.
    - c. Selecione Atribuir nova LF-tag.
    - d. Para Chaves atribuídas, escolha a `Confidential` etiqueta LF que você criou anteriormente.
    - e. Em Valores, selecione `True`.
    - f. Selecione Salvar.

Isso conclui a atribuição da tag LF ao banco de dados `tag_database`.

Crie seu banco de dados e tabela para acesso em nível de coluna

Repita as etapas a seguir para criar o banco de dados `col_tag_database` e a tabela `source_data_col_lvl1` e anexar tags LF no nível da coluna.

1. Na página Bancos de dados, selecione Criar banco de dados.
2. Em Nome, digite `col_tag_database`.
3. Em Local, digite o local do Amazon S3 criado pelo AWS CloudFormation modelo (`s3://lf-tagbased-demo-Account-ID/col_tag_database/`).
4. Desmarque Usar somente controle de acesso do IAM para novas tabelas nesse banco de dados.
5. Selecione Criar banco de dados.
6. Na página Bancos de dados, selecione seu novo banco de dados (`col_tag_database`).
7. Escolha Exibir tabelas e clique em Criar tabela.
8. Em Nome, digite `source_data_col_lvl1`.
9. Em Banco de dados, selecione seu novo banco de dados (`col_tag_database`).
10. Em Formato de tabela, escolha AWS GlueTabela padrão.
11. Em Dados localizados em, selecione Caminho especificado em minha conta.
12. Digite o caminho do Amazon S3 para `col_tag_database` (`s3://lf-tagbased-demo-Account-ID/col_tag_database/`).

13. Em Formato de dados, selecione CSV.
14. Em Upload schema, digite o seguinte esquema JSON:

```
[
  {
    "Name": "vendorid",
    "Type": "string"
  },
  {
    "Name": "lpep_pickup_datetime",
    "Type": "string"
  },
  {
    "Name": "lpep_dropoff_datetime",
    "Type": "string"
  },
  {
    "Name": "store_and_fwd_flag",
    "Type": "string"
  },
  {
    "Name": "ratecodeid",
    "Type": "string"
  },
  {
    "Name": "pulocationid",
    "Type": "string"
  },
  {
    "Name": "dolocationid",
    "Type": "string"
  }
]
```

```
    },  
    {  
      "Name": "passenger_count",  
      "Type": "string"  
    },  
    {  
      "Name": "trip_distance",  
      "Type": "string"  
    },  
    {  
      "Name": "fare_amount",  
      "Type": "string"  
    },  
    {  
      "Name": "extra",  
      "Type": "string"  
    },  
    {  
      "Name": "mta_tax",  
      "Type": "string"  
    },  
    {  
      "Name": "tip_amount",  
      "Type": "string"  
    },  
    {  
      "Name": "tolls_amount",  
      "Type": "string"  
    }
```

```
    },  
    {  
      "Name": "ehail_fee",  
      "Type": "string"  
    },  
    {  
      "Name": "improvement_surcharge",  
      "Type": "string"  
    },  
    {  
      "Name": "total_amount",  
      "Type": "string"  
    },  
    {  
      "Name": "payment_type",  
      "Type": "string"  
    }  
  ]
```

15. Selecione Upload. Após fazer o upload do esquema, o esquema da tabela deve ter a aparência da captura de tela a seguir.



#	Column Name	▼	Data type
1	vendorid		string
2	lpep_pickup_datetime		string
3	lpep_dropoff_datetime		string
4	store_and_fwd_flag		string
5	ratecodeid		string
6	pulocationid		string
7	dolocationid		string
8	passenger_count		string
9	trip_distance		string
10	fare_amount		string
11	extra		string
12	mta_tax		string
13	tip_amount		string
14	tolls_amount		string
15	ehail_fee		string
16	improvement_surcharge		string
17	total_amount		string
18	payment_type		string

16. Selecione Enviar para concluir a criação da tabela.
17. Agora, associe a `Sensitive=True` tag LF às colunas `e.vendorid` e `fare_amount`
  - a. Na página Tabelas, selecione a tabela que você criou (`source_data_col_lvl1`).
  - b. No menu Ações, escolha Esquema.
  - c. Selecione a coluna `vendorid` e escolha Editar tags LF.
  - d. Em Chaves atribuídas, selecione Sensível.
  - e. Em Valores, selecione Verdadeiro.
  - f. Escolha Salvar.
18. Em seguida, associe a `Confidential=False` etiqueta LF a `col_tag_database`. Isso é necessário para `lf-data-analyst` para poder descrever o banco de dados `col_tag_database` quando conectado. Amazon Athena
  - a. Na página Bancos de dados, localize e selecione `col_tag_database`.
  - b. No menu Ações, escolha Editar tags LF.
  - c. Escolha Atribuir nova LF-Tag.
  - d. Em Chaves atribuídas, escolha a `Confidential` etiqueta LF que você criou anteriormente.
  - e. Em Valores, selecione `False`.
  - f. Selecione Salvar.

## Etapa 4: Conceder permissões de dados

Conceda permissões aos analistas de dados para o consumo dos bancos de dados `tag_database` e da tabela `col_tag_database` usando as LF-tags `Confidential` e `Sensitive`.

1. Siga estas etapas para conceder permissões ao `lf-data-analyst` usuário nos objetos associados à tag LF `Confidential=True` (`Database:tag_database`) para ter `Describe` o banco de dados e a permissão nas tabelas. `Select`
  - a. Faça login no console do Lake Formation em <https://console.aws.amazon.com/lakeformation/> como `lf-data-engineer`.
  - b. Em Permissões, selecione Permissões do Data Lake.
  - c. Selecione Conceder.
  - d. Em Entidades principais, selecione Usuários e perfis do IAM.

- e. Para Usuários e perfis do IAM, selecione `lf-data-analyst`.
  - f. Em Etiquetas LF ou recursos do catálogo, selecione Recursos correspondentes às etiquetas LF.
  - g. Selecione Adicionar LF-tag.
  - h. Para Chave, selecione `Confidential`.
  - i. Em Valores, selecione `True`.
  - j. Para Permissões de banco de dados, selecione `Describe`.
  - k. Para Permissões de tabela, clique em Selecionar e Descrever.
  - l. Selecione Conceder.
2. Em seguida, repita as etapas para conceder permissões aos analistas de dados para a expressão LF-Tag `for. Confidential=False` Essa LF-tag é usada para descrever o `col_tag_database` e a tabela `source_data_col_lvl1` quando conectado como `lf-data-analyst` no Amazon Athena.
- a. Faça login no console do Lake Formation em <https://console.aws.amazon.com/lakeformation/> como `lf-data-engineer`.
  - b. Na página Bancos de dados, selecione o banco de dados `col_tag_database`.
  - c. Selecione Ações e Concessão.
  - d. Em Entidades principais, selecione Usuários e perfis do IAM.
  - e. Para Usuários e perfis do IAM, selecione `lf-data-analyst`.
  - f. Selecione Recursos correspondentes às etiquetas LF.
  - g. Selecione Adicionar LF-tag.
  - h. Para Chave, selecione `Confidential`.
  - i. Para Valores, selecione `False`.
  - j. Para Permissões de banco de dados, selecione `Describe`.
  - k. Para Permissões de tabela, não selecione nada.
  - l. Selecione Conceder.
3. Em seguida, repita as etapas para conceder permissões aos analistas de dados para a expressão da tag LF para `e. Confidential=False Sensitive=True` Essa LF-tag é usada para descrever o `col_tag_database` e a tabela `source_data_col_lvl1` quando conectado como `lf-data-analyst` no Amazon Athena.

- a. Entre no console do Lake Formation em <https://console.aws.amazon.com/lakeformation/> como lf-data-engineer.
- b. Na página Bancos de dados, selecione o banco de dados col\_tag\_database.
- c. Selecione Ações e Concessão.
- d. Em Entidades principais, selecione Usuários e perfis do IAM.
- e. Para Usuários e perfis do IAM, selecione lf-data-analyst.
- f. Selecione Recursos correspondentes às etiquetas LF.
- g. Selecione Adicionar LF-tag.
- h. Para Chave, selecione Confidential.
- i. Para Valores, selecione False.
- j. Selecione Adicionar LF-tag.
- k. Para Chave, selecione Sensitive.
- l. Para Valores, selecione True.
- m. Para Permissões de banco de dados, selecione Describe.
- n. Para Permissões de tabela, selecione Select e Describe.
- o. Selecione Conceder.

## Etapa 5: Executar uma consulta no Amazon Athena para verificar as permissões

Para essa etapa, use o Amazon Athena para executar consultas SELECT nas duas tabelas (source\_data and source\_data\_col\_lvl1). Use o caminho do Amazon S3 como o local do resultado da consulta (s3://lf-tagbased-demo-*Account-ID*/athena-results/).

1. Faça login no console do Athena em <https://console.aws.amazon.com/athena/> como lf-data-analyst.
2. No editor de consultas do Athena, selecione tag\_database no painel esquerdo.
3. Selecione o ícone de opções de menu adicionais (três pontos verticais) ao lado de source\_data e selecione Exibir tabela.
4. Selecione Executar consulta.

A consulta deve levar alguns minutos para ser executada. A consulta exibe todas as colunas na saída porque a LF-tag está associada no nível do banco de dados e a tabela `source_data` herdou automaticamente LF-tag do banco de dados `tag_database`.

5. Execute outra consulta usando `col_tag_database` e `source_data_col_lvl`.

A segunda consulta retorna as duas colunas que foram marcadas como `Non-Confidential` e `Sensitive`.

6. Você também pode verificar o comportamento da política de acesso baseada em tags do Lake Formation em colunas para as quais você não tem concessões de políticas. Quando uma coluna não marcada é selecionada na tabela `source_data_col_lvl`, o Athena retorna um erro. Por exemplo, você pode executar a seguinte consulta para escolher colunas não marcadas `geolocationid`:

```
SELECT geolocationid FROM "col_tag_database"."source_data_col_lvl" limit 10;
```

## Etapa 6: Limpar os recursos do AWS

Para evitar cobranças indesejadas em seu Conta da AWS, você pode excluir os recursos AWS que usou para este tutorial.

1. Faça login no console do Lake Formation como `lf-data-engineer` e exclua os bancos de dados `tag_database` e `col_tag_database`.
2. Em seguida, faça login como `lf-data-steward` e limpe todas as permissões de LF-tags, Permissões de dados e Permissões de localização de dados que foram concedidas acima e que foram concedidas a `lf-data-engineer` e `lf-data-analyst`.
3. Faça login no console do Amazon S3 como proprietário da conta usando as credenciais do IAM que você usou para implantar a pilha AWS CloudFormation.
4. Exclua os seguintes buckets:
  - `lf-tagbased-demo-accesslogs`- ID da *conta*
  - `lf-tagbased-demo`- ID da *conta*
5. Faça login no console AWS CloudFormation em <https://console.aws.amazon.com/cloudformation> e exclua a pilha que criou. Aguarde até que o status da pilha mude para `DELETE_COMPLETE`.

# Como proteger os data lakes com controle de acesso em nível de linha

As permissões em nível de linha do AWS Lake Formation permitem que você forneça acesso a linhas específicas em uma tabela com base em políticas de conformidade e governança de dados. Se você tem tabelas grandes armazenando bilhões de registros, você precisa de uma forma de permitir que diferentes usuários e equipes acessem somente os dados que eles têm permissão para ver. O controle de acesso em nível de linha é uma maneira simples e eficaz de proteger os dados e, ao mesmo tempo, dar aos usuários acesso aos dados de que precisam para realizar seu trabalho. O Lake Formation fornece relatórios centralizados de auditoria e conformidade, identificando quais entidades principais acessaram quais dados, quando, e por meio de quais serviços.

Neste tutorial, você aprenderá como os controles de acesso em nível de linha funcionam no Lake Formation e como configurá-los.

Este tutorial inclui um modelo do AWS CloudFormation para configurar rapidamente os recursos necessários. É possível revisá-lo e personalizá-lo para atender às suas necessidades.

## Tópicos

- [Público-alvo](#)
- [Pré-requisitos](#)
- [Etapa 1: Provisionar os recursos](#)
- [Etapa 2: Consulta sem filtros de dados](#)
- [Etapa 3: Configurar filtros de dados e conceder permissões](#)
- [Etapa 4: Consulta com filtros de dados](#)
- [Etapa 5: Limpar os recursos do AWS](#)

## Público-alvo

Este tutorial é destinado a administradores de dados, engenheiros de dados e analistas de dados. A tabela a seguir lista perfis e responsabilidades de um proprietário e um consumidor de dados.

Perfil	Descrição
Administrador do IAM	Um usuário que pode criar usuários e perfis e buckets do Amazon Simple Storage Service

Perfil	Descrição
	(Amazon S3). Usa a política <code>AdministratorAccess</code> gerenciada pelo AWS.
Administrador do data lake	Um usuário responsável por configurar o data lake, criar filtros de dados e conceder permissões aos analistas de dados.
Analista de dados	Um usuário que pode executar consultas no data lake. Analistas de dados residentes em países diferentes (para nosso caso de uso, EUA e Japão) só podem analisar avaliações de produtos de clientes localizados em seus próprios países e, por motivos de conformidade, não devem ser capazes de ver dados de clientes localizados em outros países.

## Pré-requisitos

Antes de iniciar este tutorial, você deve ter uma conta Conta da AWS que possa ser usada para fazer login como um usuário administrativo com as permissões corretas. Para obter mais informações, consulte [Concluir as tarefas iniciais de configuração do AWS](#).

O tutorial pressupõe que você esteja familiarizado com o IAM. Para obter informações sobre o IAM, consulte o [Guia do Usuário do IAM](#).

### Alterar as configurações do Lake Formation

#### Important

Antes de iniciar o modelo AWS CloudFormation, desative a opção Usar somente o controle de acesso do IAM para novos bancos de dados/tabelas no Lake Formation seguindo as etapas abaixo:

1. Fazer login no console do Lake Formation em <https://console.aws.amazon.com/lakeformation/> na região Leste dos EUA (Norte da Virgínia) ou Oeste dos EUA (Oregon).

2. Em Catálogo de Dados, selecione Configurações.
3. Desmarque Usar somente controle de acesso IAM para novos bancos de dados e Usar somente controle de acesso IAM para novas tabelas em novos bancos de dados.
4. Selecione Salvar.

## Etapa 1: Provisionar os recursos

Este tutorial inclui um modelo AWS CloudFormation para uma configuração rápida. É possível revisá-lo e personalizá-lo para atender às suas necessidades. O modelo AWS CloudFormation gera os seguintes recursos:

- Usuários e políticas para:
  - Administrador do Data Lake
  - Analista de dados dos EUA
  - Analista de dados do Japão
- Configurações e permissões do data lake do Lake Formation
- Uma função do Lambda (para recursos personalizados do AWS CloudFormation baseados em Lambda) usada para copiar arquivos de dados de amostra do bucket público do Amazon S3 para o seu bucket do Amazon S3
- Um bucket do Amazon S3 para servir como Data Lake.
- Um banco de dados do AWS Glue Data Catalog, uma tabela e uma partição

### Criar seus recursos

Siga estas etapas para criar seus recursos usando o AWS CloudFormation modelo.

1. Fazer login no console do AWS CloudFormation em <https://console.aws.amazon.com/cloudformation> na região Leste dos EUA (Norte da Virgínia).
2. Selecione [Iniciar Pilha](#).
3. Na página Criar pilha, selecione Avançar.
4. Digite um Nome de pilha.
5. Para Nome de usuário do administrador do data lake e Senha do usuário do administrador do data lake, digite o nome de usuário e a senha do IAM para o usuário administrador do data lake.



6. Para Nome de usuário do analista de dados dos EUA e Senha de usuário do analista de dados dos EUA, digite o nome de usuário e a senha para o nome de usuário e a senha desejados para o usuário analista de dados responsável pelo mercado dos EUA.
7. Para Nome de usuário do analista de dados do Japão e Senha de usuário do analista de dados do Japão, digite o nome de usuário e a senha para o nome de usuário e a senha desejados para o usuário analista de dados responsável pelo mercado do Japão.
8. Em Data Lake Nome do Bucket, digite o nome do seu bucket de dados.
9. Para Nome do banco de dados e Nome da tabela, deixe como padrão.
10. Selecione Avançar.
11. Na página seguinte, selecione Avançar.
12. Revise os detalhes na página final e selecione Eu reconheço que AWS CloudFormation pode criar recursos do IAM.
13. Selecione Criar.

A criação da pilha pode levar um minuto para ser concluída.

## Etapa 2: Consulta sem filtros de dados

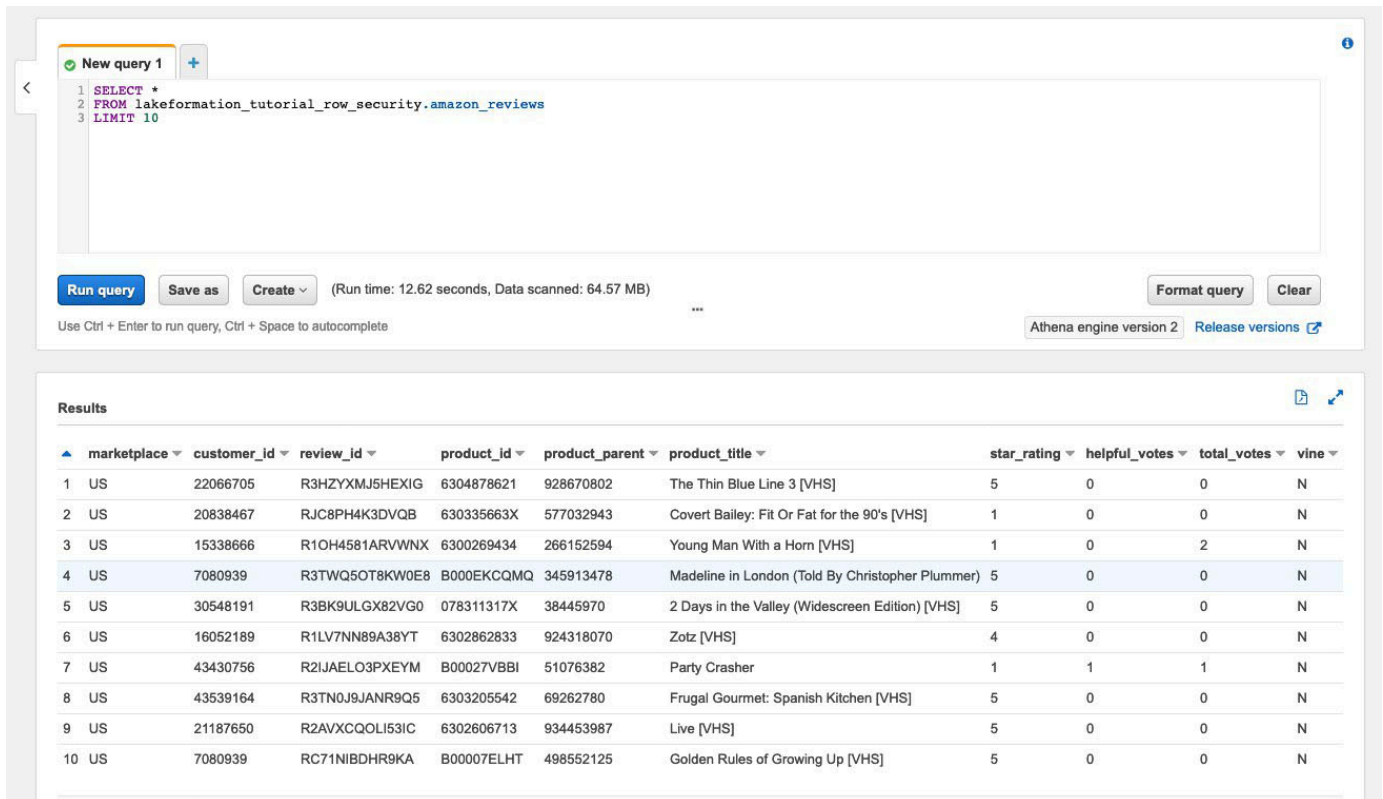
Depois que você configurar o ambiente, poderá consultar a tabela de avaliações de produtos. Primeiro, consulte a tabela sem controles de acesso em nível de linha para garantir que você possa ver os dados. Se você estiver executando consultas no Amazon Athena pela primeira vez, precisará configurar a localização do resultado da consulta.

Consulte a tabela sem controle de acesso em nível de linha

1. Faça login no console do Athena em <https://console.aws.amazon.com/athena/> como usuárioDataLakeAdmin e execute a seguinte consulta:

```
SELECT *
FROM lakeformation_tutorial_row_security.amazon_reviews
LIMIT 10
```

A captura de tela a seguir mostra o resultado da consulta. Essa tabela tem apenas uma partição, `product_category=Video`, portanto, cada registro é um comentário de avaliação de um produto de vídeo.



The screenshot displays the AWS Athena query editor and results page. The query executed is:

```
1 SELECT *
2 FROM lakeformation_tutorial_row_security.amazon_reviews
3 LIMIT 10
```

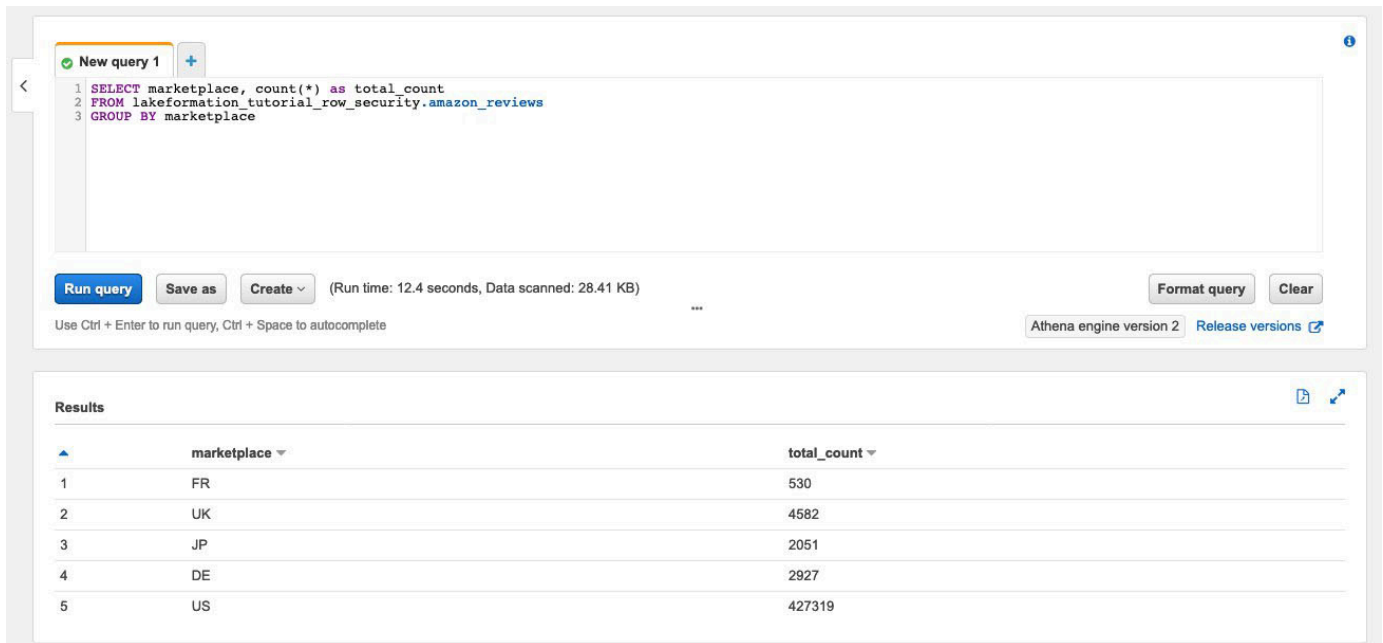
The results table contains 10 rows of data:

	marketplace	customer_id	review_id	product_id	product_parent	product_title	star_rating	helpful_votes	total_votes	vine
1	US	22066705	R3HZYXMJ5HEXIG	6304878621	928670802	The Thin Blue Line 3 [VHS]	5	0	0	N
2	US	20838467	RJC8PH4K3DVQB	630335663X	577032943	Covert Bailey: Fit Or Fat for the 90's [VHS]	1	0	0	N
3	US	15338666	R1OH4581ARVWNX	6300269434	266152594	Young Man With a Horn [VHS]	1	0	2	N
4	US	7080939	R3TWQ5OT8KW0E8	B000EKCQMQ	345913478	Madeline in London (Told By Christopher Plummer)	5	0	0	N
5	US	30548191	R3BK9ULGX82VG0	078311317X	38445970	2 Days in the Valley (Widescreen Edition) [VHS]	5	0	0	N
6	US	16052189	R1LV7NN89A38YT	6302862833	924318070	Zotz [VHS]	4	0	0	N
7	US	43430756	R2JAELO3PXEYM	B00027VBBI	51076382	Party Crasher	1	1	1	N
8	US	43539164	R3TNQ9JANR9Q5	6303205542	69262780	Frugal Gourmet: Spanish Kitchen [VHS]	5	0	0	N
9	US	21187650	R2AVXCQQLI53IC	6302606713	934453987	Live [VHS]	5	0	0	N
10	US	7080939	RC71NIBDHR9KA	B00007ELHT	498552125	Golden Rules of Growing Up [VHS]	5	0	0	N

2. Em seguida, execute uma consulta de agregação para recuperar o número total de registros por marketplace.

```
SELECT marketplace, count(*) as total_count
FROM lakeformation_tutorial_row_security.amazon_reviews
GROUP BY marketplace
```

A captura de tela a seguir mostra o resultado da consulta. A coluna marketplace tem cinco valores diferentes. Nas etapas subsequentes, você configurará filtros baseados em linhas usando a coluna marketplace.



The screenshot shows the AWS Athena console interface. At the top, there is a text area for a SQL query with the following content:

```
1 SELECT marketplace, count(*) as total_count
2 FROM lakeformation_tutorial_row_security.amazon_reviews
3 GROUP BY marketplace
```

Below the query editor, there are buttons for "Run query", "Save as", and "Create". To the right, it displays "(Run time: 12.4 seconds, Data scanned: 28.41 KB)". Further right are buttons for "Format query" and "Clear". At the bottom right, it says "Athena engine version 2" and "Release versions".

Below the query editor, the "Results" section shows a table with the following data:

	marketplace	total_count
1	FR	530
2	UK	4582
3	JP	2051
4	DE	2927
5	US	427319

### Etapa 3: Configurar filtros de dados e conceder permissões

Este tutorial usa dois analistas de dados: um responsável pelo mercado dos EUA e outro pelo mercado japonês. Cada analista usa o Athena para analisar avaliações de clientes somente para seu mercado específico. Crie dois filtros de dados diferentes, um para o analista responsável pelo mercado dos EUA e outro para o responsável pelo mercado japonês. Em seguida, conceda aos analistas suas respectivas permissões.

Crie filtros de dados e conceda permissões

1. Crie um filtro para restringir o acesso do US aos dados do marketplace.
  - a. Faça login no console do Lake Formation em <https://console.aws.amazon.com/lakeformation/> na região Leste dos EUA (Norte da Virgínia) como usuário DataLakeAdmin.
  - b. Selecione Filtros de dados.
  - c. Selecione Criar novo filtro.
  - d. Em Nome do filtro de dados, digite `amazon_reviews_US`.
  - e. Em Banco de dados de destino, selecione o banco de dados `lakeformation_tutorial_row_security`.
  - f. Em Tabela de destino, selecione a tabela `amazon_reviews`.
  - g. Para acesso em nível de coluna, deixe como padrão.

- h. Em Expressão de filtro de linha, digite `marketplace= 'US'`.
    - i. Selecione Criar filtro.
  2. Crie um filtro para restringir o acesso aos dados japoneses do marketplace.
    - a. Na página Filtros de dados, selecione Criar novo filtro.
    - b. Em Nome do filtro de dados, digite `amazon_reviews_JP`.
    - c. Em Banco de dados de destino, selecione o banco de dados `lakeformation_tutorial_row_security`.
    - d. Em Tabela de destino, selecione o `table amazon_reviews`.
    - e. Para acesso em nível de coluna, deixe como padrão.
    - f. Em Expressão de filtro de linha, digite `marketplace= 'JP'`.
    - g. Selecione Criar filtro.
  3. Em seguida, conceda permissões aos analistas de dados usando esses filtros de dados. Siga estas etapas para conceder permissões ao analista de dados dos EUA (`DataAnalystUS`):
    - a. Em Permissões, escolha Permissões do Data lake.
    - b. Em Permissão de dados, selecione Conceder.
    - c. Para Entidades principais, selecione Usuários e perfis do IAM e selecione o perfil `DataAnalystUS`.
    - d. Para LF-tags ou recursos de catálogo, selecione Recursos nomeados de Catálogo de Dados.
    - e. Para o Banco de dados, selecione `lakeformation_tutorial_row_security`.
    - f. Para Tabelas opcionais, selecione `amazon_reviews`.
    - g. Para Filtros de dados — opcional, selecione `amazon_reviews_US`.
    - h. Para Permissões de filtro de dados, marque Selecionar.
    - i. Selecione Conceder.
  4. Siga estas etapas para conceder permissões ao analista de dados japonês (`DataAnalystJP`):
    - a. Em Permissões, escolha Permissões do Data lake.
    - b. Em Permissão de dados, selecione Conceder.
    - c. Para Entidades principais, selecione Usuários e perfis do IAM e selecione o perfil `DataAnalystJP`.

- d. Para LF-tags ou recursos de catálogo, selecione Recursos nomeados de Catálogo de Dados.
- e. Para o Banco de dados, selecione `lakeformation_tutorial_row_security`.
- f. Para Tabelas opcionais, selecione `amazon_reviews`.
- g. Para Filtros de dados — opcional, selecione `amazon_reviews_JP`.
- h. Para Permissões de filtro de dados, marque Selecionar.
- i. Selecione Conceder.

## Etapa 4: Consulta com filtros de dados

Com os filtros de dados anexados à tabela de avaliações de produtos, faça algumas consultas e veja como as permissões são aplicadas pelo Lake Formation.

1. Faça login no console do Athena em <https://console.aws.amazon.com/athena/> como o usuário `DataAnalystUS`.
2. Execute a consulta a seguir para recuperar alguns registros, que são filtrados com base nas permissões em nível de linha que definimos:

```
SELECT *  
FROM lakeformation_tutorial_row_security.amazon_reviews  
LIMIT 10
```

A captura de tela a seguir mostra o resultado da consulta.

The screenshot shows the AWS Athena console interface. At the top, there are tabs for 'New query 1' and 'New query 2'. The SQL editor contains the following query:

```
1 SELECT *
2 FROM lakeformation_tutorial_row_security.amazon_reviews
3 LIMIT 10
```

Below the editor, there are buttons for 'Run query', 'Save as', and 'Create'. A status bar indicates '(Run time: 11.9 seconds, Data scanned: 0 KB)'. There are also buttons for 'Format query' and 'Clear'. At the bottom right, it says 'Athena engine version 2' and 'Release versions'.

The 'Results' section displays a table with 10 rows and 12 columns. The columns are: marketplace, customer\_id, review\_id, product\_id, product\_parent, product\_title, star\_rating, helpful\_votes, total\_votes, vine, verified\_purchase, and review\_text. The data is as follows:

marketplace	customer_id	review_id	product_id	product_parent	product_title	star_rating	helpful_votes	total_votes	vine	verified_purchase	review_text
US	43836277	R2NUBTTUO60VYU	B00068S41I	653409458	The Notebook [VHS]	4	0	0	N	Y	KL
US	20261976	R2QTOLZUQERU5B	6303060013	176265879	American Cyborg: Steel Warrior [VHS]	5	0	1	N	Y	it
US	15947067	R1PHKR75RKZNSU	6303927319	850909689	Biography - Darryl Zanuck [VHS]	5	0	0	N	N	G
US	19288153	R1BL2WVE5X34UN	6304032153	479446069	Timon & Pumbaa: Quit Buggin Me [VHS]	5	0	0	N	N	FI
US	19712967	R2DKOCIBS5FSP7	0784017743	35164822	Denise Austin - Hit the Spot: Arms & Bust [VHS]	5	0	0	N	Y	G
US	51047097	R2XF5HQATT4IVR	0793960142	233936597	I Love Lucy - Lucy's Italian Movie/Ballet [VHS]	5	0	0	N	N	FI
US	43836277	R2NUBTTUO60VYU	B00068S41I	653409458	The Notebook [VHS]	4	0	0	N	Y	KL
US	51047097	R1C0H0G6NATZOX	6304872585	233936597	I Love Lucy: Lucy Meets Superman/Freez [VHS]	5	0	1	N	N	FI
US	42808630	R2HXW7UD4IGZLN	6303060013	176265879	American Cyborg: Steel Warrior [VHS]	5	0	1	N	Y	M
US	11682952	R18IURLUPYI4DP	6302993717	42308924	Songs of Christmas [VHS]	1	0	0	N	Y	R

- Da mesma forma, execute uma consulta para contar o número total de registros por loja.

```
SELECT marketplace , count ( * ) as total_count
FROM lakeformation_tutorial_row_security .amazon_reviews
GROUP BY marketplace
```

O resultado da consulta mostra apenas o marketplace US nos resultados. Isso ocorre porque o usuário só pode ver as linhas em que o valor da coluna marketplace é igual a US.

- Mude para o usuário DataAnalystJP e execute a mesma consulta.

```
SELECT *
FROM lakeformation_tutorial_row_security.amazon_reviews
LIMIT 10
```

O resultado da consulta mostra apenas os registros que pertencem ao domínio JP marketplace.

- Execute a consulta para contar o número total de registros por marketplace.

```
SELECT marketplace, count(*) as total_count
FROM lakeformation_tutorial_row_security.amazon_reviews
GROUP BY marketplace
```

O resultado da consulta mostra apenas a linha pertencente ao domínio JP marketplace.

## Etapa 5: Limpar os recursos do AWS

### Limpar os recursos

Para evitar cobranças indesejadas em sua Conta da AWS, você pode excluir os recursos AWS que usou para este tutorial.

- [Exclua a pilha do Cloud Formation.](#)

## Compartilhamento de um Data Lake usando controle de acesso baseado em tags do Lake Formation e recursos nomeados

Este tutorial demonstra como você pode configurar AWS Lake Formation para compartilhar com segurança os dados armazenados em um Data Lake com várias empresas, organizações ou unidades de negócios, sem precisar copiar o banco de dados inteiro. Há duas opções para compartilhar seus bancos de dados e tabelas com outra Conta da AWS usando o controle de acesso entre contas do Lake Formation:

- Controle de acesso baseado em tags do Lake Formation (LF-TBAC) (recomendado)

O controle de acesso baseado em tags do Lake Formation é uma estratégia de autorização que define permissões com base em atributos. No Lake Formation, esses atributos são chamados de LF-tags. Para obter mais detalhes, consulte [Como gerenciar um data lake usando o controle de acesso baseado em tags do Lake Formation.](#)

- Recursos nomeados do Lake Formation

O método de recurso chamado Lake Formation é uma estratégia de autorização que define permissões para recursos. Os recursos incluem bancos de dados, tabelas e colunas. Os administradores do Data Lake podem atribuir e revogar permissões em recursos do Lake Formation. Para obter mais detalhes, consulte [Compartilhamento de dados entre contas no Lake Formation.](#)

Recomendamos usar recursos nomeados se o administrador do Data Lake preferir conceder permissões explicitamente a recursos individuais. Quando você concede permissões do Lake Formation em um recurso do Catálogo de Dados para uma conta externa ou diretamente para uma

entidade principal do IAM em outra conta, o Lake Formation usa o serviço AWS Resource Access Manager (AWS RAM) para compartilhar o recurso.

## Tópicos

- [Público-alvo](#)
- [Definir as configurações do Catálogo de Dados do Lake Formation na conta do produtor](#)
- [Etapa 1: Provisionar seus recursos usando modelos AWS CloudFormation](#)
- [Etapa 2: Pré-requisitos de compartilhamento entre contas do Lake Formation](#)
- [Etapa 3: Implementar o compartilhamento entre contas usando o método de controle de acesso baseado em tags](#)
- [Etapa 4: Implementar o método de recurso nomeado](#)
- [Etapa 5: Limpar os recursos do AWS](#)

## Público-alvo

Este guia é destinado a administradores de dados, engenheiros de dados e analistas de dados. Quando se trata de compartilhar tabelas do Catálogo de Dados do AWS Glue e administrar permissões no Lake Formation, os administradores de dados das contas produtoras têm propriedade funcional com base nas funções às quais dão suporte e podem conceder acesso a vários consumidores, organizações externas e contas. A tabela a seguir lista os perfis usados neste tutorial:

Perfil	Descrição
DataLakeAdminProducer	<p>O usuário do IAM administrador do Data Lake tem o seguinte acesso:</p> <ul style="list-style-type: none"><li>• Acesso completo de leitura, gravação e atualização a todos os recursos do catálogo de dados</li><li>• Capacidade de conceder permissões aos recursos</li><li>• Criar um link de recurso para a tabela compartilhada</li></ul>




Perfil	Descrição
	<ul style="list-style-type: none"> <li>• Pode anexar LF-tags aos recursos, fornecendo acesso às entidades principais com base em quaisquer políticas criadas por administradores de dados</li> </ul>
Consumidor administrativo do DataLake	<p>O usuário do IAM administrador do Data Lake tem o seguinte acesso:</p> <ul style="list-style-type: none"> <li>• Acesso completo de leitura, gravação e atualização a todos os recursos do catálogo de dados</li> <li>• Capacidade de conceder permissões aos recursos</li> <li>• Criar um link de recurso para a tabela compartilhada</li> <li>• Pode anexar LF-tags aos recursos, fornecendo acesso às entidades principais com base em quaisquer políticas criadas por administradores de dados</li> </ul>
DataAnalyst	<p>O usuário do IAM administrador do Data Lake tem o seguinte acesso:</p> <ul style="list-style-type: none"> <li>• Acesso refinado a recursos compartilhados pelas políticas de acesso baseadas em tags do Lake Formation ou usando o método de recursos nomeados</li> </ul>

## Definir as configurações do Catálogo de Dados do Lake Formation na conta do produtor

Antes de iniciar este tutorial, você deve ter uma Conta da AWS que possa ser usada para fazer login como um usuário administrativo com as permissões corretas. Para obter mais informações, consulte [Concluir as tarefas iniciais de configuração do AWS](#).


O tutorial pressupõe que você esteja familiarizado com o IAM. Para obter mais informações sobre o IAM, consulte o [Guia do Usuário do IAM](#).

Definir as configurações do Catálogo de Dados do Lake Formation na conta do produtor

 Note

Neste tutorial, a conta que tem a tabela de origem é chamada de conta do produtor, e a conta que precisa acessar a tabela de origem é chamada de conta do consumidor.

O Lake Formation fornece seu próprio modelo de gerenciamento de permissões. Para manter a compatibilidade com versões anteriores do modelo de permissão do IAM, a permissão Super é concedida ao grupo IAMAllowedPrincipals em todos os recursos AWS Glue Data Catalog existentes por padrão. Além disso, a opção Usar somente as configurações de controle de acesso do IAM estão habilitadas para novos recursos do Catálogo de Dados. Este tutorial usa controle de acesso refinado usando permissões do Lake Formation e usa políticas do IAM para controle de acesso de único fator. Para mais detalhes, consulte [Métodos para controle de acesso de alta granularidade](#). Portanto, antes de usar um modelo AWS CloudFormation para uma configuração rápida, você precisa alterar as configurações do Lake Formation Data Catalog na conta do produtor.

 Important

Essa configuração afeta todos os bancos de dados e tabelas recém-criados, portanto, é altamente recomendável concluir este tutorial em uma conta que não seja de produção ou em uma nova conta. Além disso, se você estiver usando uma conta compartilhada (como a conta de desenvolvimento da sua empresa), certifique-se de que ela não afete outros recursos. Se você preferir manter as configurações de segurança padrão, conclua uma etapa adicional ao compartilhar recursos com outras contas, na qual você revoga a Super permissão padrão do banco de dados ou da tabela do IAMAllowedPrincipals. Entraremos em detalhes mais adiante neste tutorial.

Para definir as configurações do Catálogo de Dados do Lake Formation na conta do produtor, conclua as etapas a seguir:

1. Faça login no AWS Management Console usando a conta do produtor como usuário administrador ou como usuário com permissão da API PutDataLakeSettings do Lake Formation.
2. No console do Lake Formation, no painel de navegação, em Catálogo de dados, selecione Configurações.
3. Desmarque Usar somente controle de acesso IAM para novos bancos de dados e Usar somente controle de acesso IAM para novas tabelas em novos bancos de dados

Selecione Salvar.

**AWS Lake Formation** > Data catalog settings

## Data catalog settings

### Default permissions for newly created databases and tables

These settings maintain existing AWS Glue Data Catalog behavior. You can still set individual permissions on databases and tables, which will take effect when you revoke the Super permission from IAMAllowedPrincipals. See [Changing Default Settings for Your Data Lake](#).

Use only IAM access control for new databases

Use only IAM access control for new tables in new databases

### Default permissions for AWS CloudTrail

These settings specify the information being shown in AWS CloudTrail.

**Resource owners**  
Enter resource owners you wish to share your CloudTrail access details with.

Enter one or more AWS account IDs. Press Enter after each ID.

Cancel **Save**

Além disso, você pode remover as permissões de CREATE\_DATABASE para IAMAllowedPrincipals em Perfis e tarefas administrativas, Criadores de banco de dados. Só então, você pode determinar quem pode criar um novo banco de dados por meio das permissões do Lake Formation.

## Etapa 1: Provisionar seus recursos usando modelos AWS CloudFormation

O modelo do CloudFormation para a conta do produtor gera os seguintes recursos:

- Um bucket do Amazon S3 para servir como Data Lake.
- Uma função do Lambda (para recursos personalizados baseados em Lambda AWS CloudFormation). Usamos a função para copiar exemplos de arquivos de dados do bucket público do Amazon S3 para o bucket do Amazon S3.
- Usuários e políticas do IAM: DataLakeAdminProducer.
- As configurações e permissões apropriadas do Lake Formation, incluindo:
  - Definir o administrador do Data Lake Lake Formation na conta do produtor
  - Registrar um bucket do Amazon S3 como local do Data Lake do Lake Formation (conta do produtor)
- Um banco de dados AWS Glue Data Catalog, uma tabela e uma partição. Como há duas opções para compartilhar recursos Contas da AWS, esse modelo cria dois conjuntos separados de banco de dados e tabela.

O modelo AWS CloudFormation da conta do consumidor gera os seguintes recursos:

- Usuários e políticas do IAM:
  - Consumidor administrativo do DataLake
  - DataAnalyst
- Um banco de dados do AWS Glue Data Catalog. Esse banco de dados serve para criar links de recursos para recursos compartilhados.

Crie seus recursos na conta do produtor

1. Entre no console do AWS CloudFormation em <https://console.aws.amazon.com/cloudformation> na região Leste dos EUA (Norte da Virgínia).
2. Selecione [Iniciar Pilha](#).
3. Selecione Avançar.
4. Em Nome da pilha, digite um nome de pilha, como `stack-producer`.
5. Na seção Configuração do usuário, insira o nome de usuário e a senha para `ProducerDataLakeAdminUserName` e `ProducerDataLakeAdminUserPassword`.


6. Em `DataLakeBucketName`, insira o nome do seu bucket do Data Lake. Esse nome precisa ser globalmente exclusivo.
7. Para `DatabaseName` e `TableName`, deixe os valores-padrão.
8. Selecione Avançar.
9. Na próxima página, selecione Avançar.
10. Revise os detalhes na página final e selecione Eu reconheço que AWS CloudFormation pode criar recursos do IAM.
11. Selecione Criar.

A criação da pilha pode levar até um minuto.

Crie seus recursos na conta do consumidor

1. Entre no console do AWS CloudFormation em <https://console.aws.amazon.com/cloudformation> na região Leste dos EUA (Norte da Virgínia).
2. Selecione [Iniciar Pilha](#).
3. Selecione Avançar.
4. Em Nome da pilha, digite um nome de pilha, como `stack-consumer`.
5. Na seção Configuração do usuário, insira o nome de usuário e a senha para `ConsumerDataLakeAdminUserName` e `ConsumerDataLakeAdminUserPassword`.
6. Para `DataAnalystUserName` e `DataAnalystUserPassword`, insira o nome de usuário e a senha que você deseja para o usuário do IAM do analista de dados.
7. Em `DataLakeBucketName`, insira o nome do seu bucket do Data Lake. Esse nome precisa ser globalmente exclusivo.
8. Para `DatabaseName`, deixe os valores-padrão.
9. Em `AthenaQueryResultS3BucketName`, insira o nome do bucket do Amazon S3 que armazena resultados da consulta do Amazon Athena. Se você não tiver um, [crie um bucket da Amazon S3](#).
10. Selecione Avançar.
11. Na próxima página, selecione Avançar.
12. Revise os detalhes na página final e selecione Eu reconheço que AWS CloudFormation pode criar recursos do IAM.
13. Selecione Criar.

A criação da pilha pode levar até um minuto.

 Note

Depois de concluir o tutorial, exclua a pilha AWS CloudFormation para evitar cobranças. Verifique se os recursos foram excluídos com sucesso no status de eventos da pilha.

## Etapa 2: Pré-requisitos de compartilhamento entre contas do Lake Formation

Antes de compartilhar recursos com o Lake Formation, há pré-requisitos tanto para o método de controle de acesso baseado em tags quanto para o método de recursos nomeados.

Controle de acesso completo com base em tags: pré-requisitos do para compartilhamento de dados do entre contas

- Para obter mais informações sobre os requisitos de compartilhamento de dados entre contas, consulte a seção [Pré-requisitos](#) no capítulo Compartilhamento de dados entre contas.

Para compartilhar os recursos do Catálogo de Dados com a versão 3 ou superior das Configurações de Versão Entre Contas, o concedente precisa ter as permissões de IAM definidas na política `AWSLakeFormationCrossAccountManager` gerenciada por AWS em sua conta.

Se você estiver usando a versão 1 ou a versão 2 das Configurações de versão entre contas, antes de usar o método de controle de acesso baseado em tags para conceder acesso entre contas aos recursos, você deve adicionar o seguinte objeto de permissões JSON à política de recursos do Catálogo de Dados na conta do produtor. Isso dá à conta do consumidor permissão para acessar o catálogo de dados quando `glue:EvaluatedByLakeFormationTags` é verdadeiro. Além disso, essa condição se torna verdadeira para recursos nos quais você concedeu permissão usando tags de permissão do Lake Formation para a conta do consumidor. Essa política é necessária para todas Conta da AWS as quais você está concedendo permissões.

A política a seguir deve estar dentro de um elemento `Statement`. Discutiremos a Política do IAM completa na próxima seção.

```
{
  "Effect": "Allow",
  "Action": [
    "glue:*"
  ],
  "Principal": {
    "AWS": [
      "consumer-account-id"
    ]
  },
  "Resource": [
    "arn:aws:glue:region:account-id:table/*",
    "arn:aws:glue:region:account-id:database/*",
    "arn:aws:glue:region:account-id:catalog"
  ],
  "Condition": {
    "Bool": {
      "glue:EvaluatedByLakeFormationTags": true
    }
  }
}
```

Preencha os pré-requisitos de compartilhamento entre contas do método de recurso nomeado

1. Se não houver uma política de recursos do Catálogo de Dados em sua conta, as concessões entre contas do Lake Formation que você fizer procederão normalmente. No entanto, se existir uma política de recursos do Catálogo de Dados, você deverá adicionar a seguinte declaração a ela para permitir que suas concessões entre contas sejam bem-sucedidas se forem feitas com o método de recurso nomeado. Se você planeja usar somente o método de recurso nomeado ou somente o método de controle de acesso baseado em tag, você pode pular esta etapa. Neste tutorial, avaliamos os dois métodos e precisamos adicionar a política a seguir.

A política a seguir deve estar dentro de um elementoStatement. Discutiremos a Política do IAM completa na próxima seção.

```
{
  "Effect": "Allow",
  "Action": [
    "glue:ShareResource"
```

```

    ],
    "Principal": {
      "Service": "ram.amazonaws.com"
    },
    "Resource": [
      "arn:aws:glue:region:account-id:table/*/*",
      "arn:aws:glue:region:account-id:database/*",
      "arn:aws:glue:region:account-id:catalog"
    ]
  }
}

```

2. Em seguida, adicione a política de recursos AWS Glue Data Catalog usando o AWS Command Line Interface (AWS CLI).

Se você conceder permissões entre contas usando o método de controle de acesso baseado em tags e o método de recurso nomeado, defina o argumento `EnableHybrid` como `"true"` ao adicionar as políticas anteriores. Como essa opção não é compatível com o console no momento, é necessário usar a API `glue:PutResourcePolicy` e a AWS CLI.

Primeiro, crie um documento de política (como `policy.json`) e adicione as duas políticas anteriores. Substitua `consumer-account-id` pelo *ID da conta* do Conta da AWS que está recebendo a concessão, `region` pela região do catálogo de dados que contém os bancos de dados e tabelas aos quais você está concedendo permissões e `account-id` pelo ID da Conta da AWS do produtor.

```

{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Principal": {
        "Service": "ram.amazonaws.com"
      },
      "Action": "glue:ShareResource",
      "Resource": [
        "arn:aws:glue:region:account-id:table/*/*",
        "arn:aws:glue:region:account-id:database/*",
        "arn:aws:glue:region:account-id:catalog"
      ]
    },
    {

```



```
    "Effect": "Allow",
    "Principal": {
      "AWS": "region:account-id"
    },
    "Action": "glue:*",
    "Resource": [
      "arn:aws:glue:region:account-id:table/*/*",
      "arn:aws:glue:region:account-id:database/*",
      "arn:aws:glue:region:account-id:catalog"
    ],
    "Condition": {
      "Bool": {
        "glue:EvaluatedByLakeFormationTags": "true"
      }
    }
  }
]
```

Insira o comando AWS CLI a seguir. Substitua *glue-resource-policy* pelos valores corretos (como `file://policy.json`).

```
aws glue put-resource-policy --policy-in-json glue-resource-policy --enable-hybrid TRUE
```

Para obter mais informações, consulte [put-resource-policy](#).


## Etapa 3: Implementar o compartilhamento entre contas usando o método de controle de acesso baseado em tags

Nesta seção, orientamos sobre estas etapas de alto nível:

1. Defina uma LF-tag.
2. Atribua a LF-tag ao recurso de destino.
3. Conceda permissão da LF-tag à conta do consumidor.
4. Conceda permissão de dados à conta do consumidor.
5. Opcionalmente, revogue as permissões para `IAMAllowedPrincipals` no banco de dados, nas tabelas e nas colunas.

6. Crie um link de recurso para a tabela compartilhada.
7. Crie uma LF-tag e atribua-a ao banco de dados de destino.
8. Conceda permissão de dados LF-tag à conta do consumidor.

Defina uma LF-tag

 Note

Se você estiver conectado à sua conta de produtor, saia antes de concluir as etapas a seguir.

1. Faça login na conta do produtor como administrador do Data Lake em <https://console.aws.amazon.com/lakeformation/>. Use o número da conta do produtor, o nome de usuário do IAM (o padrão é `DataLakeAdminProducer`) e a senha que você especificou durante a criação da pilha AWS CloudFormation.
2. No console do Lake Formation (<https://console.aws.amazon.com/lakeformation/>), no painel de navegação, em Permissões e em Perfis e tarefas administrativas, selecione LF-tags.
3. Selecione Adicionar LF-tag.

Atribua a LF-tag ao recurso de destino

Atribua a LF-tag ao recurso de destino e conceda permissões de dados para outra conta

Como administrador do Data Lake, você pode anexar tags aos recursos. Se você planeja usar uma função separada, talvez seja necessário conceder permissões, descrever e anexar ao perfil separado.

1. No painel de navegação, em Catálogo de Dados, selecione Bancos de Dados.
2. Selecione o banco de dados de destino (`lakeformation_tutorial_cross_account_database_tbac`) e, no menu Ações, selecione Editar LF-tags.

Neste tutorial, você atribui uma LF-tag a um banco de dados, mas também pode atribuir LF-tags a tabelas e colunas.

3. Selecione Atribuir nova LF-tag.
4. Adicione a chave `Confidentiality` e o valor `public`.

## 5. Selecione Salvar.

Conceda permissão da LF-tag à conta do consumidor.

Ainda na conta do produtor, conceda permissões à conta do consumidor para acessar a LF-tag.

1. No painel de navegação, em Permissões, Perfis e tarefas administrativas, Permissões LF-tags, selecione Conceder.
2. Para Entidades principais, selecione Contas externas.
3. Insira o ID da Conta da AWS de destino.

Contas da AWS dentro da mesma organização aparecem automaticamente. Caso contrário, você precisará inserir manualmente o ID da Conta da AWS. No momento em que este artigo foi escrito, o controle de acesso baseado em tags do Lake Formation não dava suporte à concessão de permissão a organizações ou unidades organizacionais.

4. Para LF-tags, selecione a chave e os valores da LF-tag que está sendo compartilhada com a conta do consumidor (chave Confidentiality e valor public).
5. Em Permissões, selecione Descrever para Permissões da LF-tag.

As permissões LF-tag são permissões concedidas à conta do consumidor. Permissões concedidas são permissões que a conta do consumidor pode conceder a outras entidades principais.

6. Selecione Conceder.

Nesse ponto, o administrador do Data Lake do consumidor deve ser capaz de encontrar a tag de política que está sendo compartilhada por meio do console Lake Formation da conta do consumidor, em Permissões, Perfis e tarefas administrativas, LF-tags.

Conceda a permissão de dados à conta do consumidor.

Agora, forneceremos acesso aos dados da conta do consumidor especificando uma expressão de LF-tag, concedendo à conta do consumidor acesso a qualquer tabela ou banco de dados que corresponda à expressão.

1. No painel de navegação, em Permissões, Permissões do Data Lake, selecione Conceder.
2. Em Entidades Principais, selecione Contas externas e insira a ID da Conta da AWS de destino.

3. Para LF-tags ou recursos de catálogo, escolha a chave e os valores da LF-tag que está sendo compartilhada com a conta do consumidor (chave Confidentiality e valor public).
4. Para Permissões, em Recursos combinados com LF-tags (recomendado), selecione Adicionar LF-tag.
5. Selecione a chave e o valor da tag que está sendo compartilhada com a conta do consumidor (chave Confidentiality e valor public).
6. Para permissões de banco de dados, selecione Descrever em Permissões de banco de dados para conceder permissões de acesso no nível do banco de dados.
7. O administrador do Data Lake do consumidor deve ser capaz de encontrar a tag de política que está sendo compartilhada por meio da conta do consumidor no console do Lake Formation em <https://console.aws.amazon.com/lakeformation/>, em Permissões, Perfis e tarefas administrativas, LF-tags.
8. Selecione Descrever em Permissões concedíveis para que a conta do consumidor possa conceder permissões em nível de banco de dados a seus usuários.
9. Para Permissões de tabela e coluna, selecione Selecionar e Descrever em Permissões de tabela.
10. Selecione Selecionar e Descrever em Permissões concedíveis.
11. Selecione Conceder.

Revogue a permissão para **IAMAllowedPrincipals** no banco de dados, nas tabelas e nas colunas (opcional).

No início deste tutorial, você alterou as configurações do Lake Formation Data Catalog. Se você pulou essa parte, esta etapa é necessária. Se tiver alterado as configurações do Catálogo de Dados do Lake Formation, ignore esta etapa.

Nesta etapa, precisamos revogar a Super permissão padrão do banco de dados IAMAllowedPrincipals ou da tabela. Para mais detalhes, consulte [Etapa 4: mude seus armazenamentos de dados para o modelo de permissões do Lake Formation](#).

Antes de revogar a permissão de IAMAllowedPrincipals, certifique-se de ter concedido às entidades principais existentes do IAM a permissão necessária por meio do Lake Formation. Isso inclui três etapas:

1. Adicione a permissão do IAM ao usuário ou perfil do IAM de destino com a ação GetDataAccess do Lake Formation (com a política do IAM).

2. Conceda ao usuário ou perfil de destino do IAM com permissões de dados do Lake Formation (alterar, selecionar e assim por diante).
3. Em seguida, revogue as permissões para `IAMAllowedPrincipals`. Caso contrário, depois de revogar as permissões para `IAMAllowedPrincipals`, as entidades principais existentes do IAM talvez não consigam mais acessar o banco de dados ou o catálogo de dados de destino.

A revogação da Super permissão para `IAMAllowedPrincipals` é necessária quando você deseja aplicar o modelo de permissão do Lake Formation (em vez do modelo de política do IAM) para gerenciar o acesso do usuário em uma única conta ou entre várias contas usando o modelo de permissão do Lake Formation. Você não precisa revogar a permissão de `IAMAllowedPrincipals` outras tabelas nas quais deseja manter o modelo de política tradicional do IAM.

Nesse ponto, o administrador do Data Lake da conta do consumidor deve ser capaz de encontrar o banco de dados e a tabela que estão sendo compartilhados por meio da conta do consumidor no console do Lake Formation em <https://console.aws.amazon.com/lakeformation/>, em Catálogo de dados, bancos de dados. Caso contrário, confirme se o seguinte está configurado corretamente:

1. A tag e os valores de política corretos são atribuídos aos bancos de dados e tabelas de destino.
2. A permissão correta de tag e a permissão de dados são atribuídas à conta do consumidor.
3. Revogue a super permissão padrão do banco de dados `IAMAllowedPrincipals` ou da tabela.

### Criar um link de recurso para a tabela compartilhada

Quando um recurso é compartilhado entre contas e os recursos compartilhados não são colocados no Catálogo de Dados das contas do consumidor. Para disponibilizá-los e consultar os dados subjacentes de uma tabela compartilhada usando serviços como o Athena, precisamos criar um link de recurso para a tabela compartilhada. Um link de recurso é um objeto do Catálogo de Dados que é um link para um banco de dados, ou uma tabela local ou compartilhada. Para obter mais detalhes, consulte [Criação de links de recursos](#). Ao criar um link de recurso, você pode:

- Atribuir um nome diferente a um banco de dados ou tabela que esteja alinhado às políticas de nomenclatura de recursos do Catálogo de Dados.

- Usar serviços como Athena e Redshift Spectrum para consultar bancos de dados ou tabelas compartilhados.

Para criar um link de recurso, conclua as etapas a seguir:

1. Se você estiver conectado à sua conta de consumidor, saia.
2. Faça login como administrador do Data Lake da conta do consumidor. Use o ID da conta do consumidor, o nome de usuário do IAM (DataLakeAdminConsumer padrão) e a senha que você especificou durante a criação da pilha AWS CloudFormation.
3. No console do Lake Formation (<https://console.aws.amazon.com/lakeformation/>), no painel de navegação, em Catálogo de Dados, Bancos de Dados, selecione o banco de dados compartilhado `lakeformation_tutorial_cross_account_database_tbac`.

Se você não vir o banco de dados, revise as etapas anteriores para ver se tudo está configurado corretamente.

4. Selecione Visualizar Tabelas.
5. Selecione a tabela compartilhada `amazon_reviews_table_tbac`.
6. No menu Ações, selecione Criar link de recurso.
7. Em Nome do link de recurso, insira um nome (para este tutorial, `amazon_reviews_table_tbac_resource_link`).
8. Em Banco de dados, selecione o banco de dados no qual o link do recurso foi criado (neste caso, a pilha AWS CloudFormation criou o banco de dados `lakeformation_tutorial_cross_account_database_consumer`).
9. Selecione Criar.

O link do recurso aparece em Catálogo de dados, Tabelas.

Crie uma LF-tag e atribua-a ao banco de dados de destino

As tags do Lake Formation residem no mesmo Catálogo de Dados dos recursos. Isso significa que as tags criadas na conta do produtor não estão disponíveis para uso ao conceder acesso aos links de recursos na conta do consumidor. Você precisa criar um conjunto separado de LF-tags na conta do consumidor para usar o controle de acesso baseado em LF-tags ao compartilhar os links de recursos na conta do consumidor.

1. Defina a LF-tag na conta do consumidor. Para este tutorial, usamos a chave `Division` e os valores `sales`, `marketing` e `analyst`.
2. Atribua a chave da LF-tag `Division` e o valor `analyst` ao banco de dados C, onde o link de recurso é criado.

Conceda permissão de dados da LF-tag ao consumidor

Como etapa final, conceda permissão de dados da LF-tag ao consumidor.

1. No painel de navegação, em `Permissões`, `Permissões do Data Lake`, selecione `Conceder`.
2. Para `Entidade principais`, selecione `Usuários e perfis do IAM` e selecione o usuário `DataAnalyst`.
3. Para `Recursos de LF-tags` ou de catálogo, selecione `Recursos` correspondentes a LF-tags(recomendado).
4. Selecione `Chave Divisão` e `analista de valor`.
5. Para `Permissões de banco de dados`, selecione `Descrever` em `Permissões de banco de dados`.
6. Para `Permissões de tabela e coluna`, selecione `Selecionar` e `Descrever` em `Permissões de tabela`.
7. Selecione `Conceder`.
8. Repita essas etapas para o usuário `DataAnalyst`, onde a chave LF-tag é `Confidentiality` e o valor é `public`.

Nesse ponto, o usuário analista de dados na conta do consumidor deve ser capaz de encontrar o link do banco de dados e do recurso, e consultar a tabela compartilhada por meio do console do Athena em <https://console.aws.amazon.com/athena/>. Caso contrário, confirme se o seguinte está configurado corretamente:

- O link do recurso é criado para a tabela compartilhada
- Você concedeu ao usuário acesso à LF-tag compartilhada pela conta do produtor
- Você concedeu ao usuário acesso à LF-tag associada ao link do recurso e ao banco de dados no qual o link do recurso foi criado
- Verifique se você atribuiu a LF-tag correta ao link do recurso e ao banco de dados no qual o link do recurso foi criado

## Etapa 4: Implementar o método de recurso nomeado

Para usar o método de recurso nomeado, orientamos você nas seguintes etapas de alto nível:


1. Opcionalmente, revogue a permissão para `IAMAllowedPrincipals` no banco de dados, nas tabelas e nas colunas.
2. Conceda a permissão de dados à conta do consumidor.
3. Aceitar um compartilhamento de recursos de AWS Resource Access Manager.
4. Crie um link de recurso para a tabela compartilhada.
5. Conceda permissão de dados para a tabela compartilhada ao consumidor.
6. Conceda permissão de dados para o link de recurso ao consumidor.

Revogue a permissão para **`IAMAllowedPrincipals`** no banco de dados, nas tabelas e nas colunas (opcional).

- No início deste tutorial, alteramos as configurações do Lake Formation Data Catalog. Se você pulou essa parte, esta etapa é necessária. Para obter instruções, consulte a etapa opcional na seção anterior.

Conceda a permissão de dados à conta do consumidor.

1.

 Note

Se você estiver conectado à conta do produtor como outro usuário, saia primeiro.

Faça login no console do Lake Formation em <https://console.aws.amazon.com/lakeformation/> usando a conta do produtor administrador da data lake usando o ID Conta da AWS, o nome de usuário do IAM (o padrão é `DataLakeAdminProducer`) e a senha especificada durante a criação da pilha AWS CloudFormation.

2. Na página Permissões, em Permissões do Data Lake, selecione Conceder.
3. Em Entidades Principais, selecione Contas externas e insira uma ou mais IDs Conta da AWS ou IDs de organizações AWS. Para obter mais informações, consulte: [OrganizaçõesAWS](#).



As organizações às quais a conta do produtor pertence e as Contas da AWS dentro da mesma organização aparecem automaticamente. Caso contrário, insira manualmente o ID da conta ou o ID da organização.

4. Para LF-tags ou recursos de catálogo, selecione `Named data catalog resources`.
5. Em Bancos de dados, selecione o banco de dados `lakeformation_tutorial_cross_account_database_named_resource`.
6. Selecione Adicionar LF-tag.
7. Em Tabelas, selecione Todas as tabelas.
8. Para Permissões de coluna de tabela, selecione Selecionar e Descrever em Permissões de tabela.
9. Selecione Selecionar e Descrever em Permissões concedíveis.
10. Opcionalmente, para Permissões de dados, escolha Acesso simples baseado em coluna se for necessário o gerenciamento de permissões em nível de coluna.
11. Selecione Conceder.

Se você não revogou a permissão para `IAMAllowedPrincipals`, receberá um erro de falha na Concessão de permissões. Nesse ponto, você deve ver a tabela de destino sendo compartilhada via AWS RAM com a conta do consumidor em Permissões, Permissões de dados.

Aceitar um compartilhamento de recursos de AWS RAM

#### Note

Essa etapa é necessária somente para o compartilhamento baseado em Conta da AWS, não para o compartilhamento baseado na organização.

1. Faça login no console AWS em <https://console.aws.amazon.com/connect/> usando o administrador do data lake da conta do consumidor usando o nome de usuário do IAM (o padrão é `DatalakeAdminConsumer`) e a senha especificada durante a criação da pilha AWS CloudFormation.
2. No console AWS RAM, no painel de navegação, em Compartilhado comigo, Compartilhamento de recursos, selecione o recurso compartilhado do Lake Formation. O status deve ser Pendente.
3. Selecione Ações e Concessão.

#### 4. Confirme os detalhes do recurso e selecione Aceitar compartilhamento de recursos.

Nesse momento, o administrador do Data Lake da conta do consumidor deve conseguir encontrar o recurso compartilhado no console do Lake Formation (<https://console.aws.amazon.com/lakeformation/>) em Catálogo de dados, Bancos de dados.

#### Criar um link de recurso para a tabela compartilhada

- Siga as instruções em [Etapa 3: Implementar o compartilhamento entre contas usando o método de controle de acesso baseado em tags](#) (etapa 6) para criar um link de recurso para uma tabela compartilhada. Dê um nome ao link do recurso `amazon_reviews_table_named_resource_resource_link`. Crie o link do recurso no banco de dados `lakeformation_tutorial_cross_account_database_consumer`.

#### Conceda permissão de dados para a tabela compartilhada ao consumidor

Para conceder permissão de dados para a tabela compartilhada ao consumidor, conclua as etapas a seguir:

1. No console do Lake Formation (<https://console.aws.amazon.com/lakeformation/>), em Permissões, Permissões do data lake, selecione Conceder.
2. Para Entidade principais, selecione Usuários e perfis do IAM e selecione o usuário `DataAnalyst`.
3. Para LF-tags ou recursos de catálogo, selecione Recursos de catálogo de dados nomeados.
4. Em Bancos de dados, selecione o banco de dados `lakeformation_tutorial_cross_account_database_named_resource`. Se você não encontrar o banco de dados na lista suspensa, selecione Carregar mais.
5. Em Tabelas, selecione a tabela `amazon_reviews_table_named_resource`.
6. Para Permissões de tabela e coluna, selecione Selecionar e Descrever em Permissões de tabela.
7. Selecione Conceder.

#### Conceda permissão de dados para o link de recurso ao consumidor

Além de conceder permissão ao usuário do Data Lake para acessar a tabela compartilhada, você também precisa conceder permissão ao usuário do Data Lake para acessar o link do recurso.

1. No console do Lake Formation (<https://console.aws.amazon.com/lakeformation/>), em Permissões, Permissões do data lake, selecione Conceder.
2. Para Entidade principais, selecione Usuários e perfis do IAM e selecione o usuário DataAnalyst.
3. Para LF-tags ou recursos de catálogo, selecione Recursos de catálogo de dados nomeados.
4. Em Bancos de dados, selecione o banco de dados lakeformation\_tutorial\_cross\_account\_database\_consumer. Se você não encontrar o banco de dados na lista suspensa, selecione Carregar mais.
5. Em Tabelas, selecione a tabela amazon\_reviews\_table\_named\_resource\_resource\_link.
6. Para Permissões de links de recursos, selecione Descrever em Permissões de links de recursos.
7. Selecione Conceder.

Nesse ponto, o usuário analista de dados na conta do consumidor deve ser capaz de encontrar o banco de dados e o link do recurso e consultar a tabela compartilhada por meio do console do Athena.

Caso contrário, confirme se o seguinte está configurado corretamente:

- O link do recurso é criado para a tabela compartilhada
- Você concedeu ao usuário acesso à tabela compartilhada pela conta do produtor
- Você concedeu ao usuário acesso ao link de recurso e ao banco de dados para o qual o link de recurso foi criado

## Etapa 5: Limpar os recursos do AWS

Para evitar cobranças indesejadas em seu Conta da AWS, você pode excluir os recursos AWS que usou para este tutorial.

1. Faça login no console do Lake Formation em <https://console.aws.amazon.com/lakeformation/> usando a conta do produtor e exclua ou altere o seguinte:
  - Compartilhamento de recursos AWS Resource Access Manager
  - Tags do Lake Formation
  - Pilha AWS CloudFormation
  - Configurações do Lake Formation

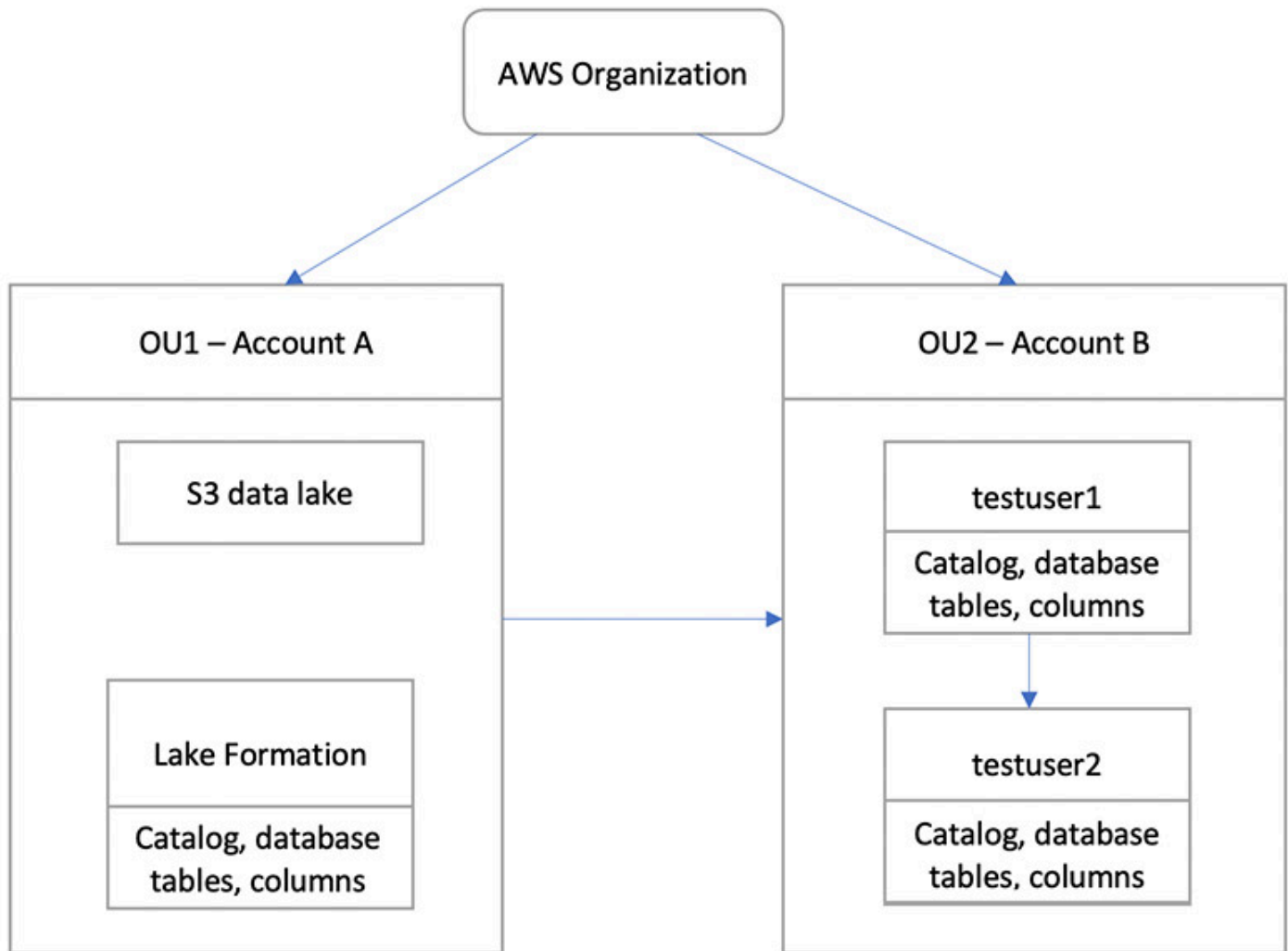
- AWS Glue Data Catalog
2. Faça login no console do Lake Formation em <https://console.aws.amazon.com/lakeformation/> usando a conta do consumidor e exclua ou altere o seguinte:
    - Tags do Lake Formation
    - Pilha AWS CloudFormation

## Como compartilhar um data lake usando o controle de acesso refinado do Lake Formation

Este tutorial fornece instruções passo a passo sobre como você pode compartilhar conjuntos de dados de forma rápida e fácil usando o Lake Formation ao gerenciar múltiplas Contas da AWS com o AWS Organizations. Você define permissões específicas para controlar o acesso a dados confidenciais.

Os procedimentos a seguir também mostram como um administrador de data lake da Conta A pode fornecer acesso refinado à Conta B, e como um usuário na Conta B, atuando como administrador de dados, pode conceder acesso refinado à tabela compartilhada para outros usuários em suas contas. Os administradores de dados em cada conta podem delegar o acesso de forma independente a seus próprios usuários, dando autonomia a cada equipe ou linha de negócios (LOB).

Este caso específico pressupõe que você esteja usando o AWS Organizations para gerenciar suas Contas da AWS. O usuário da Conta A em uma unidade organizacional (OU1) concede acesso aos usuários da Conta B na OU2. Você pode usar a mesma abordagem quando não estiver usando organizações, como quando você tem apenas algumas contas. O diagrama a seguir ilustra o controle de acesso refinado de conjuntos de dados em um data lake. O data lake está disponível na Conta A. O administrador do data lake da Conta A fornece acesso refinado à Conta B. O diagrama também mostra que um usuário da Conta B fornece acesso em nível de coluna à tabela de data lake da Conta A para outro usuário na Conta B.



## Tópicos

- [Público-alvo](#)
- [Pré-requisitos](#)
- [Etapa 1: Forneça acesso refinado a outra conta](#)
- [Etapa 2: Forneça acesso refinado a um usuário na mesma conta](#)

## Público-alvo

Este tutorial é destinado a administradores de dados, engenheiros de dados e analistas de dados. A tabela a seguir lista os perfis usados neste tutorial:

Perfil	Descrição
Administrador do IAM	Usuário que tem a política gerenciada pela <code>AWS: AdministratorAccess</code> .
Administrador do data lake	Usuário que tem a política AWS gerenciada: anexada ao perfil <code>AWSLakeFormationDataAdmin</code> .
Analista de dados	Usuário que tem a política AWS gerenciada: <code>AmazonAthenaFullAccess</code> anexada.

## Pré-requisitos

Antes de iniciar este tutorial, você deve ter uma conta Conta da AWS que possa ser usada para fazer login como um usuário administrativo com as permissões corretas. Para obter mais informações, consulte [Concluir as tarefas iniciais de configuração do AWS](#).

O tutorial pressupõe que você esteja familiarizado com o IAM. Para obter informações sobre o IAM, consulte o [Guia do Usuário do IAM](#).

Você precisa dos seguintes recursos para este tutorial:

- Duas unidades organizacionais:
  - OU1 — Contém a Conta A
  - OU2 — Contém a Conta B
- Um local (bucket) do data lake do Amazon S3 na Conta A.
- Um usuário administrador de data lake na Conta A. Você pode criar um administrador de data lake usando o console do Lake Formation (<https://console.aws.amazon.com/lakeformation/>) ou a operação `PutDataLakeSettings` da API do Lake Formation.
- Lake Formation configurado na Conta A e o local do data lake do Amazon S3 registrada com o Lake Formation na Conta A.
- Dois usuários na Conta B com as seguintes políticas gerenciadas pelo IAM:
  - `testuser1` — Tem as políticas gerenciadas pela AWS anexadas ao `AWSLakeFormationDataAdmin`.

- testuser2 — Tem as políticas gerenciadas pela AWS anexadas ao AmazonAthenaFullAccess.
- Um banco de dados testdb no banco de dados Lake Formation para a Conta B.

## Etapa 1: Forneça acesso refinado a outra conta

Saiba como um administrador de data lake da Conta A fornece acesso refinado à Conta B.

Conceda acesso refinado a outra conta

1. Faça login no AWS Management Console em <https://console.aws.amazon.com/connect/> com a Conta A como administrador do data lake.
2. Abra o console Lake Formation (<https://console.aws.amazon.com/lakeformation/>) e selecione Começar.
3. No painel de navegação, selecione Bancos de dados.
4. Selecione Criar banco de dados.
5. Na seção informações do Banco de dados, selecione Banco de dados.
6. Em Nome, digite um nome (para este tutorial, usamos sampledb01).
7. Certifique-se de que a opção Usar somente o controle de acesso do IAM para novas tabelas nesse banco de dados não esteja selecionada. Deixar essa opção desmarcada nos permite controlar o acesso a partir do Lake Formation.
8. Selecione Criar banco de dados.
9. Na página Bancos de dados, selecione seu banco de dados sampledb01.
10. No menu Ações, selecione Conceder.
11. Na seção Conceder permissões, selecione Conta externa.
12. Para a ID da Conta da AWS ou ID da organização da AWS, insira o ID da Conta B na OU2.
13. Em Tabela, selecione a tabela à qual você deseja que a Conta B tenha acesso (neste caso, usamos a tabela acc\_a\_area). Opcionalmente, você pode conceder acesso às colunas dentro da tabela, o que fazemos neste caso.
14. Em Incluir colunas, selecione as colunas às quais você deseja que a Conta B tenha acesso (para este caso, concedemos permissões para digitar, nomear e identificadores).
15. Em Colunas, selecione Incluir colunas.
16. Em Permissões de tabela, selecione Selecionar.

17. Para Permissões concedidas, marque Selecionar. As permissões concedidas são necessárias para que os usuários administradores na Conta B possam conceder permissões a outros usuários na Conta B.
18. Selecione Conceder.
19. No painel de navegação, selecione TableTabelas.
20. É possível ver uma conexão ativa na seção de organizações da Contas da AWS e da AWS com acesso.

### Crie um link de recurso

Serviços integrados, como o Amazon Athena, não podem acessar diretamente bancos de dados ou tabelas entre contas. Portanto, você precisa criar um link de recurso para que o Athena possa acessar links de recursos em sua conta para bancos de dados e tabelas em outras contas. Crie um link de recurso para a tabela (`acc_a_area`) para que os usuários da Conta B possam consultar seus dados com o Athena.

1. Faça login no console da AWS em <https://console.aws.amazon.com/connect/> na Conta B como `testuser1`.
2. No console do Lake Formation (<https://console.aws.amazon.com/lakeformation/>), no painel de navegação, selecione Tabelas. Você deve ver as tabelas às quais a Conta A forneceu acesso.
3. Selecione a tabela `acc_a_area`.
4. No menu Ações, selecione Criar link de recurso
5. Em Nome do link do recurso, digite um nome (para este tutorial, `acc_a_area_rl`).
6. Em Banco de dados, selecione o banco de dados `testdb`.
7. Selecione Criar.
8. No painel de navegação, selecione TableTabelas.
9. Selecione a tabela `acc_b_area_rl`.
10. No menu Ações, selecione Exibir dados

Você será redirecionado para o console do Athena, onde deverá ver o banco de dados e a tabela.

Agora você pode executar uma consulta na tabela para ver o valor da coluna para o qual o acesso foi fornecido ao `testuser1` da Conta B.



## Etapa 2: Forneça acesso refinado a um usuário na mesma conta

Esta seção mostra como um usuário na Conta B (testuser1), atuando como administrador de dados, fornece acesso refinado a outro usuário na mesma conta (testuser2) ao nome da coluna na tabela compartilhada `aac_b_area_r1`.

Conceda acesso refinado a um usuário na mesma conta

1. Faça login no console da AWS em <https://console.aws.amazon.com/connect/> na Conta B como testuser1.
2. No console do Lake Formation, no painel de navegação, selecione Tabelas.

É possível conceder permissões em uma tabela por meio do link do recurso. Para fazer isso, na página Tabelas, selecione o link do recurso `acc_b_area_r1`, e no menu Ações, selecione Conceder no destino.

3. Na seção Conceder permissões, selecione Minha conta.
4. Para Usuários e perfis do IAM, selecione o usuário testuser2.
5. Em Coluna, escolha o nome da coluna.
6. Em Permissões de tabela, selecione Selecionar.
7. Selecione Conceder.

Após criar um link de recurso, somente você poderá visualizá-lo e acessá-lo. Para permitir que outros usuários da sua conta acessem o link do recurso, você precisa conceder permissões no próprio link do recurso. Você precisa conceder as permissões Descrever ou Descartar. Na página Tabelas, selecione sua tabela novamente e, no menu Ações, selecione Conceder.

8. Na seção Conceder permissões, selecione Minha conta.
9. Para Usuários e perfis do IAM, selecione o usuário testuser2.
10. Para permissões de links de recursos, selecione Descrever.
11. Selecione Conceder.
12. Faça login no console da AWS na Conta B como testuser2.

No console do Athena (<https://console.aws.amazon.com/athena/>), você deve ver o banco de dados e a tabela `acc_b_area_r1`. Agora você pode executar uma consulta na tabela para ver o valor da coluna que testuser2 tem acesso.

# Permissões de integração ao Lake Formation

AWS Lake Formation usa o AWS Glue Data Catalog para armazenar metadados para os dados do Amazon S3 na forma de bancos de dados e tabelas. As tabelas armazenam informações sobre os dados subjacentes, incluindo informações de esquema, informações de partição e localização dos dados. Bancos de dados são coleções de tabelas. O Catálogo de dados também contém links de recursos, que são links para bancos de dados e tabelas compartilhados em contas externas e são usados para acesso entre contas aos dados no data lake. Cada conta AWS tem um catálogo de dados por região AWS.

O Lake Formation fornece um modelo de permissões do sistema de gerenciamento de banco de dados relacional (RDBMS) para conceder ou revogar o acesso a bancos de dados, tabelas e colunas no Catálogo de dados com dados subjacentes no Amazon S3.

Antes de aprender sobre os detalhes do modelo de permissões do Lake Formation, é útil revisar as seguintes informações básicas:

- Data lakes gerenciados pelo Lake Formation residem em locais designados no Amazon Simple Storage Service (Amazon S3).
- O Lake Formation mantém um catálogo de dados que contém metadados sobre dados de origem a serem importados para seus data lakes, como dados em logs e bancos de dados relacionais, e sobre dados em seus lagos de dados no Amazon S3. Os metadados são organizados como bancos de dados e tabelas. As tabelas de metadados contêm esquema, localização, particionamento e outras informações sobre os dados que elas representam. Bancos de dados de metadados são coleções de tabelas.
- O Catálogo de dados do Lake Formation é o mesmo Catálogo de dados usado por AWS Glue. Você pode usar crawlers AWS Glue para criar tabelas do Catálogo de dados e pode usar tarefas AWS Glue de extração, transformação e carregamento (ETL) para preencher os dados subjacentes em seus data lakes.
- Os bancos de dados e tabelas no Catálogo de dados são chamados de recursos do Catálogo de dados. As tabelas no Catálogo de dados são chamadas de tabelas de metadados para diferenciá-las das tabelas nas fontes de dados ou dos dados tabulares no Amazon S3. Os dados para os quais as tabelas de metadados apontam no Amazon S3 ou nas fontes de dados são chamados de dados subjacentes.
- Uma entidade principal é um usuário ou função, um usuário ou grupo do Amazon QuickSight, um usuário ou grupo que se autentica com o Lake Formation por meio de um provedor de SAML ou,

para controle de acesso entre contas, um ID de conta AWS, ID da organização ou ID da unidade organizacional.

- Crawlers AWS Glue criam tabelas de metadados, mas você também pode criar tabelas de metadados manualmente com o console do Lake Formation, a API ou o AWS Command Line Interface (AWS CLI). Ao criar uma tabela de metadados, você deve especificar uma localização. Quando você cria um banco de dados, o local é opcional. Os locais das tabelas podem ser locais do Amazon S3 ou locais de fonte de dados, como um banco de dados do Amazon Relational Database Service (Amazon RDS). Os locais do banco de dados são sempre locais do Amazon S3.
- Serviços que se integram ao Lake Formation, como Amazon Athena e Amazon Redshift, podem acessar o Catálogo de dados para obter metadados e verificar a autorização para executar consultas. Para obter uma lista completa de serviços integrados, consulte [Integrações de serviços da AWS com o Lake Formation](#).

## Tópicos

- [Visão geral das permissões do Lake Formation](#)
- [Referência de personas e permissões do IAM do Lake Formation](#)
- [Alterando as configurações padrão do seu data lake](#)
- [Permissões implícitas do Lake Formation](#)
- [Referência de permissões do Lake Formation](#)
- [Integrar o Centro de Identidade do IAM](#)
- [Adicionar uma localização do Amazon S3 ao seu data lake](#)
- [Modo de acesso híbrido](#)
- [Criando tabelas e bancos de dados do Catálogo de dados](#)
- [Importação de dados usando fluxos de trabalho no Lake Formation](#)

## Visão geral das permissões do Lake Formation

Há dois tipos principais de permissões em AWS Lake Formation:

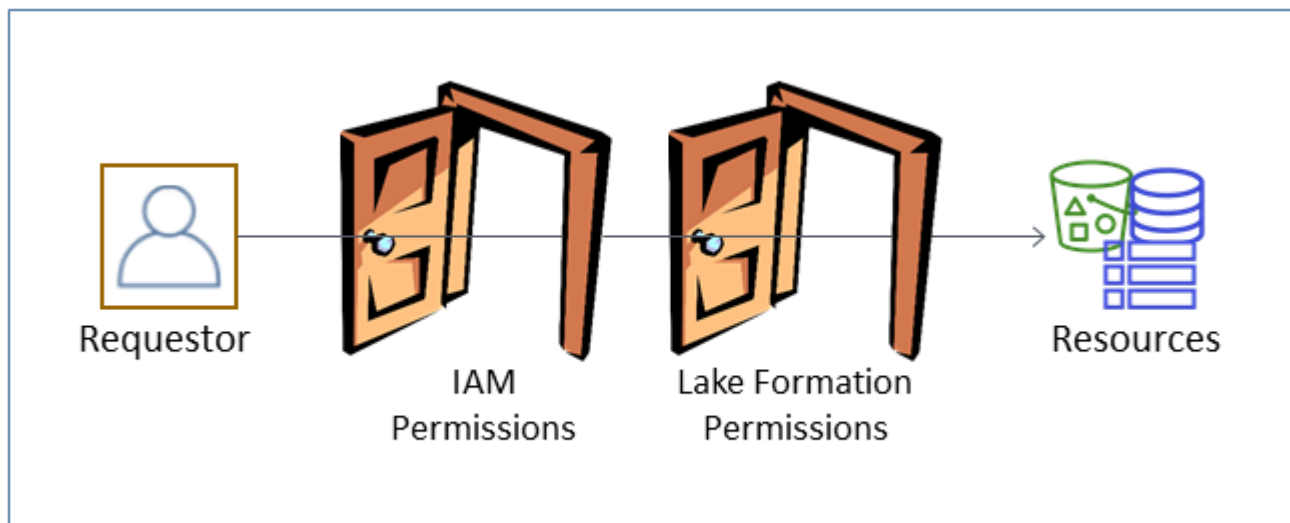
- Acesso aos metadados – Permissões nos recursos do Catálogo de dados (Permissões do Catálogo de dados).

Essas permissões permitem que as entidades principais criem, leiam, atualizem e excluam bancos de dados e tabelas de metadados no Catálogo de dados.

- Acesso aos dados subjacentes – Permissões em locais no Amazon Simple Storage Service (Amazon S3) (Permissões de acesso a dados e Permissões de localização de dados).
  - As permissões do data lake permitem que as entidades principais leiam e gravem dados em locais subjacentes do Amazon S3 – dados apontados pelos recursos do Catálogo de dados.
  - As permissões de localização de dados permitem que as entidades principais criem e alterem bancos de dados e tabelas de metadados que apontam para os locais específicos do Amazon S3.

Para ambas as áreas, o Lake Formation usa uma combinação de permissões e permissões AWS Identity and Access Management do Lake Formation (IAM). O modelo de permissões do IAM consiste em políticas do IAM. O modelo de permissões do Lake Formation é implementado como comandos CONCEDER/REVOGAR no estilo DBMS, como `Grant SELECT on tableName to userName`.

Quando uma entidade principal faz uma solicitação para acessar os recursos do Catálogo de dados ou os dados subjacentes, para que a solicitação seja bem-sucedida, ela deve passar pelas verificações de permissão do IAM e do Lake Formation.



As permissões do Lake Formation controlam o acesso aos recursos do Catálogo de dados, aos locais do Amazon S3 e aos dados subjacentes nesses locais. As permissões do IAM controlam o acesso ao Lake Formation, APIs AWS Glue e recursos. Portanto, embora você possa ter a permissão do Lake Formation para criar uma tabela de metadados no Catálogo de dados (`CREATE_TABLE`), sua operação falhará se você não tiver a permissão do IAM na API

`glue:CreateTable`. (Por que uma permissão `glue:?` Porque o Lake Formation usa o Catálogo de dados AWS Glue.)

### Note

As permissões do Lake Formation se aplicam somente na região em que foram concedidas.

AWS Lake Formation exige que cada entidade principal (usuário ou perfil) seja autorizada a realizar ações nos recursos gerenciados pelo Lake Formation. Uma entidade principal recebe as autorizações necessárias do administrador do data lake ou de outra entidade principal com as permissões para conceder as permissões do Lake Formation.

Ao conceder uma permissão de Lake Formation a uma entidade principal, você pode, opcionalmente, conceder a capacidade de passar essa permissão para outra entidade principal.

Você pode usar a API do Lake Formation, a AWS Command Line Interface (AWS CLI) ou as páginas Permissões de dados e Localizações de dados do console do Lake Formation para conceder e revogar as permissões do Lake Formation.

## Métodos para controle de acesso de alta granularidade

Com um data lake, o objetivo é ter um controle de acesso de alta granularidade aos dados. No Lake Formation, isso significa controle de acesso de alta granularidade aos recursos do Catálogo de dados e aos locais do Amazon S3. Você pode obter um controle de acesso de alta granularidade com um dos seguintes métodos.

Método	Permissões do Lake Formation	Permissões do IAM	Comentários
Método 1	Aberto	Alta granularidade	<p>Esse é o método padrão para compatibilidade com versões anteriores com AWS Glue.</p> <ul style="list-style-type: none"> <li>Aberto significa que a permissão especial <code>Super</code> é concedida ao grupo <code>IAMAllowedPrincipals</code>, onde <code>IAMAllowedPrincipals</code> é criada</li> </ul>

Método	Permissões do Lake Formation	Permissões do IAM	Comentários
			<p>automaticamente e inclui todos os usuários e perfis do IAM que têm permissão para acessar seus recursos do Catálogo de dados por meio de suas políticas do IAM, e a permissão Super permite que uma entidade principal execute todas as operações suportadas do Lake Formation no banco de dados ou na tabela em que foi concedida. Isso efetivamente faz com que o acesso aos recursos do Catálogo de dados e aos locais do Amazon S3 seja controlado exclusivamente pelas políticas do IAM. Para obter mais informações, consulte <a href="#">Alterando as configurações padrão do seu data lake</a> e <a href="#">Atualizar as permissões AWS Glue de dados para o modelo AWS Lake Formation</a>.</p> <ul style="list-style-type: none"><li>Alta granularidade significa que as políticas do IAM controlam todo o acesso aos recursos do Catálogo de dados e aos buckets individuais do Amazon S3.</li></ul> <p>No console do Lake Formation, esse método aparece como Use apenas controle de acesso do IAM.</p>

Método	Permissões do Lake Formation	Permissões do IAM	Comentários
Método 2	Alta granularidade	Baixa granularidade	<p>Este é o método recomendado.</p> <ul style="list-style-type: none"> <li>Alta granularidade significa conceder permissões limitadas do Lake Formation a entidades principais individuais sobre os recursos do Catálogo de dados, os locais do Amazon S3 e os dados subjacentes nesses locais.</li> <li>Baixa granularidade significa permissões mais amplas em operações individuais e no acesso aos locais do Amazon S3. Por exemplo, uma política de IAM de baixa granularidade pode incluir <code>"glue:*"</code> ou <code>"glue:Create*"</code> ao invés de <code>"glue:CreateTables"</code>, deixando que as permissões do Lake Formation controlem se uma entidade principal pode ou não criar objetos de catálogo. Isso também significa dar às entidades principais acesso às APIs de que precisam para realizar seu trabalho, mas bloquear outras APIs e recursos. Por exemplo, você pode criar uma política do IAM que permita que uma entidade principal crie recursos do Catálogo de dados e crie e execute fluxos de trabalho, mas não permita a criação de conexões AWS Glue ou funções definidas pelo usuário. Veja os exemplos mais adiante nesta seção.</li> </ul>

**⚠ Important**

Esteja ciente do seguinte:

- Por padrão, o Lake Formation tem as configurações de Controle de acesso apenas para uso do IAM habilitadas para compatibilidade com o comportamento existente do Catálogo de dados AWS Glue. Recomendamos que você desative essas configurações após a transição para o uso das permissões do Lake Formation. Para obter mais informações, consulte [Alterando as configurações padrão do seu data lake](#).
- Os administradores do Data Lake e os criadores de bancos de dados têm permissões implícitas do Lake Formation que você precisa entender. Para obter mais informações, consulte [Permissões implícitas do Lake Formation](#).

## Controle de acesso a metadados

Para controle de acesso aos recursos do Catálogo de dados, a discussão a seguir pressupõe controle de acesso de alta granularidade com permissões do Lake Formation e controle de acesso de baixa granularidade com políticas do IAM.

Há dois métodos distintos para conceder permissões do Lake Formation nos recursos do Catálogo de dados:

- Controle de acesso a recursos nomeados – Com esse método, você concede permissões em bancos de dados ou tabelas específicos especificando nomes de bancos de dados ou tabelas. As concessões têm o seguinte formato:

Conceda Permissões a entidades principais sobre recursos [com opção de concessão].

Com a opção de concessão, você pode permitir que o beneficiário conceda as permissões a outras entidades principais.

- Controle de acesso baseado em tags – Com esse método, você atribui uma ou mais LF-Tags aos bancos de dados, tabelas e colunas do Catálogo de dados e concede permissões em uma ou mais LF-Tags às entidades principais. Cada LF-tag é um par de chave-valor, como `department=sales`. Uma entidade principal que tenha LF-Tags que correspondam às LF-Tags em um recurso do Catálogo de dados pode acessar esse recurso. Esse método é recomendado para data lakes com um grande número de bancos de dados e tabelas. Isso é explicado em detalhes em [Controle de acesso baseado em tags do Lake Formation](#).



As permissões que uma entidade principal tem sobre um recurso são a união das permissões concedidas pelos dois métodos.

A tabela a seguir resume as permissões disponíveis do Lake Formation nos recursos do Catálogo de dados. Os títulos das colunas indicam o recurso no qual a permissão é concedida.

Catálogo	Banco de dados	Tabela
CREATE_DATABASE	CREATE_TABLE	ALTER
	ALTER	DROP
	DROP	DESCRIBE
	DESCRIBE	SELECT*
		INSERT*
		DELETE*

Por exemplo, a permissão CREATE\_TABLE é concedida em um banco de dados. Isso significa que a entidade principal tem permissão para criar tabelas nesse banco de dados.

As permissões com um asterisco (\*) são concedidas nos recursos do Catálogo de dados, mas se aplicam aos dados subjacentes. Por exemplo, a permissão DROP em uma tabela de metadados permite que você remova a tabela do Catálogo de dados. No entanto, a permissão DELETE concedida na mesma tabela permite que você exclua os dados subjacentes da tabela no Amazon S3, usando, por exemplo, uma instrução SQL. Com essas permissões, você também pode visualizar a tabela no console do Lake Formation e recuperar informações sobre a tabela com a API AWS Glue. Portanto, SELECT, INSERT e DELETE são permissões do Catálogo de dados e permissões de acesso aos dados.

Ao conceder SELECT em uma tabela, você pode adicionar um filtro que inclua ou exclua uma ou mais colunas. Isso permite um controle de acesso de alta granularidade nas colunas da tabela de metadados, limitando as colunas que os usuários de serviços integrados podem ver ao executar consultas. Esse recurso não está disponível usando apenas as políticas do IAM.

Também há uma permissão especial chamada Super. A permissão Super permite que uma entidade principal execute todas as operações suportadas do Lake Formation no banco de dados

ou na tabela em que ela foi concedida. Essa permissão pode coexistir com as outras permissões do Lake Formation. Por exemplo, você pode conceder `Super`, `SELECT` e `INSERT` em uma tabela de metadados. A entidade principal pode realizar todas as ações suportadas na tabela e, quando você revoga `Super`, as permissões `SELECT` e `INSERT` permanecem.

Para obter detalhes sobre cada permissão, consulte [Referência de permissões do Lake Formation](#).

#### Important

Para poder ver uma tabela do Catálogo de dados criada por outro usuário, você deve ter pelo menos uma permissão do Lake Formation na tabela. Se você tiver pelo menos uma permissão na tabela, também poderá ver o banco de dados que contém a tabela.

Você pode conceder ou revogar as permissões do Catálogo de dados usando o console do Lake Formation, a API ou o AWS Command Line Interface (AWS CLI). Veja a seguir um exemplo de um comando AWS CLI que concede ao usuário permissão `datalake_user1` para criar tabelas no banco de dados `retail`.

```
aws lakeformation grant-permissions --principal
DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1
--permissions "CREATE_TABLE" --resource '{ "Database": {"Name":"retail"} }'
```

Veja a seguir um exemplo de uma política do IAM de controle de acesso de baixa granularidade que complementa o controle de acesso de alta granularidade com as permissões do Lake Formation. Ele permite todas as operações em qualquer banco de dados ou tabela de metadados.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "glue:*Database*",
        "glue:*Table*",
        "glue:*Partition*"
      ],
      "Resource": "*"
    }
  ]
}
```

```
}
```

O próximo exemplo também é de baixa granularidade, mas um pouco mais restritivo. Ele permite operações somente para leitura em todos os bancos de dados e tabelas de metadados no Catálogo de dados na conta e região designadas.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "glue:GetTables",
        "glue:SearchTables",
        "glue:GetTable",
        "glue:GetDatabase",
        "glue:GetDatabases"
      ],
      "Resource": "arn:aws:glue:us-east-1:111122223333:*"
    }
  ]
}
```

Compare essas políticas com a política a seguir, que implementa o controle de acesso de alta granularidade baseado em IAM. Ele concede permissões somente em um subconjunto de tabelas no banco de dados de metadados de gerenciamento de relacionamento com o cliente (CRM) na conta e região designadas.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "glue:GetTables",
        "glue:SearchTables",
        "glue:GetTable",
        "glue:GetDatabase",
        "glue:GetDatabases"
      ],
      "Resource": [
        "arn:aws:glue:us-east-1:111122223333:catalog",

```

```
        "arn:aws:glue:us-east-1:111122223333:database/CRM",  
        "arn:aws:glue:us-east-1:111122223333:table/CRM/P*" ]  
    }  
]  
}
```

Para obter mais exemplos de políticas de controle de acesso de baixa granularidade, consulte [Referência de personas e permissões do IAM do Lake Formation](#).

## Controle de acesso a dados subjacente

Quando um serviço AWS integrado solicita acesso aos dados em um local do Amazon S3 que é controlado por AWS Lake Formation, o Lake Formation fornece credenciais temporárias para acessar os dados.

Para permitir que o Lake Formation controle o acesso aos dados subjacentes em um local do Amazon S3, você registra esse local no Lake Formation.

Depois de registrar uma localização no Amazon S3, você pode começar a conceder as seguintes permissões do Lake Formation:

- Permissões de acesso aos dados (SELECT, INSERT e DELETE) nas tabelas do Catálogo de dados que apontam para esse local.
- Permissões de localização de dados nesse local.

As permissões de localização de dados do Lake Formation controlam a capacidade de criar recursos do Catálogo de dados que apontam para locais específicos do Amazon S3. As permissões de localização de dados fornecem uma camada extra de segurança aos locais dentro do data lake. Ao conceder a permissão CREATE\_TABLE ou ALTER a uma entidade principal, você também concede permissões de localização de dados para limitar os locais para os quais a entidade principal pode criar ou alterar tabelas de metadados.

Os locais do Amazon S3 são buckets ou prefixos em um bucket, mas não objetos individuais do Amazon S3.

Você pode conceder permissões de localização de dados a uma entidade principal usando o console do Lake Formation, a API ou o AWS CLI. A forma geral de uma subvenção é a seguinte:

```
grant DATA_LOCATION_ACCESS to principal on S3 location [with grant option]
```

Se você incluir `with grant option`, o beneficiário poderá conceder as permissões a outras entidades principais.

Lembre-se de que as permissões do Lake Formation sempre funcionam em combinação com as permissões AWS Identity and Access Management (IAM) para um controle de acesso de alta granularidade. Para permissões de leitura/gravação em dados subjacentes do Amazon S3, as permissões do IAM são concedidas da seguinte forma:

Ao registrar um local, você especifica um perfil do IAM que concede permissões de leitura/gravação nesse local. A Lake Formation assume essa função ao fornecer credenciais temporárias para serviços integrados AWS. Uma função típica pode ter a seguinte política anexada, em que o local registrado é o bucket `awsexamplebucket`.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "s3:PutObject",
        "s3:GetObject",
        "s3:DeleteObject"
      ],
      "Resource": [
        "arn:aws:s3:::awsexamplebucket/*"
      ]
    },
    {
      "Effect": "Allow",
      "Action": [
        "s3:ListBucket"
      ],
      "Resource": [
        "arn:aws:s3:::awsexamplebucket"
      ]
    }
  ]
}
```

O Lake Formation fornece uma função vinculada ao serviço que você pode usar durante o registro para criar automaticamente políticas como essa. Para obter mais informações, consulte [Usar perfis vinculados ao serviço para o Lake Formation](#).

Portanto, registrar um local do Amazon S3 concede as permissões necessárias `s3`: do IAM nesse local, onde as permissões são especificadas pela função usada para registrar o local.

#### Important

Evite registrar um bucket do Amazon S3 que tenha o Solicitante paga ativado. Para buckets registrados no Lake Formation, a função usada para registrar o bucket é sempre vista como solicitante. Se o bucket for acessado por outra conta AWS, o proprietário do bucket será cobrado pelo acesso aos dados se a função pertencer à mesma conta do proprietário do bucket.

Para acesso de leitura/gravação aos dados subjacentes, além das permissões do Lake Formation, as entidades principais também precisam da seguinte permissão do IAM:

`lakeformation:GetDataAccess`

Com essa permissão, o Lake Formation concede a solicitação de credenciais temporárias para acessar os dados.

#### Note

O Amazon Athena exige que o usuário tenha a permissão `lakeformation:GetDataAccess`. Outros serviços integrados exigem que sua função de execução subjacente tenha a permissão `lakeformation:GetDataAccess`.

Essa permissão está incluída nas políticas sugeridas no [Referência de personas e permissões do IAM do Lake Formation](#).

Resumindo, para permitir que as entidades principais do Lake Formation leiam e gravem dados subjacentes com acesso controlado pelas permissões do Lake Formation:

- Registre os locais do Amazon S3 que contêm os dados com o Lake Formation.
- As entidades principais que criam tabelas do Catálogo de dados que apontam para locais de dados subjacentes devem ter permissões de localização de dados.
- As entidades principais que lêem e gravam dados subjacentes devem ter permissões de acesso aos dados do Lake Formation nas tabelas do Catálogo de dados que apontam para os locais de dados subjacentes.

- As entidades principais que lêem e gravam dados subjacentes devem ter a permissão `lakeformation:GetDataAccess` do IAM quando o local dos dados subjacentes é registrado no Lake Formation.

### Note

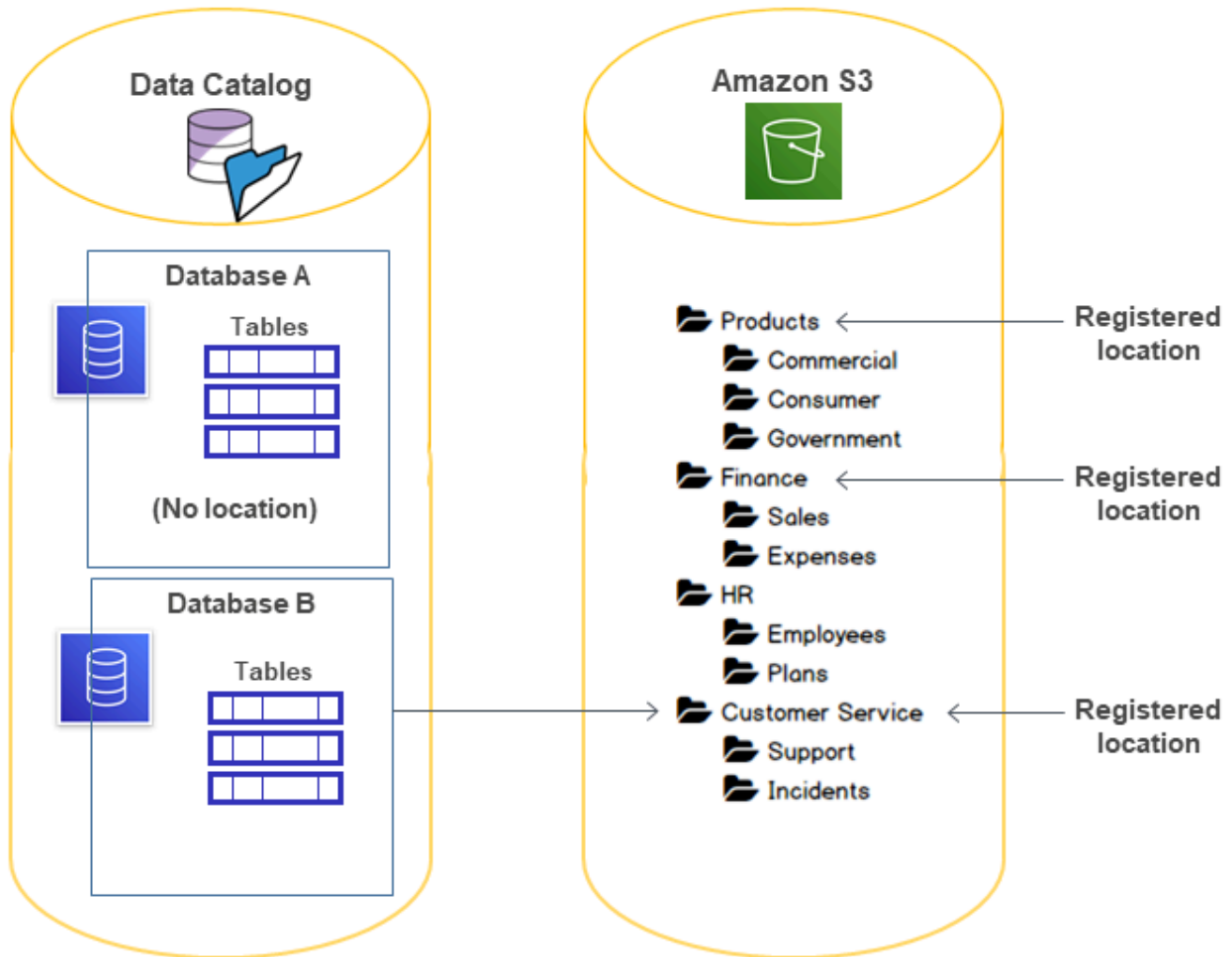
O modelo de permissões do Lake Formation não impede o acesso aos locais do Amazon S3 por meio da API ou do console do Amazon S3 se você tiver acesso a eles por meio de políticas do IAM ou do Amazon S3. Você pode anexar políticas do IAM às entidades principais para bloquear esse acesso.

## Saiba mais sobre permissões de localização de dados

As permissões de localização de dados governam o resultado das operações de criação e atualização nos bancos de dados e tabelas do Catálogo de dados. As regras são as seguintes:

- Uma entidade principal deve ter permissões de localização de dados explícitas ou implícitas em um local do Amazon S3 para criar ou atualizar um banco de dados ou tabela que especifique esse local.
- A permissão explícita `DATA_LOCATION_ACCESS` é concedida usando o console, a API ou AWS CLI.
- As permissões implícitas são concedidas quando um banco de dados tem uma propriedade de localização que aponta para um local registrado, a entidade principal tem a permissão `CREATE_TABLE` no banco de dados e a entidade principal tenta criar uma tabela nesse local ou em um local secundário.
- Se uma entidade principal receber permissões de localização de dados em um local, a entidade principal terá permissões de localização de dados em todos os locais secundários.
- Uma entidade principal não precisa de permissões de localização de dados para realizar operações de leitura/gravação nos dados subjacentes. É suficiente ter as permissões de acesso aos dados `SELECT` ou `INSERT`. As permissões de localização de dados se aplicam somente à criação de recursos do Catálogo de dados que apontam para o local.

Considere o cenário mostrado no diagrama a seguir.



Neste diagrama:

- Os buckets do Amazon S3 Products, Finance e Customer Service estão registrados no Lake Formation.
- Database A não tem propriedade de localização e Database B tem uma propriedade de localização que aponta para o bucket Customer Service.
- O usuário `datalake_user` tem `CREATE_TABLE` nos dois bancos de dados.
- O usuário `datalake_user` recebeu permissões de localização de dados somente no bucket Products.



A seguir estão os resultados quando o usuário `dataLake_user` tenta criar uma tabela de catálogo em um banco de dados específico em um determinado local.

Local onde **dataLake\_user** tenta criar uma tabela

Banco de dados e localização	Sucesso ou falha	Motivo
Banco de dados A em Finance/Sales	Falha	Sem permissão de localização de dados
Banco de dados A em Products	Sucesso	Possui permissão de localização de dados
Banco de dados A em HR/Plans	Sucesso	O local não está registrado
Banco de dados B em Customer Service/Incidents	Sucesso	O banco de dados tem propriedade de localização em Customer Service

Para obter mais informações, consulte as informações a seguir.

- [Adicionar uma localização do Amazon S3 ao seu data lake](#)
- [Referência de permissões do Lake Formation](#)
- [Referência de personas e permissões do IAM do Lake Formation](#)

## Referência de personas e permissões do IAM do Lake Formation

Esta seção lista algumas personas sugeridas de Lake Formation e suas permissões sugeridas (IAM) AWS Identity and Access Management . Para obter informações sobre permissões do Lake Formation, consulte [the section called “Referência de permissões do Lake Formation”](#).

### AWS Lake Formation personas

A tabela a seguir lista as AWS Lake Formation personas sugeridas.

## Personas do Lake Formation

Pessoa	Descrição
IAM administrator (superuser)	(Obrigatório) Usuário que pode criar usuários e perfis do IAM. Tem a política <code>AdministratorAccess</code> AWS gerenciada. Tem todas as permissões em todos os recursos do Lake Formation. Pode adicionar administradores de data lake. Não é possível conceder permissões do Lake Formation se também não for designado como administrador do data lake.
Administrador de data lake	(Obrigatório) Usuário que pode registrar locais do Amazon S3, acessar o catálogo de dados, criar bancos de dados, criar e executar fluxos de trabalho, conceder permissões do Lake Formation a outros usuários e visualizar registros. AWS CloudTrail Tem menos permissões do IAM do que o administrador do IAM, mas o suficiente para administrar o data lake. Não é possível adicionar outros administradores de data lake.
Administrador somente para leitura	(Opcional) Usuário que pode visualizar as entidades principais, os recursos, as permissões e os logs AWS CloudTrail do Catálogo de dados, sem a permissão para fazer atualizações.
Engenheiro de dados	(Opcional) Usuário que pode criar bancos de dados, criar e executar crawlers e fluxos de trabalho e conceder permissões do Lake Formation nas tabelas do Catálogo de dados que os crawlers e fluxos de trabalho criam. Recomendamos tornar todos os engenheiros de dados criadores de bancos de dados. Para ter mais informações, consulte <a href="#">Criação de um banco de dados</a> .
Analista de dados	(Opcional) Usuário que pode executar consultas no data lake usando, por exemplo, Amazon Athena. Tem permissões suficientes apenas para executar consultas.
Papel do fluxo de trabalho	(Obrigatório) Função que executa um fluxo de trabalho em nome de um usuário. Você especifica esse perfil ao criar um fluxo de trabalho a partir de um esquema.

## AWS políticas gerenciadas para Lake Formation

Você pode conceder as permissões AWS Identity and Access Management (IAM) necessárias para trabalhar AWS Lake Formation usando políticas AWS gerenciadas e políticas em linha. As seguintes políticas AWS gerenciadas estão disponíveis para Lake Formation.

### AWS política gerenciada: `AWSLakeFormationDataAdmin`

[AWSLakeFormationDataAdmin](#) política concede acesso administrativo AWS Lake Formation e serviços relacionados, como AWS Glue o gerenciamento de lagos de dados.

Você pode vincular a `AWSLakeFormationDataAdmin` aos seus usuários, grupos e perfis.

#### Detalhes da permissão

- `CloudTrail`— Permite que os diretores visualizem os AWS CloudTrail registros. Isso é necessário para analisar quaisquer erros na configuração do data lake.
- `Glue` – Permite que as entidades principais visualizem, criem e atualizem tabelas de metadados e bancos de dados no Catálogo de dados. Isso inclui operações de API que começam com `Get`, `List`, `Create`, `Update`, `Delete` e `Search`. Isso é necessário para gerenciar os metadados das tabelas do data lake.
- `IAM` – Permite que as entidades principais recuperem informações sobre usuários, perfis e políticas do IAM vinculadas às funções. Isso é necessário para que o administrador de dados revise e liste os usuários e perfis do IAM para conceder permissões ao Lake Formation.
- `Lake Formation` – Concede aos administradores do data lake as permissões necessárias do Lake Formation para gerenciar os data lakes.
- `S3` – Permite que as entidades principais recuperem informações sobre os buckets do Amazon S3 e suas localizações para configurar a localização dos dados para os data lakes.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "lakeformation:*",
        "cloudtrail:DescribeTrails",
        "cloudtrail:LookupEvents",
        "glue:GetDatabase",
```

```

        "glue:GetDatabases",
        "glue:CreateDatabase",
        "glue:UpdateDatabase",
        "glue>DeleteDatabase",
        "glue:GetConnections",
        "glue:SearchTables",
        "glue:GetTable",
        "glue:CreateTable",
        "glue:UpdateTable",
        "glue>DeleteTable",
        "glue:GetTableVersions",
        "glue:GetPartitions",
        "glue:GetTables",
        "glue:GetWorkflow",
        "glue:ListWorkflows",
        "glue:BatchGetWorkflows",
        "glue>DeleteWorkflow",
        "glue:GetWorkflowRuns",
        "glue:StartWorkflowRun",
        "glue:GetWorkflow",
        "s3:ListBucket",
        "s3:GetBucketLocation",
        "s3:ListAllMyBuckets",
        "s3:GetBucketAcl",
        "iam:ListUsers",
        "iam:ListRoles",
        "iam:GetRole",
        "iam:GetRolePolicy"
    ],
    "Resource": "*"
},
{
    "Effect": "Deny",
    "Action": [
        "lakeformation:PutDataLakeSettings"
    ],
    "Resource": "*"
}
]
}

```

**Note**

A política `AWSLakeFormationDataAdmin` não concede todas as permissões necessárias para administradores de data lake. Permissões adicionais são necessárias para criar e executar fluxos de trabalho e registrar locais com a função `AWSServiceRoleForLakeFormationDataAccess` vinculada ao serviço. Para ter mais informações, consulte [Crie um administrador de data lake](#) e [Usar perfis vinculados ao serviço para o Lake Formation](#).

## AWS política gerenciada: `AWSLakeFormationCrossAccountManager`

[`AWSLakeFormationCrossAccountManager`](#) política fornece acesso entre contas a AWS Glue recursos por meio do Lake Formation e concede acesso de leitura a outros serviços necessários, como AWS Organizations AWS RAM e.

Você pode vincular a `AWSLakeFormationCrossAccountManager` aos seus usuários, grupos e perfis.

### Detalhes da permissão

Esta política inclui as seguintes permissões:

- `Glue` – Permite que as entidades principais definam ou excluam a política de recursos do Catálogo de dados para controle de acesso.
- `Organizations` – Permite que as entidades principais recuperem as informações da conta e da unidade organizacional (OU) de uma organização.
- `ram:CreateResourceShare` – Permite que as entidades principais criem um compartilhamento de recursos.
- `ram:UpdateResourceShare` – Permite que as entidades principais modifiquem algumas propriedades do compartilhamento de recursos especificado.
- `ram>DeleteResourceShare` – Permite que as entidades principais excluam o compartilhamento de recursos especificado.
- `ram:AssociateResourceShare` – Permite que as entidades principais adicionem a lista especificada de entidades principais e a lista de recursos a um compartilhamento de recursos.

- `ram:DisassociateResourceShare` – Permite que as entidades principais removam as entidades principais ou recursos especificados da participação no compartilhamento de recursos especificado.
- `ram:GetResourceShares` – Permite que as entidades principais recuperem detalhes sobre os compartilhamentos de recursos que você possui ou que são compartilhados com você.
- `ram:RequestedResourceType` – Permite que as entidades principais recuperem o tipo de recurso (banco de dados, tabela ou catálogo).
- `AssociateResourceSharePermission`— Permite que os diretores adicionem ou substituam a AWS RAM permissão de um tipo de recurso incluído em um compartilhamento de recursos. Você pode ter exatamente uma permissão associada a cada tipo de recurso no compartilhamento de recursos.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "ram:CreateResourceShare"
      ],
      "Resource": "*",
      "Condition": {
        "StringLikeIfExists": {
          "ram:RequestedResourceType": [
            "glue:Table",
            "glue:Database",
            "glue:Catalog"
          ]
        }
      }
    },
    {
      "Effect": "Allow",
      "Action": [
        "ram:UpdateResourceShare",
        "ram>DeleteResourceShare",
        "ram:AssociateResourceShare",
        "ram:DisassociateResourceShare",
        "ram:GetResourceShares"
      ],
    }
  ]
}
```

```

    "Resource": "*",
    "Condition": {
      "StringLike": {
        "ram:ResourceShareName": [
          "LakeFormation*"
        ]
      }
    }
  },
  {
    "Effect": "Allow",
    "Action": [
      "ram:AssociateResourceSharePermission"
    ],
    "Resource": "*",
    "Condition": {
      "StringLike": {
        "ram:PermissionArn": [
          "arn:aws:ram::aws:permission/AWSRAMLFEnabled*"
        ]
      }
    }
  },
  {
    "Effect": "Allow",
    "Action": [
      "glue:PutResourcePolicy",
      "glue>DeleteResourcePolicy",
      "organizations:DescribeOrganization",
      "organizations:DescribeAccount",
      "ram:Get*",
      "ram:List*"
    ],
    "Resource": "*"
  },
  {
    "Effect": "Allow",
    "Action": [
      "organizations:ListRoots",
      "organizations:ListAccountsForParent",
      "organizations:ListOrganizationalUnitsForParent"
    ],
    "Resource": "*"
  }
}

```

```
]
}
```

## AWS política gerenciada: `AWSGlueConsoleFullAccess`

[AWSGlueConsoleFullAccess](#) é uma política que concede acesso total aos recursos do AWS Glue quando uma identidade à qual a política está anexada usa o AWS Management Console. Se você seguir a convenção de nomenclatura para os recursos especificados nesta política, os usuários poderão acessar todos os recursos do console. Essa política geralmente é anexada aos usuários do AWS Glue console.

Além disso, o AWS Glue e o Lake Formation assumem a função de serviço `AWSGlueServiceRole` para permitir o acesso a serviços relacionados, incluindo Amazon Elastic Compute Cloud (Amazon EC2), Amazon Simple Storage Service (Amazon S3) e Amazon CloudWatch.

## AWS managed policy: `LakeFormationDataAccessServiceRolePolicy`

Essa política é anexada a uma função vinculada ao serviço chamada `ServiceRoleForLakeFormationDataAccess` que permite que o serviço execute ações nos recursos conforme sua solicitação. Você não pode anexar essa política às suas identidades do IAM.

Essa política permite que os serviços AWS integrados do Lake Formation, como Amazon Athena e Amazon Redshift, usem a função vinculada ao serviço para descobrir os recursos do Amazon S3.

Para obter mais informações, consulte [Usar perfis vinculados ao serviço para o Lake Formation](#).

### Detalhes de permissões

Essa política inclui a seguinte permissão.

- `s3:ListAllMyBuckets`— Retorna uma lista de todos os buckets pertencentes ao remetente autenticado da solicitação.

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



```

    "s3:ListAllMyBuckets"
  ],
  "Resource": [
    "arn:aws:s3:::*"
  ]
}
]
}

```

## Atualizações do Lake Formation nas políticas AWS gerenciadas

Veja detalhes sobre as atualizações das políticas AWS gerenciadas do Lake Formation desde que esse serviço começou a rastrear essas mudanças.

Alteração	Descrição	Data
Política LakeFormationDataAccessServiceRolePolicy atualizada do Lake Formation.	A Lake Formation aprimorou a <a href="#">LakeFormationDataAccessServiceRolePolicy</a> política adicionando um elemento Sid à declaração de política.	Fevereiro de 2024
Política AWSLakeFormationCrossAccountManager atualizada do Lake Formation.	A Lake Formation aprimorou a <a href="#">AWSLakeFormationCrossAccountManager</a> política adicionando uma nova permissão para permitir o compartilhamento de dados entre contas no modo de acesso híbrido.	Outubro de 2023
Política AWSLakeFormationCrossAccountManager atualizada do Lake Formation.	A Lake Formation aprimorou a <a href="#">AWSLakeFormationCrossAccountManager</a> política para criar apenas um compartilhamento de recursos por conta de destinatário quando o recurso é compartilhado pela primeira vez. Todos os recursos compartilhados posteriormente com a mesma conta são	6 de maio de 2022

Alteração	Descrição	Data
	vinculados ao mesmo compartilhamento de recursos.	
O Lake Formation passou a monitorar as alterações.	A Lake Formation começou a monitorar as mudanças em suas políticas AWS gerenciadas.	6 de maio de 2022

## Permissões sugeridas por personas

A seguir estão as permissões sugeridas para cada persona. O administrador do IAM não está incluído porque esse usuário tem todas as permissões em todos os recursos.

### Tópicos

- [Permissões de administrador do data lake](#)
- [Permissões de administrador somente para leitura](#)
- [Permissões de engenheiro de dados](#)
- [Permissões de analista de dados](#)
- [Permissões da função de fluxo de trabalho](#)

## Permissões de administrador do data lake

### Important

Nas políticas a seguir, <account-id> substitua por um número de AWS conta válido e <workflow\_role> substitua pelo nome de uma função que tenha permissões para executar um fluxo de trabalho, conforme definido em [Permissões da função de fluxo de trabalho](#).

Tipo de política	Política
AWS políticas gerenciadas	<ul style="list-style-type: none"> <li>• AWSLakeFormationDataAdmin</li> <li>• LakeFormationDataAccessServiceRolePolicy (política de função vinculada ao serviço)</li> </ul>

Tipo de política	Política
	<ul style="list-style-type: none"> <li>• AWSGlueConsoleFullAccess (opcional)</li> <li>• CloudWatchLogsReadOnlyAccess (opcional)</li> <li>• AWSLakeFormationCrossAccountManager (opcional)</li> <li>• AmazonAthenaFullAccess (opcional)</li> </ul> <p>Para obter informações sobre as políticas AWS gerenciadas opcionais, consulte <a href="#">the section called “Crie um administrador de data lake”</a>.</p>

Política embutida (para criar a função vinculada ao serviço Lake Formation)

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": "iam:CreateServiceLinkedRole",
      "Resource": "*",
      "Condition": {
        "StringEquals": {
          "iam:AWSServiceName": "lakeformation.amazonaws.com"
        }
      }
    },
    {
      "Effect": "Allow",
      "Action": [
        "iam:PutRolePolicy"
      ],
      "Resource": "arn:aws:iam:: <account-id> :role/aws-service-role/lakeformation.amazonaws.com/AWSServiceRoleForLakeFormationDataAccess"
    }
  ]
}
```

Tipo de política	Política
<p>(Opcional) Política embutida (política de senha para a função de fluxo de trabalho). Isso é necessário somente se o administrador do data lake criar e executar fluxos de trabalho.</p>	<pre> {   "Version": "2012-10-17",   "Statement": [     {       "Sid": "PassRolePermissions",       "Effect": "Allow",       "Action": [         "iam:PassRole"       ],       "Resource": [         "arn:aws:iam:: &lt;account-id&gt; :role/&lt;workflow_role&gt; "       ]     }   ] } </pre>
<p>(Opcional) Política embutida (se sua conta estiver concedendo ou recebendo permissões entre contas do Lake Formation). Essa política serve para aceitar ou rejeitar convites de compartilhamento de AWS RAM recursos e para permitir a concessão de permissões entre contas às organizações. <code>ram:EnableSharingWithAwsOrganization</code> é necessário somente para administradores de data lake na conta de AWS Organizations gerenciamento.</p>	<pre> {   "Version": "2012-10-17",   "Statement": [     {       "Effect": "Allow",       "Action": [         "ram:AcceptResourceShareInvitation",         "ram:RejectResourceShareInvitation",         "ec2:DescribeAvailabilityZones",         "ram:EnableSharingWithAwsOrganization"       ],       "Resource": "*"     }   ] } </pre>

## Permissões de administrador somente para leitura

Tipo de política	Política
Política embutida (básica)	<pre> {   "Version": "2012-10-17",   "Statement": [     {       "Effect": "Allow",       "Action": [         "lakeformation:GetEffectivePermissionsForPath",         "lakeformation:ListPermissions",         "lakeformation:ListDataCellsFilter",         "lakeformation:GetDataCellsFilter",         "lakeformation:SearchDatabasesByLFTags",         "lakeformation:SearchTablesByLFTags",         "lakeformation:GetLFTag",         "lakeformation:ListLFTags",         "lakeformation:GetResourceLFTags",         "lakeformation:ListLakeFormationOptions",         "cloudtrail:DescribeTrails",         "cloudtrail:LookupEvents",         "glue:GetDatabase",         "glue:GetDatabases",         "glue:GetConnections",         "glue:SearchTables",         "glue:GetTable",         "glue:GetTableVersions",         "glue:GetPartitions",         "glue:GetTables",         "glue:GetWorkflow",         "glue:ListWorkflows",         "glue:BatchGetWorkflows",         "glue:GetWorkflowRuns",         "glue:GetWorkflow",         "s3:ListBucket",         "s3:GetBucketLocation",         "s3:ListAllMyBuckets",         "s3:GetBucketAcl",         "iam:ListUsers",       ]     }   ] } </pre>

Tipo de política	Política
	<pre>             "iam:ListRoles",             "iam:GetRole",             "iam:GetRolePolicy"         ],         "Resource": "*"     },     {         "Effect": "Deny",         "Action": [             "lakeformation:PutDataLakeSettings"         ],         "Resource": "*"     } ] } </pre>

## Permissões de engenheiro de dados

### Important

Nas políticas a seguir, <account-id> substitua por um número de AWS conta válido e <workflow\_role> substitua pelo nome da função do fluxo de trabalho.

Tipo de política	Política
AWS política gerenciada	AWSGlueConsoleFullAccess
Política embutida (básica)	<pre> {   "Version": "2012-10-17",   "Statement": [     {       "Effect": "Allow",       "Action": [         "lakeformation:GetDataAccess",         "lakeformation:GrantPermissions",         "lakeformation:RevokePermissions",         "lakeformation:BatchGrantPermissions", </pre>

Tipo de política	Política
	<pre> "lakeformation:BatchRevokePermissions", "lakeformation:ListPermissions", "lakeformation:AddLFTagsToResource", "lakeformation:RemoveLFTagsFromResource",  "lakeformation:GetResourceLFTags", "lakeformation:ListLFTags", "lakeformation:GetLFTag", "lakeformation:SearchTablesByLFTags", "lakeformation:SearchDatabasesByLFTags",  "lakeformation:GetWorkUnits", "lakeformation:GetWorkUnitResults", "lakeformation:StartQueryPlanning", "lakeformation:GetQueryState", "lakeformation:GetQueryStatistics" ], "Resource": "*" } ] } </pre>

Tipo de política	Política
Política embutida (para operações em tabelas controladas, incluindo operações dentro de transações)	<pre>{   "Version": "2012-10-17",   "Statement": [     {       "Effect": "Allow",       "Action": [         "lakeformation:StartTransaction",         "lakeformation:CommitTransaction",         "lakeformation:CancelTransaction",         "lakeformation:ExtendTransaction",         "lakeformation:DescribeTransaction",         "lakeformation:ListTransactions",         "lakeformation:GetTableObjects",         "lakeformation:UpdateTableObjects",         "lakeformation&gt;DeleteObjectsOnCancel"       ],       "Resource": "*"     }   ] }</pre>



Tipo de política	Política
<p>Política embutida (para controle de acesso a metadados usando o método de controle de acesso baseado em tags (LF-TBAC) do Lake Formation)</p>	<pre> {   "Version": "2012-10-17",   "Statement": [     {       "Effect": "Allow",       "Action": [         "lakeformation:AddLFTagsToResource",         "lakeformation:RemoveLFTagsFromResource",         "lakeformation:GetResourceLFTags",         "lakeformation:ListLFTags",         "lakeformation:GetLFTag",         "lakeformation:SearchTablesByLFTags",         "lakeformation:SearchDatabasesByLFTags"       ],       "Resource": "*"     }   ] } </pre>
<p>Política embutida (política de senha para a função de fluxo de trabalho)</p>	<pre> {   "Version": "2012-10-17",   "Statement": [     {       "Sid": "PassRolePermissions",       "Effect": "Allow",       "Action": [         "iam:PassRole"       ],       "Resource": [         "arn:aws:iam:: &lt;account-id&gt; :role/&lt;workflow _role&gt; "       ]     }   ] } </pre>

## Permissões de analista de dados

Tipo de política	Política
AWS política gerenciada	AmazonAthenaFullAccess
Política embutida (básica)	<pre> {   "Version": "2012-10-17",   "Statement": [     {       "Effect": "Allow",       "Action": [         "lakeformation:GetDataAccess",         "glue:GetTable",         "glue:GetTables",         "glue:SearchTables",         "glue:GetDatabase",         "glue:GetDatabases",         "glue:GetPartitions",         "lakeformation:GetResourceLFTags",         "lakeformation:ListLFTags",         "lakeformation:GetLFTag",         "lakeformation:SearchTablesByLFTags",         "lakeformation:SearchDatabasesByLFTags"       ],       "Resource": "*"     }   ] } </pre>
(Opcional) Política embutida (para operações em tabelas controladas, incluindo operações dentro de transações)	<pre> {   "Version": "2012-10-17",   "Statement": [     {       "Effect": "Allow",       "Action": [         "lakeformation:StartTransaction",         "lakeformation:CommitTransaction",         "lakeformation:CancelTransaction",         "lakeformation:ExtendTransaction",         "lakeformation:DescribeTransaction", </pre>

Tipo de política	Política
	<pre> "lakeformation:ListTransactions", "lakeformation:GetTableObjects", "lakeformation:UpdateTableObjects", "lakeformation&gt;DeleteObjectsOnCancel" ], "Resource": "*" } ] } </pre>

## Permissões da função de fluxo de trabalho

Essa função tem as permissões necessárias para executar um fluxo de trabalho. Você especifica uma função com essas permissões ao criar um fluxo de trabalho.

### Important

Nas políticas a seguir, <region> substitua por um identificador de AWS região válido (por exemplo `us-east-1`), por <account-id> um número de AWS conta válido, <workflow\_role> pelo nome da função do fluxo de trabalho e pelo <your-s3-cloudtrail-bucket> caminho do Amazon S3 para seus AWS CloudTrail registros.

Tipo de política	Política
AWS política gerenciada	AWSGlueServiceRole
Política embutida (acesso a dados)	<pre> {   "Version": "2012-10-17",   "Statement": [     {       "Sid": "Lakeformation",       "Effect": "Allow",       "Action": [         "lakeformation:GetDataAccess",         "lakeformation:GrantPermissions"       ],     }   ], } </pre>

Tipo de política	Política
	<pre> "Resource": "*"     }   ] } </pre>
Política embutida (política de senha para a função de fluxo de trabalho)	<pre> {   "Version": "2012-10-17",   "Statement": [     {       "Sid": "PassRolePermissions",       "Effect": "Allow",       "Action": [         "iam:PassRole"       ],       "Resource": [         "arn:aws:iam:: &lt;account-id&gt; :role/&lt;workflow _role&gt; "       ]     }   ] } </pre>
Política embutida (para ingerir dados fora do data lake, por exemplo, AWS CloudTrail registros)	<pre> {   "Version": "2012-10-17",   "Statement": [     {       "Effect": "Allow",       "Action": ["s3:GetObject", "s3:ListBucket"],       "Resource": ["arn:aws:s3::: &lt;your-s3- cloudtrail-bucket&gt; /*"]     }   ] } </pre>

## Alterando as configurações padrão do seu data lake

Para manter a compatibilidade com versões anteriores AWS Glue, AWS Lake Formation tem as seguintes configurações iniciais de segurança:

- A permissão `Super` é concedida ao grupo `IAMAllowedPrincipals` em todos os recursos existentes do Catálogo de dados AWS Glue.
- As configurações “Usar somente o controle de acesso ao IAM” estão habilitadas para novos recursos do Catálogo de dados.

Essas configurações efetivamente fazem com que o acesso aos recursos do catálogo de dados e aos locais do Amazon S3 seja controlado exclusivamente por políticas (IAM) AWS Identity and Access Management. As permissões individuais do Lake Formation não estão em vigor.

O grupo `IAMAllowedPrincipals` inclui todos os usuários e perfis do IAM que possuem permissão para acessar os recursos do seu catálogo de dados por meio de suas políticas do IAM. A permissão `Super` possibilita que uma entidade principal execute todas as operações suportadas do Lake Formation no banco de dados ou na tabela em que ela foi concedida.


Para alterar as configurações de segurança para que o acesso aos recursos do Catálogo de dados (bancos de dados e tabelas) seja gerenciado pelas permissões do Lake Formation, faça o seguinte:

1. Altere as configurações de segurança padrão para novos recursos. Para obter instruções, consulte [Altere o modelo de permissão padrão ou use o modo de acesso híbrido](#).
2. Altere as configurações dos recursos existentes do Catálogo de dados. Para obter instruções, consulte [Atualizar as permissões AWS Glue de dados para o modelo AWS Lake Formation](#).

Alterando as configurações de segurança padrão usando a operação da API do Lake Formation **PutDataLakeSettings**

Você também pode alterar as configurações de segurança padrão usando a operação da API do Lake Formation [putDataLakeSettings](#). Essa ação usa como argumentos um ID de catálogo opcional e uma estrutura [DataLakeSettings](#).

Para impor metadados e controle de acesso aos dados subjacentes pelo Lake Formation em novos bancos de dados e tabelas, codifique a estrutura `DataLakeSettings` da seguinte forma.

 Note

Substitua `<AccountID>` por um ID de conta válido AWS e `<Username>` por um nome de usuário válido do IAM. Você pode especificar mais de um usuário como administrador de data lake.

```
{
  "DataLakeSettings": {
    "DataLakeAdmins": [
      {
        "DataLakePrincipalIdentifier":
"arn:aws:iam::<AccountId>:user/<Username>"
      }
    ],
    "CreateDatabaseDefaultPermissions": [],
    "CreateTableDefaultPermissions": []
  }
}
```

Você também pode codificar a estrutura da seguinte maneira. Omitir o parâmetro `CreateDatabaseDefaultPermissions` ou `CreateTableDefaultPermissions` é equivalente a passar uma lista vazia.

```
{
  "DataLakeSettings": {
    "DataLakeAdmins": [
      {
        "DataLakePrincipalIdentifier":
"arn:aws:iam::<AccountId>:user/<Username>"
      }
    ]
  }
}
```

Essa ação revoga efetivamente todas as permissões do grupo `IAMAllowedPrincipals Lake Formation` em novos bancos de dados e tabelas. Ao criar um banco de dados, você pode substituir essa configuração.

Para aplicar metadados e controle de acesso aos dados subjacentes somente pelo IAM em novos bancos de dados e tabelas, codifique a estrutura `DataLakeSettings` da seguinte forma.

```
{
  "DataLakeSettings": {
    "DataLakeAdmins": [
      {
        "DataLakePrincipalIdentifier":
"arn:aws:iam::<AccountId>:user/<Username>"
      }
    ]
  }
}
```

```
    }
  ],
  "CreateDatabaseDefaultPermissions": [
    {
      "Principal": {
        "DataLakePrincipalIdentifier": "IAM_ALLOWED_PRINCIPALS"
      },
      "Permissions": [
        "ALL"
      ]
    }
  ],
  "CreateTableDefaultPermissions": [
    {
      "Principal": {
        "DataLakePrincipalIdentifier": "IAM_ALLOWED_PRINCIPALS"
      },
      "Permissions": [
        "ALL"
      ]
    }
  ]
}
```

Isso concede ao Lake Formation Super permissão para o grupo IAMAllowedPrincipals em novos bancos de dados e tabelas. Ao criar um banco de dados, você pode substituir essa configuração.

#### Note

Na estrutura `DataLakeSettings` anterior, o único valor permitido para `DataLakePrincipalIdentifier` é `IAM_ALLOWED_PRINCIPALS` e o único valor permitido para `Permissions` é `ALL`.

## Permissões implícitas do Lake Formation

AWS Lake Formation concede as seguintes permissões implícitas aos administradores do data lake, criadores de banco de dados e criadores de tabelas.

## Administradores de data lake

- Tenha acesso `Describe` a todos os recursos do Catálogo de dados, exceto aos recursos compartilhados de outra conta diretamente com uma entidade principal diferente. Esse acesso não pode ser revogado por um administrador.
- Tenha permissões de localização de dados em todos os lugares no data lake.
- Pode conceder ou revogar o acesso a quaisquer recursos no Catálogo de Dados a qualquer entidade principal (inclusive a si mesmo). Esse acesso não pode ser revogado por um administrador.
- Pode criar bancos de dados no Catálogo de dados.
- Pode conceder permissão para criar um banco de dados para outro usuário.

### Note

Os administradores do data lake podem registrar locais do Amazon S3 somente se tiverem permissões do IAM para fazer isso. As políticas de administrador de data lake sugeridas neste guia concedem essas permissões. Além disso, os administradores do data lake não têm permissões implícitas para eliminar bancos de dados ou alterar/eliminar tabelas criadas por outras pessoas. No entanto, eles podem conceder a si mesmos permissões para fazer isso.

Para obter mais informações sobre administradores de data lake, consulte [Crie um administrador de data lake](#).

## Criadores de banco de dados

- Têm todas as permissões de banco de dados nos bancos de dados que eles criam, têm permissões nas tabelas que eles criam no banco de dados e podem conceder a outras entidades principais na mesma permissão de conta AWS para criar tabelas no banco de dados. Um criador de banco de dados que também tenha a política `AWSLakeFormationCrossAccountManager` gerenciada por AWS pode conceder permissões no banco de dados para outras contas AWS ou organizações.

Os administradores do data lake podem usar o console ou a API do Lake Formation para designar criadores de banco de dados.



**Note**

Os criadores de banco de dados não possui permissões implícitas em tabelas que outras pessoas criam no banco de dados.

Para obter mais informações, consulte [Criação de um banco de dados](#).

### Criadores de tabelas

- Tenha todas as permissões nas tabelas que eles criam.
- Podem conceder permissões em todas as tabelas que eles criam para entidades principais na mesma conta AWS.
- Podem conceder permissões em todas as tabelas que eles criam para outras contas AWS ou organizações se tiverem a política `AWSLakeFormationCrossAccountManager` gerenciada por AWS.
- Pode visualizar os bancos de dados que contêm as tabelas que eles criam.

## Referência de permissões do Lake Formation

Para realizar AWS Lake Formation operações, os diretores precisam tanto das permissões do Lake Formation quanto das permissões AWS Identity and Access Management (IAM). Normalmente, você concede permissões do IAM usando políticas de controle de acesso de baixa granularidade, conforme descrito em [the section called “Visão geral das permissões do Lake Formation”](#). Você pode conceder permissões ao Lake Formation usando o console, a API ou o AWS Command Line Interface (AWS CLI).

Para saber como conceder ou revogar permissões do Lake Formation, consulte [the section called “Conceder e revogar permissões do Catálogo de dados”](#) e [the section called “Concedendo permissões de localização de dados”](#).

**Note**

Os exemplos nesta seção mostram como conceder permissões aos principais na mesma conta AWS. Para obter exemplos de concessões entre contas, consulte [the section called “Compartilhamento de dados entre contas”](#).

## Permissões do Lake Formation por tipo de recurso

A seguir estão as permissões válidas do Lake Formation disponíveis para cada tipo de recurso:

Recurso	Permissão
Database	ALL (Super)
	ALTER
	CREATE_TABLE
	DESCRIBE
	DROP
Table	ALL (Super)
	ALTER
	DELETE
	DESCRIBE
	DROP
	INSERT
	SELECT
View	ALL (Super)
	SELECT
	DESCRIBE
	DROP
Data Catalog	CREATE_DATABASE
Amazon S3 location	DATA_LOCATION_ACCESS

Recurso	Permissão	
LF-Tags	DROP	
	ALTER	
LF-Tag values	ASSOCIATE	
	DESCRIBE	
	GrantWithLFTagExpression	
LF-Tag policy - Database	ALL (Super)	
	ALTER	
	CREATE_TABLE	
	DESCRIBE	
	DROP	
LF-Tag policy - Table	ALL (Super)	
	ALTER	
	DESCRIBE	
	DELETE	
	DROP	
	INSERT	
	SELECT	
Resource link - Database or Table	DESCRIBE	
	DROP	

Recurso	Permissão	
Table with data filters	DESCRIBE	
	DROP	
	SELECT	
Table with column filter	SELECT	

## Tópicos

- [Lake Formation concede e revoga comandos AWS CLI](#)
- [Permissões do Lake Formation](#)

## Lake Formation concede e revoga comandos AWS CLI

Cada descrição de permissão nesta seção inclui exemplos de como conceder a permissão usando um AWS CLI comando. A seguir estão as sinopses da Lake Formation e dos comandos. `grant-permissions` `revoke-permissions` AWS CLI

```
grant-permissions
[--catalog-id <value>]
--principal <value>
--resource <value>
--permissions <value>
[--permissions-with-grant-option <value>]
[--cli-input-json <value>]
[--generate-cli-skeleton <value>]
```

```
revoke-permissions
[--catalog-id <value>]
--principal <value>
--resource <value>
--permissions <value>
[--permissions-with-grant-option <value>]
[--cli-input-json <value>]
[--generate-cli-skeleton <value>]
```

Para obter descrições detalhadas desses comandos, consulte [grant-permissions](#) e [revoke-permissions](#) na Referência de comandos AWS CLI . Esta seção fornece informações adicionais sobre a opção `--principal`.

O valor da opção `--principal` é um dos seguintes:

- Nome de recurso da Amazon (ARN) para um usuário ou função AWS Identity and Access Management (IAM)
- ARN para um usuário ou grupo que se autentica por meio de um provedor SAML, como o Microsoft Active Directory Federation Service (AD FS)
- ARN para um QuickSight usuário ou grupo da Amazon
- Para permissões entre contas, uma ID AWS da conta, uma ID da organização ou uma ID da unidade organizacional

Veja a seguir a sintaxe e exemplos para todos os tipos `--principal`.

Entidade principal é um usuário do IAM

Sintaxe:

```
--principal DataLakePrincipalIdentifier=arn:aws:iam::<account-id>:user/<user-name>
```

Exemplo:

```
--principal DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/  
datalake_user1
```

Entidade principal é um perfil do IAM

Sintaxe:

```
--principal DataLakePrincipalIdentifier=arn:aws:iam::<account-id>:role/<role-name>
```

Exemplo:

```
--principal DataLakePrincipalIdentifier=arn:aws:iam::111122223333:role/workflowrole
```

Entidade principal é um usuário que se autentica por meio de um provedor SAML

Sintaxe:

```
--principal DataLakePrincipalIdentifier=arn:aws:iam::<account-id>:saml-provider/<SAMLproviderName>:user/<user-name>
```

Exemplos:

```
--principal DataLakePrincipalIdentifier=arn:aws:iam::111122223333:saml-provider/idp1:user/datalake_user1
```

```
--principal DataLakePrincipalIdentifier=arn:aws:iam::111122223333:saml-provider/AthenaLakeFormation0kta:user/athena-user@example.com
```

Entidade principal é um grupo que se autentica por meio de um provedor SAML

Sintaxe:

```
--principal DataLakePrincipalIdentifier=arn:aws:iam::<account-id>:saml-provider/<SAMLproviderName>:group/<group-name>
```

Exemplos:


```
--principal DataLakePrincipalIdentifier=arn:aws:iam::111122223333:saml-provider/idp1:group/data-scientists
```

```
--principal DataLakePrincipalIdentifier=arn:aws:iam::111122223333:saml-provider/AthenaLakeFormation0kta:group/my-group
```

Principal é usuário da Amazon QuickSight Enterprise Edition

Sintaxe:

```
--principal DataLakePrincipalIdentifier=arn:aws:quicksight:<region>:<account-id>:user/<namespace>/<user-name>
```

 Note

Para *<namespace>*, você deve especificar default.

### Exemplo:

```
--principal DataLakePrincipalIdentifier=arn:aws:quicksight:us-east-1:111122223333:user/default/bi_user1
```

Principal é um grupo da Amazon QuickSight Enterprise Edition

### Sintaxe:

```
--principal DataLakePrincipalIdentifier=arn:aws:quicksight:<region>:<account-id>:group/<namespace>/<group-name>
```

#### Note

Para *<namespace>*, você deve especificar default.

### Exemplo:

```
--principal DataLakePrincipalIdentifier=arn:aws:quicksight:us-east-1:111122223333:group/default/data_scientists
```

Principal é uma AWS conta

### Sintaxe:

```
--principal DataLakePrincipalIdentifier=<account-id>
```

### Exemplo:

```
--principal DataLakePrincipalIdentifier=111122223333
```

Entidade principal é uma organização

### Sintaxe:

```
--principal DataLakePrincipalIdentifier=arn:aws:organizations::<account-id>:organization/<organization-id>
```

### Exemplo:

```
--principal  
DataLakePrincipalIdentifier=arn:aws:organizations::111122223333:organization/o-  
abcdefghijkl
```

Entidade principal é uma unidade organizacional

### Sintaxe:

```
--principal DataLakePrincipalIdentifier=arn:aws:organizations::<account-  
id>:ou/<organization-id>/<organizational-unit-id>
```

### Exemplo:

```
--principal DataLakePrincipalIdentifier=arn:aws:organizations::111122223333:ou/o-  
abcdefghijkl/ou-ab00-cdefghij
```

Principal é um usuário ou grupo de identidade do IAM Identity Center

### Exemplo: Usuário

```
--principal DataLakePrincipalIdentifier=arn:aws:identitystore:::user/<UserID>
```

### Exemplo: Grupo:

```
--principal DataLakePrincipalIdentifier=arn:aws:identitystore:::group/<GroupID>
```

O diretor é um grupo do IAM - **IAMAllowedPrincipals**

O Lake Formation define a Super permissão em todos os bancos de dados e tabelas no Catálogo de Dados para um grupo chamado IAMAllowedPrincipals por padrão. Se essa permissão de grupo existir em um banco de dados ou em uma tabela, todos os diretores da sua conta terão acesso ao recurso por meio das políticas principais do IAM para AWS Glue. Ele fornece compatibilidade com versões anteriores quando você começa a usar as permissões do Lake Formation para proteger os recursos do catálogo de dados que antes eram protegidos pelas políticas do IAM para AWS Glue.

Ao usar o Lake Formation para gerenciar permissões para seus recursos do Catálogo de Dados, você precisa primeiro revogar a IAMAllowedPrincipals permissão dos recursos ou optar pelo



modo de acesso híbrido dos diretores e dos recursos para que as permissões do Lake Formation funcionem.

Exemplo:

```
--principal DataLakePrincipalIdentifier=IAM_Allowed_Principals
```

O diretor é um grupo do IAM - **ALLIAMPrincipals**

Quando você concede permissões para `ALLIAMPrincipals` agrupar em um recurso do Catálogo de Dados, cada diretor da conta tem acesso ao recurso do Catálogo de Dados usando as permissões do Lake Formation e as permissões do IAM.

Exemplo:

```
--principal DataLakePrincipalIdentifier=123456789012:IAMPrincipals
```

## Permissões do Lake Formation

Esta seção contém as permissões disponíveis do Lake Formation que você pode conceder às entidades principais.

### ALTER

Permissão	Concedido neste recurso	O beneficiário também precisa
ALTER	DATABASE	glue:UpdateDatabase
ALTER	TABLE	glue:UpdateTable
ALTER	LF-Tag	lakeformation:UpdateLFTag

Uma entidade principal com essa permissão pode alterar os metadados de um banco de dados ou tabela no Catálogo de dados. Para tabelas, você pode alterar o esquema da coluna e adicionar parâmetros da coluna. Você não pode alterar as colunas nos dados subjacentes para os quais uma tabela de metadados aponta.

Se a propriedade que está sendo alterada for um local registrado do Amazon Simple Storage Service (Amazon S3), a entidade principal deverá ter permissões de localização de dados no novo local.

### Example

O exemplo a seguir concede a ALTER permissão ao usuário `datalake_user1` no banco de dados `retail` na AWS conta 1111-2222-3333.

```
aws lakeformation grant-permissions --principal
  DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1 --
permissions "ALTER" --resource '{ "Database": {"Name":"retail"}}'
```

### Example

O exemplo a seguir concede ALTER ao usuário `datalake_user1` na tabela `inventory` no banco de dados `retail`.

```
aws lakeformation grant-permissions --principal
  DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1
--permissions "ALTER" --resource '{ "Table": {"DatabaseName":"retail",
"Name":"inventory"}}'
```

## CREATE\_DATABASE

Permissão	Concedido neste recurso	O beneficiário também precisa
CREATE_DATABASE	Catálogo de dados	glue:CreateDatabase

Uma entidade principal com essa permissão pode criar um banco de dados de metadados ou um link de recurso no Catálogo de dados. A entidade principal também pode criar tabelas no banco de dados.

### Example

O exemplo a seguir concede CREATE\_DATABASE ao usuário `datalake_user1` na AWS conta 1111-2222-3333.

```
aws lakeformation grant-permissions --principal
  DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1 --
permissions "CREATE_DATABASE" --resource '{ "Catalog": {}}'
```

Quando uma entidade principal cria um banco de dados no Catálogo de dados, nenhuma permissão para os dados subjacentes é concedida. As seguintes permissões adicionais de metadados são concedidas (junto com a capacidade de conceder essas permissões a outras pessoas):

- CREATE\_TABLE no banco de dados
- Banco de dados da ALTER
- Banco de dados da DROP

Ao criar um banco de dados, a entidade principal pode opcionalmente especificar um local do Amazon S3. Dependendo se a entidade principal tem permissões de localização de dados, a permissão CREATE\_DATABASE pode não ser suficiente para criar bancos de dados em todos os casos. É importante ter em mente os três casos a seguir.

Crie um caso de uso de banco de dados	Permissões necessárias
A propriedade do local não é especificada.	CREATE_DATABASE é suficiente.
A propriedade de localização é especificada e a localização não é gerenciada pelo Lake Formation (não está registrada).	CREATE_DATABASE é suficiente.
A propriedade de localização é especificada e a localização é gerenciada pelo Lake Formation (está registrada).	CREATE_DATABASE é obrigatório, além de permissões de localização de dados no local especificado.

## CREATE\_TABLE

Permissão	Concedido neste recurso	O beneficiário também precisa
CREATE_TABLE	DATABASE	glue:CreateTable

Uma entidade principal com essa permissão pode criar uma tabela de metadados ou um link de recurso no Catálogo de dados dentro do banco de dados especificado.

## Example

O exemplo a seguir concede ao usuário `datalake_user1` permissão para criar tabelas no `retail` banco de dados na AWS conta 1111-2222-3333.

```
aws lakeformation grant-permissions --principal
DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1
--permissions "CREATE_TABLE" --resource '{ "Database": {"Name":"retail"} }'
```

Quando uma entidade principal cria uma tabela no Catálogo de dados, todas as permissões do Lake Formation na tabela são concedidas à entidade principal, com a capacidade de conceder essas permissões a outras pessoas.

## Subsídios entre contas

Se uma conta do proprietário do banco de dados conceder `CREATE_TABLE` a uma conta do destinatário e um usuário na conta do destinatário criar com êxito uma tabela no banco de dados da conta do proprietário, as seguintes regras se aplicam:

- O usuário e os administradores do data lake na conta do destinatário têm todas as permissões do Lake Formation disponíveis. Eles podem conceder permissões na tabela a outras entidades principais de suas contas. Eles não podem conceder permissões às entidades principais na conta do proprietário ou em qualquer outra conta.
- Os administradores do data lake na conta do proprietário podem conceder permissões na tabela a outras entidades principais da conta.

## Permissões de localização de dados

Quando você tenta criar uma tabela que aponta para um local do Amazon S3, dependendo se você tem permissões de localização de dados, a permissão `CREATE_TABLE` pode não ser suficiente para criar uma tabela. É importante ter em mente os três casos a seguir.

Crie um caso de uso de tabela	Permissões necessárias
O local especificado não é gerenciado pelo Lake Formation (não está registrado).	<code>CREATE_TABLE</code> é suficiente.

Crie um caso de uso de tabela	Permissões necessárias
O local especificado é gerenciado pelo Lake Formation (está registrado), e o banco de dados que o contém não tem propriedade de localização ou tem uma propriedade de localização que não seja um prefixo do Amazon S3 da localização da tabela.	CREATE_TABLE é obrigatório, além de permissões de localização de dados no local especificado.
O local especificado é gerenciado pelo Lake Formation (está registrado), e o banco de dados que o contém tem uma propriedade de localização que aponta para um local registrado e é um prefixo Amazon S3 da localização da tabela.	CREATE_TABLE é suficiente.

## DATA\_LOCATION\_ACCESS

Permissão	Concedido neste recurso	O beneficiário também precisa
DATA_LOCATION_ACCESS	Local do Amazon S3	(Permissões do Amazon S3 no local, que devem ser especificadas pela função usada para registrar o local.)

Essa é a única permissão de localização de dados. Uma entidade principal com essa permissão pode criar um banco de dados ou tabela de metadados que aponte para a localização especificada do Amazon S3. O local deve ser registrado. Uma entidade principal que tem permissões de localização de dados em um local também tem permissões de localização em locais secundários.

### Example

O exemplo a seguir concede permissões de localização de dados de `s3://products/retail` ao usuário `datalake_user1` na conta AWS 1111-2222-3333.

```
aws lakeformation grant-permissions --principal
  DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1
```

```
--permissions "DATA_LOCATION_ACCESS" --resource '{ "DataLocation":
{"ResourceArn":"arn:aws:s3:::products/retail"} }'
```

DATA\_LOCATION\_ACCESS não é necessário consultar ou atualizar dados subjacentes. Essa permissão se aplica somente à criação de recursos do Catálogo de dados.

Para obter mais informações sobre permissões de localização de dados, consulte [Underlying data access control](#).

## DELETE

Permissão	Concedido neste recurso	O beneficiário também precisa
DELETE	TABLE	(Nenhuma permissão adicional do IAM é necessária se o local estiver registrado.)

Uma entidade principal com essa permissão pode excluir dados subjacentes no local do Amazon S3 especificado pela tabela. A entidade principal também pode visualizar a tabela no console do Lake Formation e recuperar informações sobre a tabela com a API AWS Glue.

### Example

O exemplo a seguir concede a DELETE permissão ao usuário na tabela `dataLake_user1` no banco de dados `inventory retail` na AWS conta `1111-2222-3333`.

```
aws lakeformation grant-permissions --principal
DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1
--permissions "DELETE" --resource '{ "Table": {"DatabaseName":"retail",
"Name":"inventory"} }'
```

Essa permissão se aplica somente aos dados no Amazon S3 e não aos dados de outros armazenamentos de dados, como o Amazon Relational Database Service (Amazon RDS).

## DESCRIBE

Permissão	Concedido neste recurso	O beneficiário também precisa
DESCRIBE	Link de recurso de tabela	<code>glue:GetTable</code>

Permissão	Concedido neste recurso	O beneficiário também precisa
	Link de recurso de banco de dados	<code>glue:GetDatabase</code>
DESCRIBE	DATABASE	<code>glue:GetDatabase</code>
DESCRIBE	TABLE	<code>glue:GetTable</code>
DESCRIBE	LF-Tag	<code>glue:GetTable</code> <code>glue:GetDatabase</code> <code>lakeformation:GetResourceLFTags</code> <code>lakeformation:ListLFTags</code> <code>lakeformation:GetLFTag</code> <code>lakeformation:SearchTablesByLFTags</code> <code>lakeformation:SearchDatabasesByLFTags</code>

Uma entidade principal com essa permissão pode visualizar o banco de dados, a tabela ou o link do recurso especificado. Nenhuma outra permissão do Catálogo de dados é concedida implicitamente e nenhuma permissão de acesso aos dados é concedida implicitamente. Bancos de dados e tabelas aparecem nos editores de consultas dos serviços integrados, mas nenhuma consulta pode ser feita neles, a menos que outras permissões do Lake Formation (por exemplo, `SELECT`) sejam concedidas.

Por exemplo, um usuário que tem `DESCRIBE` em um banco de dados pode ver o banco de dados e todos os metadados do banco de dados (descrição, localização e assim por diante). No entanto, o usuário não consegue descobrir quais tabelas o banco de dados contém e não pode descartar, alterar ou criar tabelas no banco de dados. Da mesma forma, um usuário que tem `DESCRIBE` em

uma tabela pode ver a tabela e os metadados da tabela (descrição, esquema, localização e assim por diante), mas não pode descartar, alterar ou executar consultas na tabela.

A seguir estão algumas regras adicionais para DESCRIBE:

- Se um usuário tiver outras permissões do Lake Formation em um banco de dados, tabela ou link de recurso, DESCRIBE será concedida implicitamente.
- Se um usuário tiver SELECT em apenas um subconjunto de colunas para uma tabela (parcial SELECT), o usuário estará restrito a ver apenas essas colunas.
- Você não pode conceder DESCRIBE a um usuário que tenha seleção parcial em uma tabela. Por outro lado, você não pode especificar listas de inclusão ou exclusão de colunas para tabelas concedidas a DESCRIBE.

## Example

O exemplo a seguir concede a DESCRIBE permissão ao usuário `datalake_user1` no link do recurso de tabela no banco de dados `inventory-link` `retail` na AWS conta `1111-2222-3333`.

```
aws lakeformation grant-permissions --principal
  DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1
  --permissions "DESCRIBE" --resource '{ "Table": {"DatabaseName":"retail",
  "Name":"inventory-link"} }'
```

## DROP

Permissão	Concedido neste recurso	O beneficiário também precisa
DROP	DATABASE	glue:DeleteDatabase
DROP	TABLE	glue:DeleteTable
DROP	LF-Tag	lakeformation:DeleteLFTag
DROP	Link de recurso de banco de dados	glue:DeleteDatabase
	Link de recurso de tabela	glue:DeleteTable



Uma entidade principal com essa permissão pode colocar um link de banco de dados, tabela ou recurso no Catálogo de Dados. Você não pode conceder DROP em um banco de dados a uma conta ou organização externa.

### Warning

Eliminar um banco de dados elimina todas as tabelas no banco de dados.

### Example

O exemplo a seguir concede a DROP permissão ao usuário no banco de dados `datalake_user1 retail` na AWS conta 1111-2222-3333.

```
aws lakeformation grant-permissions --principal
  DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1 --
permissions "DROP" --resource '{ "Database": {"Name":"retail"}}'
```

### Example

O exemplo a seguir concede DROP ao usuário `datalake_user1` na tabela `inventory` do banco de dados `retail`.

```
aws lakeformation grant-permissions --principal
  DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1
--permissions "DROP" --resource '{ "Table": {"DatabaseName":"retail",
"Name":"inventory"}}'
```

### Example

O exemplo a seguir concede DROP ao usuário `datalake_user1` na tabela o link do recurso `inventory-link` no banco de dados `retail`.

```
aws lakeformation grant-permissions --principal
  DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1 --
permissions "DROP" --resource '{ "Table": {"DatabaseName":"retail", "Name":"inventory-
link"}}'
```

## INSERT

Permissão	Concedido neste recurso	O beneficiário também precisa
INSERT	TABLE	(Nenhuma permissão adicional do IAM é necessária se o local estiver registrado.)

Uma entidade principal com essa permissão pode inserir, atualizar e ler dados subjacentes no local do Amazon S3 especificado pela tabela. A entidade principal também pode visualizar a tabela no console do Lake Formation e recuperar informações sobre a tabela com a API AWS Glue.

### Example

O exemplo a seguir concede a INSERT permissão ao usuário na tabela `dataLake_user1` no banco de dados `inventory retail` na AWS conta 1111-2222-3333.

```
aws lakeformation grant-permissions --principal
DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/dataLake_user1
--permissions "INSERT" --resource '{ "Table": {"DatabaseName":"retail",
"Name":"inventory"} }'
```

Essa permissão se aplica somente aos dados no Amazon S3 e não aos dados em outros armazenamentos de dados, como o Amazon RDS.

## SELECT

Permissão	Concedido neste recurso	O beneficiário também precisa
SELECT	<ul style="list-style-type: none"> <li>TABLE</li> </ul>	(Nenhuma permissão adicional do IAM é necessária se o local estiver registrado.)

Uma entidade principal com essa permissão pode visualizar uma tabela no Catálogo de dados e consultar os dados subjacentes no Amazon S3 no local especificado pela tabela. A entidade principal pode visualizar a tabela no console do Lake Formation e recuperar informações sobre a tabela com a API AWS Glue. Se a filtragem de colunas foi aplicada quando essa permissão foi concedida, a

entidade poderá visualizar os metadados somente das colunas incluídas e poderá consultar dados somente das colunas incluídas.

### Note

É responsabilidade do serviço de análise integrada aplicar a filtragem de colunas ao processar uma consulta.

## Example

O exemplo a seguir concede a SELECT permissão ao usuário na tabela `datalake_user1` no banco de dados `inventory retail` na AWS conta 1111-2222-3333.

```
aws lakeformation grant-permissions --principal
  DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1
  --permissions "SELECT" --resource '{ "Table": {"DatabaseName":"retail",
  "Name":"inventory"}}'
```

Essa permissão se aplica somente aos dados no Amazon S3 e não aos dados em outros armazenamentos de dados, como o Amazon RDS.

Você pode filtrar (restringir o acesso a) colunas específicas com uma lista de inclusão opcional ou uma lista de exclusão. Uma lista de inclusão especifica as colunas que podem ser acessadas. Uma lista de exclusão especifica as colunas que não podem ser acessadas. Na ausência de uma lista de inclusão ou exclusão, todas as colunas da tabela estão acessíveis.

Os resultados de `glue:GetTable` retornam somente as colunas que o autor da chamada tem permissão para visualizar. Serviços integrados, como Amazon Athena e Amazon Redshift, honram as listas de inclusão e exclusão de colunas.

## Example

O exemplo a seguir concede SELECT ao usuário `datalake_user1` na tabela `inventory` usando uma lista de inclusão.

```
aws lakeformation grant-permissions --principal
  DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1 --
  permissions "SELECT" --resource '{ "TableWithColumns": {"DatabaseName":"retail",
  "Name":"inventory", "ColumnNames": ["prodcode","location","period","withdrawals"]}}'
```

## Example

O próximo exemplo concede SELECT na tabela `inventory` usando uma lista de exclusão.

```
aws lakeformation grant-permissions --principal
  DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1 --
permissions "SELECT" --resource '{ "TableWithColumns": {"DatabaseName":"retail",
  "Name":"inventory", "ColumnWildcard": {"ExcludedColumnNames": ["intkey",
  "prodcode"]}}}'
```

As seguintes restrições se aplicam à permissão SELECT:

- Ao conceder SELECT, você não poderá incluir a opção de concessão se a filtragem de colunas for aplicada.
- Você não pode restringir o controle de acesso em colunas que são chaves de partição.
- Uma entidade principal com a permissão SELECT em um subconjunto de colunas em uma tabela não pode receber a permissão ALTER, DROP, DELETE ou INSERT nessa tabela. Da mesma forma, uma entidade principal com a permissão ALTER, DROP, DELETE ou INSERT em uma tabela não pode receber a permissão SELECT com a filtragem de colunas.

A permissão SELECT sempre aparece na página Permissões de dados do console do Lake Formation como uma linha separada. A imagem a seguir mostra que SELECT é concedida aos usuários `datalake_user2` e `datalake_user3` em todas as colunas da tabela `inventory`.

Principal	Principal type	Resource type	Resource	Owner account ID	Permissions
<input type="radio"/> datalake_user3	IAM user	Table	inventory	111122223333	Insert
<input type="radio"/> datalake_user3	IAM user	Column	retail.inventory.*	111122223333	Select
<input type="radio"/> datalake_user2	AD user	Table	inventory	111122223333	Delete, Insert
<input type="radio"/> datalake_user2	AD user	Column	retail.inventory.*	111122223333	Select

## Super

Permissão	Concedido neste recurso	O beneficiário também precisa
Super	DATABASE	glue:*Database*

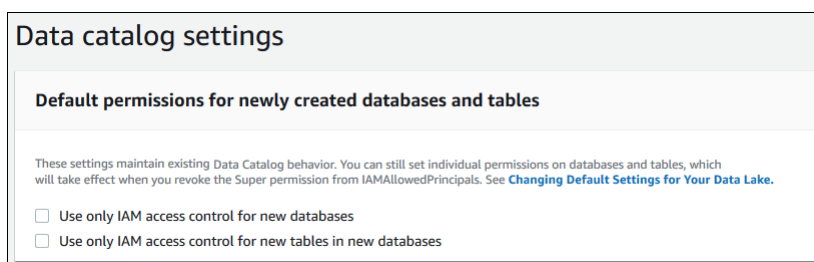
Permissão	Concedido neste recurso	O beneficiário também precisa
Super	TABLE	glue:*Table*, glue:*Partition*

Essa permissão permite que uma entidade principal execute todas as operações suportadas do Lake Formation no banco de dados ou na tabela. Você não pode conceder Super em um banco de dados para uma conta externa.

Essa permissão pode coexistir com as outras permissões do Lake Formation. Por exemplo, você pode conceder as permissões Super, SELECT e INSERT e em uma tabela de metadados. A entidade principal pode então executar todas as operações suportadas na tabela. Quando você revoga Super, as permissões SELECT e INSERT permanecem, e a entidade principal só pode realizar operações de seleção e inserção.

Em vez de conceder Super a uma entidade principal individual, você pode concedê-la ao grupo IAMAllowedPrincipals. O grupo IAMAllowedPrincipals é criado automaticamente e inclui todos os usuários e perfis do IAM que têm acesso permitido aos recursos do seu Catálogo de dados por meio de suas políticas do IAM. Quando Super é concedido a IAMAllowedPrincipals para um recurso do Catálogo de dados, o acesso ao recurso é efetivamente controlado somente pelas políticas do IAM.

Você pode ter a Super permissão de receber IAMAllowedPrincipals automaticamente novos recursos do catálogo aproveitando as opções na página Configurações do console do Lake Formation.



- Para conceder Super a IAMAllowedPrincipals para todos os novos bancos de dados, selecione Usar somente o controle de acesso do IAM para novos bancos de dados.
- Para conceder Super a IAMAllowedPrincipals para todas as novas tabelas em novos bancos de dados, selecione Usar somente o controle de acesso do IAM para novas tabelas em novos bancos de dados.

**Note**

Essa opção faz com que a caixa de seleção Usar somente o controle de acesso do IAM para novas tabelas nesse banco de dados na caixa de diálogo Criar banco de dados seja selecionada por padrão. Não faz nada mais do que isso. É a caixa de seleção na caixa de diálogo Criar banco de dados que permite a concessão de Super a IAMAllowedPrincipals.

Essas opções da página Configurações estão habilitadas por padrão. Para obter mais informações, consulte:

- [the section called “Alterando as configurações padrão do seu data lake”](#)
- [the section called “Atualizar as permissões de dados AWS Glue para o modelo do Lake Formation”](#)

**ASSOCIATE**

Permissão	Concedido neste recurso	O beneficiário também precisa
ASSOCIATE	LF-Tag	glue:GetDatabase glue:GetTable lakeformation:AddLFTagsToResource" lakeformation:RemoveLFTagsFromResource" lakeformation:GetResourceLFTags lakeformation:ListLFTags lakeformation:GetLFTag

Permissão	Concedido neste recurso	O beneficiário também precisa
		lakeformation:SearchTablesByLFTags
		lakeformation:SearchDatabasesByLFTags

Uma entidade principal com essa permissão em uma LF-Tag pode atribuir a LF-Tag a um recurso do Catálogo de dados. Concessão ASSOCIATE de concessões DESCRIBE implicitamente.

### Example

Este exemplo concede ao usuário `dataLake_user1` a permissão ASSOCIATE na LF-Tag com a chave `module`. Ele concede permissões para visualizar e atribuir todos os valores dessa chave, conforme indicado pelo asterisco (\*).

```
aws lakeformation grant-permissions --principal
DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/
dataLake_user1 --permissions "ASSOCIATE" --resource '{ "LFTag":
{"CatalogId":"111122223333","TagKey":"module","TagValues":["*"]}'
```

## Integrar o Centro de Identidade do IAM

Com AWS IAM Identity Center, você pode se conectar a provedores de identidade (IdPs) e gerenciar centralmente o acesso de usuários e grupos em todos os serviços de AWS análise. É possível integrar provedores de identidade, como Okta, Ping e Microsoft Entra ID (antigo Azure Active Directory), ao Centro de Identidade do IAM para que os usuários da organização acessem dados usando uma experiência de login único. O Centro de Identidade do IAM também aceita a conexão de outros provedores de identidade terceiros.

Para obter mais informações, consulte [Provedores de identidade compatíveis](#) no Guia AWS IAM Identity Center do usuário.

Você pode configurar AWS Lake Formation como um aplicativo habilitado no IAM Identity Center, e os administradores do data lake podem conceder permissões refinadas a usuários e grupos autorizados sobre recursos. AWS Glue Data Catalog

Os usuários da organização podem entrar em qualquer aplicação habilitada para o Centro de Identidade usando o provedor de identidade da organização e consultar conjuntos de dados aplicando as permissões do Lake Formation. Com essa integração, você pode gerenciar o acesso aos AWS serviços sem criar várias funções do IAM.

### Note

A propagação confiável de identidade permite que os membros existentes de usuários e grupos acessem dados em todos os AWS serviços de análise. Com a propagação de identidade confiável, um usuário pode entrar em um aplicativo e o aplicativo pode transmitir a identidade do usuário em solicitações para acessar dados em AWS serviços. Você não precisa realizar nenhuma configuração de provedor de identidade específica do serviço ou configuração de função do IAM. Para obter mais informações, consulte [Propagação confiável de identidade entre aplicativos](#) no Guia do AWS IAM Identity Center usuário.

Para conhecer as limitações, consulte [Limitações da integração com o Centro de Identidade do IAM](#).

## Tópicos

- [Pré-requisitos](#)
- [Conectar o Lake Formation ao Centro de Identidade do IAM](#)
- [Atualizar uma integração com o Centro de Identidade do IAM](#)
- [Excluir uma conexão do Lake Formation com o Centro de Identidade do IAM](#)
- [Conceder permissões a usuários e grupos](#)

## Pré-requisitos

Veja a seguir os pré-requisitos para integrar o Centro de Identidade do IAM ao Lake Formation.

1. Habilitar o Centro de Identidade do IAM: habilitar o Centro de Identidade do IAM é um pré-requisito para oferecer compatibilidade com a autenticação e a propagação de identidade.
2. Selecionar a fonte de identidade: depois de habilitar o Centro de Identidade do IAM, é necessário ter um provedor de identificação para gerenciar usuários e grupos. É possível usar o diretório



incorporado do Centro de Identidade como fonte de identidade ou usar IdP externo, como Microsoft Entra ID ou Okta.

Para obter mais informações, consulte [Gerenciar sua fonte de identidade](#) e [Conectar-se a um provedor de identidade externo](#) no Guia AWS IAM Identity Center do usuário.

3. Crie um perfil do IAM: a função que cria a conexão do Centro de Identidade do IAM exige permissões para criar e modificar a configuração da aplicação no Lake Formation e no Centro de Identidade do IAM, conforme a política incorporada a seguir.

É necessário adicionar permissões de acordo com as práticas recomendadas do IAM. As permissões específicas são detalhadas nos procedimentos a seguir. Para obter mais informações, consulte [Getting Started with IAM Identity Center](#).

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "lakeformation:CreateLakeFormationIdentityCenterConfiguration",
        "sso:CreateApplication",
        "sso:PutApplicationAssignmentConfiguration",
        "sso:PutApplicationAuthenticationMethod",
        "sso:PutApplicationGrant",
        "sso:PutApplicationAccessScope"
      ],
      "Resource": [
        "*"
      ]
    }
  ]
}
```

As políticas incorporadas a seguir contêm permissões específicas necessárias para visualizar, atualizar e excluir propriedades da integração do Lake Formation com o Centro de Identidade do IAM.

- Use a política incorporada a seguir para que um perfil do IAM visualize uma integração do Lake Formation ao Centro de Identidade do IAM.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "lakeformation:DescribeLakeFormationIdentityCenterConfiguration",
        "sso:DescribeApplication"
      ],
      "Resource": [
        "*"
      ]
    }
  ]
}
```

- Use a política incorporada a seguir para que um perfil do IAM atualize uma integração do Lake Formation ao Centro de Identidade do IAM.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "lakeformation:UpdateLakeFormationIdentityCenterConfiguration",
        "lakeformation:DescribeLakeFormationIdentityCenterConfiguration",
        "sso:DescribeApplication",
        "sso:UpdateApplication"
      ],
      "Resource": [
        "*"
      ]
    }
  ]
}
```

- Use a política incorporada a seguir para que um perfil do IAM exclua uma integração do Lake Formation ao Centro de Identidade do IAM.

```
{
```

```
"Version": "2012-10-17",
"Statement": [
  {
    "Effect": "Allow",
    "Action": [
      "lakeformation:DeleteLakeFormationIdentityCenterConfiguration",
      "sso:DeleteApplication"
    ],
    "Resource": [
      "*"
    ]
  }
]
```

- Para conhecer as permissões do IAM necessárias para conceder ou revogar permissões de data lake para usuários e grupos do Centro de Identidade do IAM, consulte [Permissões do IAM necessárias para conceder ou revogar as permissões do Lake Formation](#).

### Descrição das permissões

- `lakeformation:CreateLakeFormationIdentityCenterConfiguration`: cria a configuração do Lake Formation IdC.
- `lakeformation:DescribeLakeFormationIdentityCenterConfiguration`: descreve uma configuração existente do IdC.
- `lakeformation>DeleteLakeFormationIdentityCenterConfiguration`: permite excluir uma configuração existente do Lake Formation IdC.
- `lakeformation:UpdateLakeFormationIdentityCenterConfiguration`: usado para alterar uma configuração existente do Lake Formation.
- `sso:CreateApplication`: usado para criar uma aplicação IAM Identity Center.
- `sso>DeleteApplication`: usado para excluir uma aplicação IAM Identity Center.
- `sso:UpdateApplication`: usado para atualizar uma aplicação IAM Identity Center.
- `sso:PutApplicationGrant`: usado para alterar as informações do emissor de tokens confiáveis.
- `sso:PutApplicationAuthenticationMethod`: concede acesso para autenticação no Lake Formation.

- `sso:GetApplicationGrant`: usado para listar as informações do emissor de tokens confiáveis.
- `sso:DeleteApplicationGrant`: exclui as informações do emissor de tokens confiáveis.
- `sso:PutApplicationAccessScope`: adiciona ou atualiza a lista de alvos autorizados para um escopo de acesso ao Centro de Identidade do IAM para uma aplicação.
- `sso:PutApplicationAssignmentConfiguration`: usado para configurar como os usuários obtêm acesso a uma aplicação.

## Conectar o Lake Formation ao Centro de Identidade do IAM

Antes de usar o Centro de Identidade do IAM para gerenciar identidades e conceder acesso aos recursos do catálogo de dados usando o Lake Formation, siga as etapas abaixo. É possível criar a integração do Centro de Identidade do IAM usando o console do Lake Formation ou a AWS CLI.

### AWS Management Console

Como conectar o Lake Formation ao Centro de Identidade do IAM

1. Faça login no AWS Management Console e abra o console do Lake Formation em <https://console.aws.amazon.com/lakeformation/>.
2. No painel de navegação esquerdo, selecione Integração com o Centro de Identidade do IAM.

[AWS Lake Formation](#) > IAM Identity Center integration

## Create IAM Identity Center Integration

Enable IAM Identity Center and then create Lake Formation - IAM Identity Center integration to manage identities from IAM Identity Center (external IdPs like Azure AD or Okta Universal Directory). [Learn more](#)

### ▼ How it works

#### Enable IAM Identity Center

Enable IAM Identity Center for your account or organization and select an identity provider.

✔ IAM Identity Center enabled

#### Create Lake Formation integration

Integrate Lake Formation with IAM Identity Center to permit Lake Formation to access users from your selected identity provider.

#### Grant permissions

Grant permissions to users on Data Catalog databases and tables using fine-grained Lake Formation permissions.

### Connect Lake Formation to IAM Identity Center

#### IAM Identity Center

Manage access to Lake Formation by assigning users and groups from your Identity Center directory.

arn:aws:sso::instance/ssoins-69876430de32a79f

### ▶ Lake Formation application integration - *optional*

Add application IDs that can access S3 data locations registered with Lake Formation on behalf of the user.

ⓘ After this step, you can't edit the connection. You can edit AWS accounts, organizations, and applications. If you want to modify the connection, delete it and create a new connection.

Submit

- (Opcional) Na tela Criar integração com o Lake Formation, especifique os ARNs de aplicações de terceiros que podem acessar dados em locais do Amazon S3 registrados no Lake Formation. A Lake Formation vende credenciais temporárias com escopo reduzido na forma de tokens para locais AWS STS registrados do Amazon S3 com base nas permissões efetivas, para que aplicativos autorizados possam acessar dados em nome dos usuários.

#### 4. Selecione Submit (Enviar).

Depois que o administrador do Lake Formation concluir as etapas e criar a integração, as propriedades do Centro de Identidade do IAM aparecerão no console do Lake Formation. A conclusão dessas tarefas torna o Lake Formation uma aplicação habilitada para o Centro de Identidade do IAM. As propriedades no console incluem o status da integração. O status de integração indica Success quando está concluída. Esse status indica se a configuração do Centro de Identidade do IAM foi concluída.

### AWS CLI

- O exemplo a seguir mostra como criar a integração do Lake Formation ao Centro de Identidade do IAM. Também é possível especificar o Status (ENABLED, DISABLED) das aplicações.

```
aws lakeformation create-lake-formation-identity-center-configuration \  
  --catalog-id <123456789012> \  
  --instance-arn <arn:aws:sso:::instance/ssoins-112111f12ca1122p> \  
  --external-filtering '{"AuthorizedTargets": [<app arn1>", "<app arn2>"],  
  "Status": "ENABLED"}'
```

- O exemplo a seguir mostra como visualizar a integração do Lake Formation ao Centro de Identidade do IAM.

```
aws lakeformation describe-lake-formation-identity-center-configuration  
  --catalog-id <123456789012>
```

## Atualizar uma integração com o Centro de Identidade do IAM

Depois de criar a conexão, é possível adicionar aplicações de terceiros para a integração com o Centro de Identidade do IAM a fim de integrá-las ao Lake Formation e obter acesso aos dados do Amazon S3 em nome dos usuários. Também é possível remover aplicações existentes da integração com o Centro de Identidade do IAM. Você pode adicionar ou remover aplicativos usando o console Lake Formation e usando [UpdateLakeFormationIdentityCenterConfiguration](#) a operação. AWS CLI

**Note**

Depois de criar a integração com o Centro de Identidade do IAM, não é possível atualizar o ARN da instância.

## AWS Management Console

Como atualizar uma conexão existente do Centro de Identidade do IAM com o Lake Formation

1. Faça login no AWS Management Console e abra o console do Lake Formation em <https://console.aws.amazon.com/lakeformation/>.
2. No painel de navegação esquerdo, selecione Integração com o Centro de Identidade do IAM.
3. Selecione Adicionar na página Integração com o Centro de Identidade do IAM.
4. Na tela Adicionar aplicações, insira os IDs das aplicações de terceiros que você deseja integrar ao Lake Formation.
5. Selecione Adicionar.

## AWS CLI

Você pode adicionar ou remover aplicativos de terceiros para a integração do IAM Identity Center executando o AWS CLI comando a seguir. Ao definir o status de filtragem externa como ENABLED, ele permite que o Centro de Identidade do IAM forneça gerenciamento de identidade para aplicações de terceiros acessarem dados gerenciados pelo Lake Formation. Também é possível habilitar ou desabilitar a integração com o Centro de Identidade do IAM definindo o status da aplicação.

```
aws lakeformation update-lake-formation-identity-center-configuration \  
  --external-filtering '{"AuthorizedTargets": ["<app arn1>", "<app arn2>"], "Status":  
  "ENABLED"}' \  
  --application-status ENABLED
```

## Excluir uma conexão do Lake Formation com o Centro de Identidade do IAM

Se quiser excluir uma integração existente do IAM Identity Center, você pode fazer isso usando o console ou a [DeleteLakeFormationIdentityCenterConfiguration](#) operação do Lake Formation. AWS CLI

### AWS Management Console

Como excluir uma conexão existente do Centro de Identidade do IAM com o Lake Formation

1. Faça login no AWS Management Console e abra o console do Lake Formation em <https://console.aws.amazon.com/lakeformation/>.
2. No painel de navegação esquerdo, selecione Integração com o Centro de Identidade do IAM.
3. Selecione Excluir na página Integração com o Centro de Identidade do IAM.
4. Na tela Confirmar integração, confirme a ação e selecione Excluir.

### AWS CLI

Você pode excluir a integração do IAM Identity Center executando o AWS CLI comando a seguir.

```
aws lakeformation delete-lake-formation-identity-center-configuration \  
  --catalog-id <123456789012>
```

## Conceder permissões a usuários e grupos


O administrador do data lake pode conceder permissões a usuários e grupos do Centro de Identidade do IAM nos recursos do catálogo de dados (bancos de dados, tabelas e visualizações) para facilitar o acesso aos dados. Para conceder ou revogar as permissões do data lake, o concesso exige permissões para as ações a seguir do Centro de Identidade do IAM.

- [DescribeUser](#)
- [DescribeGroup](#)
- [DescribeInstance](#)

É possível conceder permissões usando o console do Lake Formation, a API ou a AWS CLI.



Para obter mais informações sobre como conceder permissões, consulte [the section called “Conceder e revogar permissões do Catálogo de dados”](#).

 Note

Só é possível conceder permissões em recursos em sua conta. Para distribuir permissões em cascata para usuários e grupos em recursos compartilhados com você, você deve usar compartilhamentos de AWS RAM recursos.

## AWS Management Console

Como conceder permissões a usuários e grupos

1. Faça login no AWS Management Console e abra o console do Lake Formation em <https://console.aws.amazon.com/lakeformation/>.
2. Selecione Permissões do data lake em Permissões no console do Lake Formation.
3. Selecione Conceder.
4. Na página Conceder permissões de data lake, selecione usuários e grupos do SSM.
5. Selecione Adicionar para escolher os usuários e os grupos aos quais conceder permissões.

[AWS Lake Formation](#) > Grant permissions

## Grant data lake permissions

### Principals

Choose the principals to grant permissions.

<input type="radio"/> <b>IAM users and roles</b> Users or roles from this AWS account.	<input checked="" type="radio"/> <b>IAM Identity Center - new</b> Users and groups configured in IAM Identity Center.	<input type="radio"/> <b>SAML users and groups</b> SAML users and group or QuickSight ARNs.	<input type="radio"/> <b>External accounts</b> AWS account, AWS organization or IAM principal outside of this account
---	--	--	--

### Users and groups (3)

Choose users and groups to grant permissions.

[Remove](#)[Add](#)

&lt;

1

&gt;



<input type="checkbox"/>	Name <a href="#">↗</a>	Type
<input type="checkbox"/>	<a href="#">DataStewards</a>	Group
<input type="checkbox"/>	<a href="#">user1</a>	User
<input type="checkbox"/>	<a href="#">user2</a>	User

- Na tela Atribuir usuários e grupos, selecione os usuários e/ou os grupos aos quais conceder permissões.

Selecione Atribuir.

**Assign users and groups** ✕

🔍 Search by user display name or group name

**Users**

user1 Remove

user2 Remove

**Groups**

DataStewards Remove

[Manage groups](#)

[Learn more about managing groups from IAM Identity Center](#)

Cancel Assign

7. Depois, escolha o método para conceder permissões.

Para obter instruções sobre como conceder permissões usando o método de recursos nomeados, consulte [Conceder permissões de data lake usando o método de recurso nomeado](#).

Para obter instruções sobre como conceder permissão usando tags do LF, consulte [Conceder permissões de data lake usando o método LF-TBAC](#).

8. Selecione os recursos do catálogo de dados nos quais deseja conceder as permissões.
9. Selecione as permissões do catálogo de dados a serem concedidas.
10. Selecione Conceder.

## AWS CLI

O exemplo a seguir mostra como conceder a permissão SELECT ao usuário do Centro de Identidade do IAM em uma tabela.

```
aws lakeformation grant-permissions \  
--principal DataLakePrincipalIdentifier=arn:aws:identitystore:::user/<UserId> \  
--permissions "SELECT" \  
--resource '{ "Table": { "DatabaseName": "retail", "TableWildcard": {} } }'
```

Para recuperar `UserId` do IAM Identity Center, consulte a [GetUserId](#) operação na Referência da API do IAM Identity Center.

## Adicionar uma localização do Amazon S3 ao seu data lake

Para adicionar um local do Amazon Simple Storage Service (Amazon S3) como armazenamento em seu data lake, você registra o local com AWS Lake Formation. Em seguida, você pode usar as permissões do Lake Formation para um controle de acesso refinado aos AWS Glue Data Catalog objetos que apontam para esse local e aos dados subjacentes no local.

O Lake Formation também permite registrar uma localização de dados no modo de acesso híbrido e fornece a flexibilidade de habilitar seletivamente as permissões do Lake Formation para bancos de dados e tabelas em seu Catálogo de dados. Com o modo de acesso híbrido, agora você tem um caminho incremental que permite definir permissões do Lake Formation para um conjunto específico de usuários sem interromper as políticas de permissão de outros usuários ou workloads existentes.

Para obter mais informações sobre como configurar o modo de acesso híbrido, consulte [Modo de acesso híbrido](#)

Quando você registra um local, esse caminho do Amazon S3 e todas as pastas sob esse caminho são registrados.

Por exemplo, digamos que você tenha uma organização de caminhos do Amazon S3 como a seguinte:

```
/mybucket/accounting/sales/
```

Se você se registrar `S3://mybucket/accounting`, a pasta `sales` também será registrada e sob o gerenciamento do Lake Formation.

Para obter mais informações sobre o registro de locais, consulte [Underlying data access control](#).

#### Note

As permissões do Lake Formation são recomendadas para dados estruturados (organizados em tabelas com linhas e colunas). Se os dados contiverem dados não estruturados baseados em objetos, pense em usar a permissão do IAM para o Amazon S3 gerenciar o acesso aos dados.

## Tópicos

- [Requisitos para funções usadas para registrar locais](#)
- [Registrando uma localização do Amazon S3](#)
- [Registrando uma localização criptografada do Amazon S3](#)
- [Registrando uma localização do Amazon S3 em outra conta AWS](#)
- [Registrando uma localização criptografada do Amazon S3 em todas as contas AWS](#)
- [Cancelar o registro de uma localização do Amazon S3](#)

## Requisitos para funções usadas para registrar locais

Você deve especificar uma função AWS Identity and Access Management (IAM) ao registrar uma localização do Amazon Simple Storage Service (Amazon S3). AWS Lake Formation assume essa função ao acessar os dados nesse local.

Você pode usar um dos seguintes tipos de perfil para registrar um local:

- A função vinculada ao serviço do Lake Formation. Esse perfil concede as permissões necessárias no local. O uso desse perfil é a maneira mais simples de registrar o local. Para ter mais informações, consulte [Usar perfis vinculados ao serviço para o Lake Formation](#).
- Um perfil definido pelo usuário. Use um perfil definido pelo usuário quando precisar conceder mais permissões do que o perfil vinculado ao serviço fornece.

Você deve usar um perfil definido pelo usuário nas seguintes circunstâncias:

- Ao registrar um local em outra conta.

Para ter mais informações, consulte [the section called “Registrando uma localização do Amazon S3 em outra conta AWS”](#) e [the section called “Registrando uma localização criptografada do Amazon S3 em todas as contas AWS”](#).

- Se você usou uma CMK AWS gerenciada (aws/s3) para criptografar a localização do Amazon S3.

Para ter mais informações, consulte [Registrando uma localização criptografada do Amazon S3](#).

- Se você planeja acessar o local usando o Amazon EMR.

Se você já registrou um local com o perfil vinculado ao serviço e deseja começar a acessar o local com o Amazon EMR, você deve cancelar o registro do local e registrá-lo novamente com um perfil definido pelo usuário. Para ter mais informações, consulte [the section called “Cancelar o registro de uma localização do Amazon S3”](#).

## Usar perfis vinculados ao serviço para o Lake Formation

AWS Lake Formation usa um perfil vinculado ao serviço (IAM). Um perfil vinculado ao serviço é um tipo exclusivo de perfil IAM vinculado diretamente ao Lake Formation. O perfil vinculado ao serviço é predefinido pelo Lake Formation e inclui todas as permissões que o serviço exige para chamar outros serviços AWS em seu nome.

Um perfil vinculado ao serviço facilita a configuração do Lake Formation porque você não precisa criar um perfil e adicionar manualmente as permissões necessárias. O Lake Formation define as permissões de seu perfil vinculado ao serviço e, a menos que definido de outra forma, apenas o Lake Formation pode assumir suas funções. As permissões definidas incluem a política de confiança e a política de permissões, e essa política de permissões não pode ser anexada a nenhuma outra entidade do IAM.

Este perfil vinculado ao serviço confia nos seguintes serviços para assumir a função:

- `lakeformation.amazonaws.com`

## Permissões de perfil vinculado ao serviço para o Lake Formation

O Lake Formation usa o perfil vinculado ao serviço chamado `AWSServiceRoleForLakeFormationDataAccess`. Essa função fornece um conjunto de permissões do Amazon Simple Storage Service (Amazon S3) que permitem que o serviço integrado

Lake Formation (como Amazon Athena) acesse locais registrados. Ao registrar um local de data lake, você deve fornecer um perfil que tenha as permissões de leitura/gravação necessárias do Amazon S3 nesse local. Em vez de criar um perfil com as permissões necessárias para o Amazon S3, você pode usar esse perfil vinculado ao serviço.

Na primeira vez que você nomeia o perfil vinculado ao serviço como o perfil com o qual registrar um caminho, o perfil vinculado ao serviço e uma nova política do IAM são criadas em seu nome. O Lake Formation adiciona o caminho à política embutida e o anexa ao perfil vinculado ao serviço. Quando você registra caminhos subsequentes com o perfil vinculado ao serviço, o Lake Formation adiciona o caminho à política existente.

Enquanto estiver conectado como administrador do data lake, registre um local do data lake. Em seguida, no console do IAM, pesquise o perfil `AWSServiceRoleForLakeFormationDataAccess` e visualize as políticas anexadas.

Por exemplo, depois de registrar o local `s3://my-kinesis-test/logs`, o Lake Formation cria a seguinte política embutida e a anexa a `AWSServiceRoleForLakeFormationDataAccess`.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "LakeFormationDataAccessPermissionsForS3",
      "Effect": "Allow",
      "Action": [
        "s3:PutObject",
        "s3:GetObject",
        "s3:DeleteObject",
        "s3:AbortMultipartUpload",
        "s3:ListMultipartUploadParts"
      ],
      "Resource": [
        "arn:aws:s3:::my-kinesis-test/logs/*"
      ]
    },
    {
      "Sid": "LakeFormationDataAccessPermissionsForS3ListBucket",
      "Effect": "Allow",
      "Action": [
        "s3:ListBucket",
        "s3:ListBucketMultipartUploads"
      ],
    }
  ]
}
```

```
        "Resource": [
            "arn:aws:s3:::my-kinesis-test"
        ]
    }
]
}
```

As seguintes permissões são necessárias para poder registrar locais com este perfil vinculado ao serviço:

- iam:CreateServiceLinkedRole
- iam:PutRolePolicy

O administrador do data lake normalmente possui essas permissões.

A seguir estão os requisitos para um perfil definido pelo usuário:

- Ao criar o novo perfil, na página Criar perfil do console do IAM, escolha Serviço AWS e, em seguida, em Escolha um caso de uso, escolha Lake Formation.

Se você criar o perfil usando um caminho diferente, certifique-se de que o perfil tenha uma relação de confiança com `lakeformation.amazonaws.com`. Para obter mais informações, consulte [Modificando uma política de confiança de função \(console\)](#).

- A função deve ter relações de confiança com as seguintes entidades:
  - `glue.amazonaws.com`
  - `lakeformation.amazonaws.com`

Para obter mais informações, consulte [Modificando uma política de confiança de função \(console\)](#).

- A função deve ter uma política em linha que conceda ao Amazon S3 permissões de leitura/gravação no local. A seguir está uma política típica.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "s3:PutObject",
        "s3:GetObject",
```



```

        "s3:DeleteObject"
    ],
    "Resource": [
        "arn:aws:s3:::awsexamplebucket/*"
    ]
},
{
    "Effect": "Allow",
    "Action": [
        "s3:ListBucket"
    ],
    "Resource": [
        "arn:aws:s3:::awsexamplebucket"
    ]
}
]
}

```

- O administrador do data lake que registra o local deve ter a permissão `iam:PassRole` sobre o perfil.

A seguir está uma política embutida que concede essa permissão. <account-id>Substitua por um número de AWS conta válido e <role-name>substitua pelo nome da função.

```

{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "PassRolePermissions",
      "Effect": "Allow",
      "Action": [
        "iam:PassRole"
      ],
      "Resource": [
        "arn:aws:iam::<account-id>:role/<role-name>"
      ]
    }
  ]
}

```

- Para permitir que o Lake Formation adicione CloudWatch registros em Logs e publique métricas, adicione a seguinte política em linha.

**Note**

A gravação no CloudWatch Logs incorre em uma cobrança.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "Sid1",
      "Effect": "Allow",
      "Action": [
        "logs:CreateLogStream",
        "logs:CreateLogGroup",
        "logs:PutLogEvents"
      ],
      "Resource": [
        "arn:aws:logs:<region>:<account-id>:log-group:/aws-lakeformation-
acceleration/*",
        "arn:aws:logs:<region>:<account-id>:log-group:/aws-lakeformation-
acceleration/*:log-stream:*"
      ]
    }
  ]
}
```

## Registrando uma localização do Amazon S3

Você deve especificar uma função AWS Identity and Access Management (IAM) ao registrar uma localização do Amazon Simple Storage Service (Amazon S3). O Lake Formation assume essa função quando concede credenciais temporárias a AWS serviços integrados que acessam os dados naquele local.

**Important**

Evite registrar um bucket do Amazon S3 que tenha o Solicitante paga ativado. Para buckets registrados no Lake Formation, a função usada para registrar o bucket é sempre vista como solicitante. Se o bucket for acessado por outra AWS conta, o proprietário do bucket será


cobrado pelo acesso aos dados se a função pertencer à mesma conta do proprietário do bucket.

Você pode usar o AWS Lake Formation console, a API Lake Formation ou AWS Command Line Interface (AWS CLI) para registrar uma localização no Amazon S3.

Antes de começar

Analise os [requisitos da função usada para registrar o local](#).

Para registrar uma localização (console)

 Important

Os procedimentos a seguir pressupõem que a localização do Amazon S3 esteja na mesma AWS conta do Catálogo de Dados e que os dados na localização não estejam criptografados. Outras seções deste capítulo abrangem o registro de várias contas e o registro de locais criptografados.

1. Abra o AWS Lake Formation console em <https://console.aws.amazon.com/lakeformation/>. Faça login como administrador do data lake ou como usuário com a permissão `lakeformation:RegisterResource` do IAM.
2. No painel de navegação, em Registrar e ingerir, selecione Locais do data lake.
3. Escolha Registrar localização e, em seguida, escolha Procurar para selecionar um caminho do Amazon Simple Storage Service (Amazon S3).
4. (Opcional, mas altamente recomendado) Selecione Revisar permissões de local para ver uma lista de todos os recursos existentes no local selecionado do Amazon S3 e as permissões.

Registrar o local selecionado pode fazer com que seus usuários do Lake Formation tenham acesso aos dados que já estão nesse local. A visualização dessa lista ajuda a garantir que os dados existentes permaneçam seguros.

5. Para o perfil do IAM, escolha a função `AWSServiceRoleForLakeFormationDataAccess` vinculada ao serviço (a padrão) ou um perfil personalizado do IAM que atenda aos requisitos em [the section called “Requisitos para funções usadas para registrar locais”](#).

É possível atualizar um local registrado ou outros detalhes somente ao registrá-lo usando um perfil personalizado do IAM. Para editar um local registrado usando um perfil vinculado ao serviço, é necessário cancelar o registro do local e registrá-lo novamente.

- Escolha a opção Ativar Federação do Catálogo de Dados para permitir que o Lake Formation assumira uma função e forneça credenciais temporárias aos AWS serviços integrados para acessar tabelas em bancos de dados federados. Se um local estiver registrado no Lake Formation e você quiser usar o mesmo local para uma tabela em um banco de dados federado, será necessário registrar o mesmo local com a opção Habilitar federação do Catálogo de dados.
- Escolha o Modo de acesso híbrido para não ativar as permissões do Lake Formation por padrão. Ao registrar a localização do Amazon S3 no modo de acesso híbrido, você pode habilitar as permissões do Lake Formation optando pelos principais bancos de dados e tabelas nesse local.

Para obter mais informações sobre como configurar o modo de acesso híbrido, consulte [Modo de acesso híbrido](#).


- Selecione Registrar local.

Para registrar um local (AWS CLI)

- Registrar o novo local no Lake Formation

Este exemplo usa um perfil vinculado ao serviço para registrar o local. Em vez disso, você pode usar o argumento `--role-arn` para fornecer sua própria função.

<s3-path>Substitua por um caminho válido do Amazon S3, o número da conta por uma AWS conta válida e por <s3-access-role>uma função do IAM que tenha permissões para registrar um local de dados.

 Note

Não será possível editar propriedades de um local registrado se ele estiver registrado usando um perfil vinculado ao serviço.

```
aws lakeformation register-resource \  
  --resource-arn arn:aws:s3:::<s3-path> \  
  --use-service-linked-role
```

O exemplo a seguir usa um perfil personalizado para registrar o local.

```
aws lakeformation register-resource \  
  --resource-arn arn:aws:s3:::<s3-path> \  
  --role-arn arn:aws:iam::<123456789012>:role/<s3-access-role>
```

## 2. Como atualizar o local registrado no Lake Formation

Será possível editar um local registrado somente se ele estiver registrado usando um perfil personalizado do IAM. Para um local registrado com um perfil vinculado ao serviço, é necessário cancelar o registro do local e registrá-lo novamente. Para ter mais informações, consulte [the section called “Cancelar o registro de uma localização do Amazon S3”](#).

```
aws lakeformation update-resource \  
  --role-arn arn:aws:iam::<123456789012>:role/<s3-access-role>\  
  --resource-arn arn:aws:s3:::<s3-path>
```

```
aws lakeformation update-resource \  
  --resource-arn arn:aws:s3:::<s3-path> \  
  --use-service-linked-role
```

## 3. Registrar um local de dados no modo de acesso híbrido com federação

```
aws lakeformation register-resource \  
  --resource-arn arn:aws:s3:::<s3-path> \  
  --role-arn arn:aws:iam::<123456789012>:role/<s3-access-role> \  
  --hybrid-access-enabled
```

```
aws lakeformation register-resource \  
  --resource-arn arn:aws:s3:::<s3-path> \  
  --role-arn arn:aws:iam::<123456789012>:role/<s3-access-role> \  
  --with-federation
```

```
aws lakeformation update-resource \  
  --resource-arn arn:aws:s3:::<s3-path> \  
  --role-arn arn:aws:iam::<123456789012>:role/<s3-access-role> \  
  --hybrid-access-enabled
```

Para obter mais informações, consulte Operação [RegisterResource](#) da API.

### Note

Depois de registrar uma localização no Amazon S3, qualquer AWS Glue tabela apontando para a localização (ou qualquer uma de suas localizações secundárias) retornará o valor do `IsRegisteredWithLakeFormation` parâmetro como `true` na `GetTable` chamada. Há uma limitação conhecida de que as operações da API do Catálogo de dados, como `GetTables` e `SearchTables` não atualizem o valor do parâmetro `IsRegisteredWithLakeFormation` e retornem o padrão, que é `false`. É recomendável usar a API `GetTable` para visualizar o valor correto do parâmetro `IsRegisteredWithLakeFormation`.

## Registrando uma localização criptografada do Amazon S3

O Lake Formation se integra com [AWS Key Management Service](#) (AWS KMS) para permitir que você configure com mais facilidade outros serviços integrados para criptografar e descriptografar dados em locais do Amazon Simple Storage Service (Amazon S3).

Tanto o cliente é gerenciado AWS KMS keys Chaves gerenciadas pela AWS quanto o suporte. Atualmente, a criptografia/descriptografia do lado do cliente é suportada somente com o Athena.

Você deve especificar uma função AWS Identity and Access Management (IAM) ao registrar uma localização no Amazon S3. Para locais criptografados do Amazon S3, a função deve ter permissão para criptografar e descriptografar dados com o. Ou a política de chaves do AWS KMS key KMS deve conceder permissões sobre a chave da função.

### Important

Evite registrar um bucket do Amazon S3 que tenha o Solicitante paga ativado. Para buckets registrados no Lake Formation, a função usada para registrar o bucket é sempre vista como solicitante. Se o bucket for acessado por outra AWS conta, o proprietário do bucket será cobrado pelo acesso aos dados se a função pertencer à mesma conta do proprietário do bucket.

A maneira mais simples de registrar a localização é usar a função vinculada ao serviço Lake Formation. Essa função concede as permissões de leitura/gravação necessárias no local. Você

também pode usar uma função personalizada para registrar o local, desde que ele atenda aos requisitos do [the section called “Requisitos para funções usadas para registrar locais”](#).

#### Important

Se você usou um Chave gerenciada pela AWS (aws/s3) para criptografar a localização do Amazon S3, não poderá usar a função vinculada ao serviço Lake Formation. Você deve usar um papel personalizado e adicionar permissões do IAM na chave do perfil. Os detalhes são fornecidos posteriormente nesta seção.

Os procedimentos a seguir explicam como registrar um local do Amazon S3 criptografado com uma chave gerenciada pelo cliente ou um Chave gerenciada pela AWS.

- [Registrar um local criptografado com uma chave gerenciada pelo cliente](#)
- [Registrando um local criptografado com um Chave gerenciada pela AWS](#)

Antes de começar

Analise os [requisitos da função usada para registrar o local](#).


Para registrar uma localização do Amazon S3 criptografada com uma chave gerenciada pelo cliente

#### Note

Se a chave KMS ou a localização do Amazon S3 não estiverem na AWS mesma conta do catálogo de dados, siga as instruções [the section called “Registrando uma localização criptografada do Amazon S3 em todas as contas AWS”](#) em vez disso.

1. Abra o AWS KMS console em <https://console.aws.amazon.com/kms> e faça login como um usuário administrativo AWS Identity and Access Management (IAM) ou como um usuário que pode modificar a política de chaves da chave KMS usada para criptografar o local.
2. No painel de navegação, selecione Chaves gerenciadas pelo cliente e selecione o nome da chave do KMS desejada.
3. Na página de detalhes da chave KMS, escolha a guia Política de chaves e, em seguida, faça o seguinte para adicionar sua função personalizada ou a função vinculada ao serviço Lake Formation como usuário da chave KMS:

- Se a visualização padrão estiver sendo exibida (com as seções Administradores de chaves, Exclusão de chaves, Usuários de chaves e Outras AWS contas), na seção Usuários principais, adicione sua função personalizada ou a função vinculada ao serviço Lake Formation. `AWSServiceRoleForLakeFormationDataAccess`
- Se a política de chaves (JSON) estiver sendo exibida – edite a política para adicionar sua função personalizada ou a função `AWSServiceRoleForLakeFormationDataAccess` vinculada ao serviço Lake Formation ao objeto “Permitir o uso da chave”, conforme mostrado no exemplo a seguir.

 Note

Se esse objeto estiver ausente, adicione-o com as permissões mostradas no exemplo. O exemplo usa a função vinculada ao serviço.

```
...
{
  "Sid": "Allow use of the key",
  "Effect": "Allow",
  "Principal": {
    "AWS": [
      "arn:aws:iam::111122223333:role/aws-service-role/
lakeformation.amazonaws.com/AWSServiceRoleForLakeFormationDataAccess",
      "arn:aws:iam::111122223333:user/keyuser"
    ]
  },
  "Action": [
    "kms:Encrypt",
    "kms:Decrypt",
    "kms:ReEncrypt*",
    "kms:GenerateDataKey*",
    "kms:DescribeKey"
  ],
  "Resource": "*"
},
...
```




4. Abra o AWS Lake Formation console em <https://console.aws.amazon.com/lakeformation/>. Faça login como administrador do data lake ou como usuário com a permissão `lakeformation:RegisterResource` do IAM.
5. No painel de navegação, em Registrar e ingerir, escolha Locais do data lake.
6. Escolha Registrar localização e, em seguida, escolha Procurar para selecionar um caminho do Amazon Simple Storage Service (Amazon S3).
7. (Opcional, mas altamente recomendado) Escolha Revisar permissões de localização para ver uma lista de todos os recursos existentes no local selecionado do Amazon S3 e suas permissões.

Registrar o local selecionado pode fazer com que seus usuários do Lake Formation tenham acesso aos dados que já estão nesse local. A visualização dessa lista ajuda a garantir que os dados existentes permaneçam seguros.

8. Para o perfil do IAM, escolha a função `AWSServiceRoleForLakeFormationDataAccess` vinculada ao serviço (a padrão) ou sua função personalizada que atenda a [the section called “Requisitos para funções usadas para registrar locais”](#).
9. Escolha Registrar local.

Para obter mais informações sobre a função vinculada ao serviço, consulte [Permissões de perfil vinculado ao serviço para o Lake Formation](#).

Para registrar uma localização do Amazon S3 criptografada com um Chave gerenciada pela AWS

 Important

Se a localização do Amazon S3 não estiver na mesma AWS conta do catálogo de dados, siga as instruções em [the section called “Registrando uma localização criptografada do Amazon S3 em todas as contas AWS”](#) vez disso.

1. Crie um perfil do IAM para usar para registrar o local. Certifique-se de que ele atenda aos requisitos listados em [the section called “Requisitos para funções usadas para registrar locais”](#).
2. Adicione a seguinte política em linha à função. Ele concede permissões sobre a chave da função. A especificação `Resource` deve designar o nome do recurso da Amazon (ARN) da Chave gerenciada pela AWS. Você pode obter o ARN no AWS KMS console. Para obter o ARN

correto, certifique-se de fazer login no AWS KMS console com a mesma AWS conta e região Chave gerenciada pela AWS que foram usadas para criptografar o local.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "kms:Encrypt",
        "kms:Decrypt",
        "kms:ReEncrypt*",
        "kms:GenerateDataKey*",
        "kms:DescribeKey"
      ],
      "Resource": "<Chave gerenciada pela AWS ARN>"
    }
  ]
}
```

3. Abra o AWS Lake Formation console em <https://console.aws.amazon.com/lakeformation/>. Faça login como administrador do data lake ou como usuário com a permissão `lakeformation:RegisterResource` do IAM.
4. No painel de navegação, em Registrar e ingerir, escolha Locais do data lake.
5. Escolha Registrar localização e, em seguida, escolha Procurar para selecionar um caminho do Amazon S3.
6. (Opcional, mas altamente recomendado) Escolha Revisar permissões de localização para ver uma lista de todos os recursos existentes no local selecionado do Amazon S3 e suas permissões.

Registrar o local selecionado pode fazer com que seus usuários do Lake Formation tenham acesso aos dados que já estão nesse local. A visualização dessa lista ajuda a garantir que os dados existentes permaneçam seguros.

7. Em Perfil do IAM, escolha a função que você criou na Etapa 1.
8. Escolha Registrar local.

## Registrando uma localização do Amazon S3 em outra conta AWS

AWS Lake Formation permite que você registre localizações do Amazon Simple Storage Service (Amazon S3) em todas as contas. Por exemplo, se AWS Glue Data Catalog estiver na conta A, um usuário na conta A poderá registrar um bucket do Amazon S3 na conta B.

Registrar um bucket do Amazon S3 AWS na conta B usando AWS Identity and Access Management (IAM) na conta A requer as seguintes permissões:

- O papel na conta A deve conceder permissões no bucket na conta B.
- A política de bucket na conta B deve conceder permissões de acesso à função na conta A.

### Important

Evite registrar um bucket do Amazon S3 que tenha o Solicitante paga ativado. Para buckets registrados no Lake Formation, a função usada para registrar o bucket é sempre vista como solicitante. Se o bucket for acessado por outra AWS conta, o proprietário do bucket será cobrado pelo acesso aos dados se a função pertencer à mesma conta do proprietário do bucket.

Você não pode usar a função vinculada ao serviço Lake Formation para registrar um local em outra conta. Em vez disso, é necessário usar uma função definida pelo usuário. A função deve atender aos requisitos do [the section called “Requisitos para funções usadas para registrar locais”](#). Para obter mais informações sobre a função vinculada ao serviço, consulte [Permissões de perfil vinculado ao serviço para o Lake Formation](#).

Antes de começar

Analise os [requisitos da função usada para registrar o local](#).

Para registrar um local em outra AWS conta

### Note

Se o local estiver criptografado, siga as instruções em [the section called “Registrando uma localização criptografada do Amazon S3 em todas as contas AWS”](#) vez disso.

O procedimento a seguir pressupõe que uma entidade principal na conta 1111-2222-3333, que contém o Catálogo de dados, queira registrar o bucket `awsexamplebucket1` do Amazon S3, que está na conta 1234-5678-9012.

1. Na conta 1111-2222-3333, faça login AWS Management Console e abra o console do IAM em <https://console.aws.amazon.com/iam/>
2. Crie uma nova função ou visualize uma função existente que atenda aos requisitos do [the section called “Requisitos para funções usadas para registrar locais”](#). Certifique-se de que a função concede permissões do Amazon S3 em `awsexamplebucket1`.
3. Abra o console do Amazon S3 em <https://console.aws.amazon.com/s3/>. Faça login com a conta 1234-5678-9012.
4. Na lista Nome do bucket, escolha o nome do bucket, `awsexamplebucket1`.
5. Escolha Permissões.
6. Na página Permissões, escolha Política de bucket.
7. No Editor de políticas do bucket, cole a política a seguir. Substitua `<role-name>` pelo nome da sua função.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Principal": {
        "AWS": "arn:aws:iam::111122223333:role/<role-name>"
      },
      "Action": "s3:ListBucket",
      "Resource": "arn:aws:s3:::awsexamplebucket1"
    },
    {
      "Effect": "Allow",
      "Principal": {
        "AWS": "arn:aws:iam::111122223333:role/<role-name>"
      },
      "Action": [
        "s3:DeleteObject",
        "s3:GetObject",
        "s3:PutObject"
      ],
      "Resource": "arn:aws:s3:::awsexamplebucket1/*"
    }
  ]
}
```

```
    }  
  ]  
}
```

8. Escolha Salvar.
9. Abra o AWS Lake Formation console em <https://console.aws.amazon.com/lakeformation/>. Faça login na conta 1111-2222-3333 como administrador do data lake ou como usuário com permissões suficientes para registrar locais.
10. No painel de navegação, em Administração em Locais de data lake.
11. Na página Locais de data lake, selecione Registrar local.
12. Na página Registrar localização, para o caminho do Amazon S3, insira o nome `s3://awsexamplebucket1` do bucket.

#### Note

Você deve digitar o nome do bucket porque os buckets de várias contas não aparecem na lista quando você escolhe Procurar.

13. Para o perfil do IAM, escolha seu perfil.
14. Escolha Registrar local.

## Registrando uma localização criptografada do Amazon S3 em todas as contas AWS

AWS Lake Formation se integra com [AWS Key Management Service](#) (AWS KMS) para permitir que você configure com mais facilidade outros serviços integrados para criptografar e descriptografar dados em locais do Amazon Simple Storage Service (Amazon S3).

Tanto as chaves gerenciadas pelo cliente quanto Chaves gerenciadas pela AWS as são suportadas. A criptografia/descriptografia do lado do cliente não é suportada.

#### Important

Evite registrar um bucket do Amazon S3 que tenha o Solicitante paga ativado. Para buckets registrados no Lake Formation, a função usada para registrar o bucket é sempre vista como solicitante. Se o bucket for acessado por outra AWS conta, o proprietário do bucket será

cobrado pelo acesso aos dados se a função pertencer à mesma conta do proprietário do bucket.

Esta seção explica como registrar uma localização do Amazon S3 nas seguintes circunstâncias:

- Os dados no local do Amazon S3 são criptografados com uma chave KMS criada em AWS KMS.
- A localização do Amazon S3 não está na mesma AWS conta do. AWS Glue Data Catalog
- A chave KMS está ou não na mesma AWS conta do Catálogo de Dados.

O registro de um bucket do Amazon S3 AWS KMS criptografado AWS na conta B usando AWS Identity and Access Management uma função (IAM) AWS na conta A requer as seguintes permissões:

- O papel na conta A deve conceder permissões no bucket na conta B.
- A política de bucket na conta B deve conceder permissões de acesso à função na conta A.
- Se a chave KMS estiver na conta B, a política de chaves deverá conceder acesso à função na conta A, e a função na conta A deverá conceder permissões na chave KMS.

No procedimento a seguir, você cria uma função na AWS conta que contém o Catálogo de Dados (conta A na discussão anterior). Em seguida, você usa essa função para registrar o local. O Lake Formation assume essa função ao acessar dados subjacentes no Amazon S3. A função assumida tem as permissões necessárias na chave do KMS. Como resultado, você não precisa conceder permissões na chave KMS às entidades principais que acessam dados subjacentes com trabalhos de ETL ou com serviços integrados, como Amazon Athena.

#### Important

Você não pode usar a função vinculada ao serviço Lake Formation para registrar um local em outra conta. Em vez disso, é necessário usar uma função definida pelo usuário. A função deve atender aos requisitos do [the section called “Requisitos para funções usadas para registrar locais”](#). Para obter mais informações sobre a função vinculada ao serviço, consulte [Permissões de perfil vinculado ao serviço para o Lake Formation](#).

## Antes de começar

Analise os [requisitos da função usada para registrar o local](#).

Para registrar uma localização criptografada do Amazon S3 em todas as contas AWS

1. Na mesma AWS conta do Catálogo de dados, faça login AWS Management Console e abra o console do IAM em <https://console.aws.amazon.com/iam/>.
2. Crie uma nova função ou visualize uma função existente que atenda aos requisitos de [the section called "Requisitos para funções usadas para registrar locais"](#). Certifique-se de que a função inclua uma política que concede permissões do Amazon S3 no local.
3. Se a chave KMS não estiver na mesma conta do Catálogo de dados, adicione à função uma política em linha que conceda as permissões necessárias na chave KMS. Veja abaixo um exemplo de política . Substitua `<cmk-region>` e `< cmk-account-id >` pela região e pelo número da conta da chave KMS. Substitua `<key-id>` pelo ID da chave.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "kms:Encrypt",
        "kms:Decrypt",
        "kms:ReEncrypt*",
        "kms:GenerateDataKey*",
        "kms:DescribeKey"
      ],
      "Resource": "arn:aws:kms:<cmk-region>:<cmk-account-id>:key/<key-id>"
    }
  ]
}
```

4. No console do Amazon S3, adicione uma política de bucket concedendo as permissões do Amazon S3 necessárias para a função. A seguir há um exemplo de política de bucket. Substitua `< catalog-account-id >` pelo número da AWS conta do Catálogo `<role-name>` de Dados, pelo nome da sua função e `<bucket-name>` pelo nome do bucket.

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

```
    "Principal": {
      "AWS": "arn:aws:iam::<catalog-account-id>:role/<role-name>"
    },
    "Action": "s3:ListBucket",
    "Resource": "arn:aws:s3:::<bucket-name>"
  },
  {
    "Effect": "Allow",
    "Principal": {
      "AWS": "arn:aws:iam::<catalog-account-id>:role/<role-name>"
    },
    "Action": [
      "s3:DeleteObject",
      "s3:GetObject",
      "s3:PutObject"
    ],
    "Resource": "arn:aws:s3:::<bucket-name>/*"
  }
]
```

5. Em AWS KMS, adicione a função como usuário da chave KMS.
  - a. Abra o AWS KMS console em <https://console.aws.amazon.com/kms>. Em seguida, faça login como usuário administrador ou como usuário que pode modificar a política de chaves da chave KMS usada para criptografar o local.
  - b. No painel de navegação, selecione Chaves gerenciadas pelo cliente e selecione o nome da chave do KMS.
  - c. Na página de detalhes da chave KMS, na guia Política de chaves, se a visualização JSON da política de chaves não estiver sendo exibida, escolha Alternar para visualização de política.
  - d. Na seção Política de chaves, escolha Editar e adicione o nome do recurso da Amazon (ARN) da função ao objeto Allow use of the key, conforme mostrado no exemplo a seguir.

 Note


Se esse objeto estiver ausente, adicione-o com as permissões mostradas no exemplo.



```
...
{
  "Sid": "Allow use of the key",
  "Effect": "Allow",
  "Principal": {
    "AWS": [
      "arn:aws:iam::<catalog-account-id>:role/<role-name>"
    ]
  },
  "Action": [
    "kms:Encrypt",
    "kms:Decrypt",
    "kms:ReEncrypt*",
    "kms:GenerateDataKey*",
    "kms:DescribeKey"
  ],
  "Resource": "*"
},
...
```

Para obter mais informações, consulte [Permitir que usuários de outras contas usem uma chave KMS](#) no Guia do desenvolvedor AWS Key Management Service .

6. Abra o AWS Lake Formation console em <https://console.aws.amazon.com/lakeformation/>. Faça login na conta AWS do Catálogo de dados como administrador do data lake.
7. No painel de navegação, em Registrar e ingerir, escolha Locais do data lake.
8. Escolha Registrar local.
9. Na página Registrar localização, para caminho do Amazon S3, insira o caminho da localização como `s3://<bucket>/<prefix>`. Substitua `<bucket>` pelo nome do bucket e `<prefix>` pelo restante do caminho do local.

 Note

Você deve digitar o caminho porque os buckets entre contas não aparecem na lista quando você escolhe Procurar.

10. Para o perfil do IAM, escolha a função na Etapa 2.
11. Escolha Registrar local.

## Cancelar o registro de uma localização do Amazon S3

Você pode cancelar o registro de uma localização do Amazon Simple Storage Service (Amazon S3) se não quiser mais que ela seja gerenciada pelo Lake Formation. O cancelamento do registro de um local não afeta as permissões de localização de dados do Lake Formation concedidas nesse local. Você pode registrar novamente um local que você cancelou e as permissões de localização de dados permanecerão em vigor. Você pode usar uma função diferente para registrar novamente o local.

Para cancelar o registro de um local (console)

1. Abra o AWS Lake Formation console em <https://console.aws.amazon.com/lakeformation/>. Faça login como administrador do data lake ou como usuário com a permissão `lakeformation:RegisterResource` do IAM.
2. No painel de navegação, em Registrar e ingerir, escolha Locais do data lake.
3. Selecione um local e, no menu Ações, escolha Remove.
4. Quando solicitada a confirmação, escolha Remove.

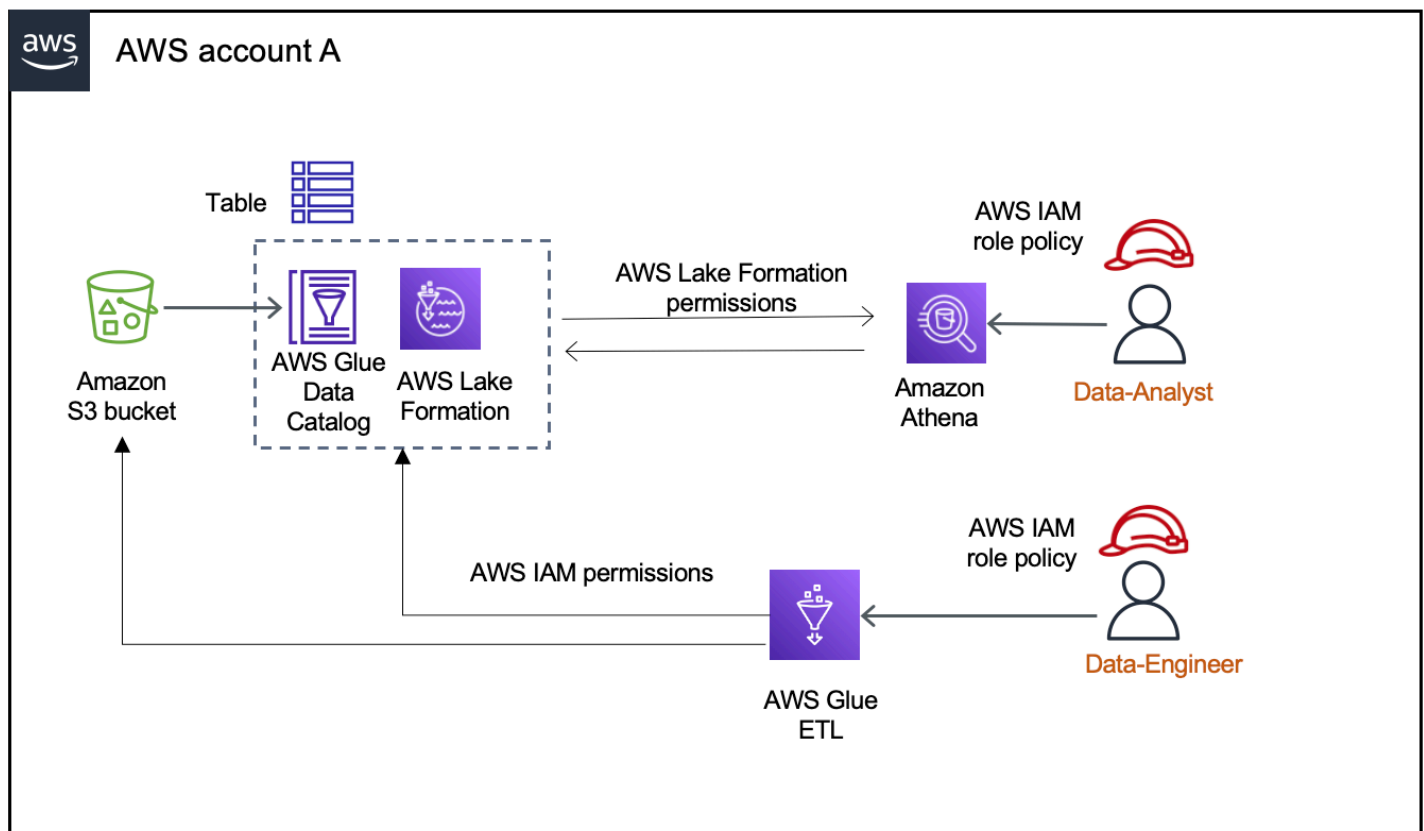
## Modo de acesso híbrido

AWS Lake Formation o modo de acesso híbrido oferece suporte a dois caminhos de permissão para os mesmos AWS Glue Data Catalog bancos de dados e tabelas.

No primeiro caminho, o Lake Formation permite que você selecione diretores específicos e conceda a eles permissões do Lake Formation para acessar bancos de dados e tabelas, optando por participar. O segundo caminho permite que todos os outros diretores acessem esses recursos por meio das políticas principais padrão do IAM para Amazon AWS Glue S3 e ações.

Ao registrar um local do Amazon S3 com o Lake Formation, você tem a opção de aplicar permissões do Lake Formation para todos os recursos desse local ou usar o modo de acesso híbrido. O modo de acesso híbrido impõe somente `CREATE_TABLE`, `CREATE_PARTITION`, `UPDATE_TABLE` permissões por padrão. Quando um local do Amazon S3 está no modo híbrido, você pode habilitar as permissões do Lake Formation optando pelas entidades principais para bancos de dados e tabelas nesse local.

Assim, o modo de acesso híbrido fornece a flexibilidade de habilitar seletivamente o Lake Formation para bancos de dados e tabelas em seu Catálogo de dados para um conjunto específico de usuários sem interromper o acesso de outros usuários ou workloads existentes.



Para ver as considerações e as limitações, consulte [Considerações e limitações do modo de acesso híbrido](#).

## Termos e definições

Aqui estão as definições dos recursos do Catálogo de dados com base em como você configura as permissões de acesso:

### Recurso do Lake Formation

Um recurso registrado no Lake Formation. Os usuários precisam de permissões do Lake Formation para acessar o recurso.

### AWS Glue recurso

Um recurso registrado no Lake Formation. Os usuários precisam apenas de permissões do IAM para acessar o recurso porque ele tem permissões de grupo `IAMAllowedPrincipals`. As permissões do Lake Formation não são aplicadas.

Para obter mais informações sobre as permissões de grupo `IAMAllowedPrincipals`, consulte [Permissões de metadados](#).

## Recurso híbrido

Um recurso registrado no modo de acesso híbrido. Com base nos usuários que acessam o recurso, ele alterna dinamicamente entre ser um recurso do Lake Formation ou um recurso do AWS Glue .

## Casos de uso comuns do modo de acesso híbrido

Você pode usar o modo de acesso híbrido para fornecer acesso em cenários de compartilhamento de dados com uma única conta e entre contas:

### Cenários de conta única

- Converter um AWS Glue recurso em um recurso híbrido — Nesse cenário, você não está usando o Lake Formation no momento, mas deseja adotar as permissões do Lake Formation para bancos de dados e tabelas do Catálogo de Dados. Ao registrar o local do Amazon S3 no modo de acesso híbrido, você pode conceder permissões do Lake Formation aos usuários que optam por bancos de dados e tabelas específicos que apontam para esse local.
- Converter um recurso do Lake Formation em um recurso híbrido — Atualmente, você está usando as permissões do Lake Formation para controlar o acesso a um banco de dados do Data Catalog, mas deseja fornecer acesso a novos diretores usando as permissões do IAM para o Amazon S3 AWS Glue e sem interromper as permissões existentes do Lake Formation.

Quando você atualiza um registro de localização de dados para o modo de acesso híbrido, novas entidades principais podem acessar o banco de dados do Catálogo de dados apontando a localização do Amazon S3 usando políticas de permissões do IAM sem interromper as permissões Lake Formation dos usuários existentes.

Antes de atualizar o registro de localização de dados para ativar o modo de acesso híbrido, você precisa primeiro optar pelas entidades principais que atualmente estão acessando o recurso com as permissões do Lake Formation.

Isso evita possíveis interrupções no fluxo de trabalho atual.

Você também precisa conceder permissão `Super` nas tabelas do banco de dados ao grupo `IAMAllowedPrincipal`.

## Cenários de compartilhamento de dados entre contas

- Compartilhe AWS Glue recursos usando o modo de acesso híbrido — Nesse cenário, a conta do produtor tem tabelas em um banco de dados que atualmente são compartilhadas com uma conta de consumidor usando políticas de permissões do IAM para Amazon S3 e AWS Glue ações. A localização dos dados do banco de dados não está registrada no Lake Formation.

Antes de registrar a localização dos dados no modo de acesso híbrido, você precisa atualizar as Configurações da versão entre contas para a versão 4. A versão 4 fornece as novas políticas de AWS RAM permissão necessárias para o compartilhamento entre contas quando o IAMAllowedPrincipal grupo tem Super permissão sobre o recurso. Para esses recursos com permissões de grupo IAMAllowedPrincipal, você pode conceder permissões do Lake Formation para contas externas e permitir que eles usem as permissões do Lake Formation. O administrador do data lake na conta do destinatário pode conceder permissões do Lake Formation às entidades principais da conta e autorizá-las a aplicar as permissões do Lake Formation.

- Compartilhe recursos do Lake Formation usando o modo de acesso híbrido – Atualmente, a conta de produtor tem tabelas em um banco de dados que são compartilhadas com uma conta de consumidor que impõe as permissões do Lake Formation. A localização dos dados do banco de dados é registrada no Lake Formation.

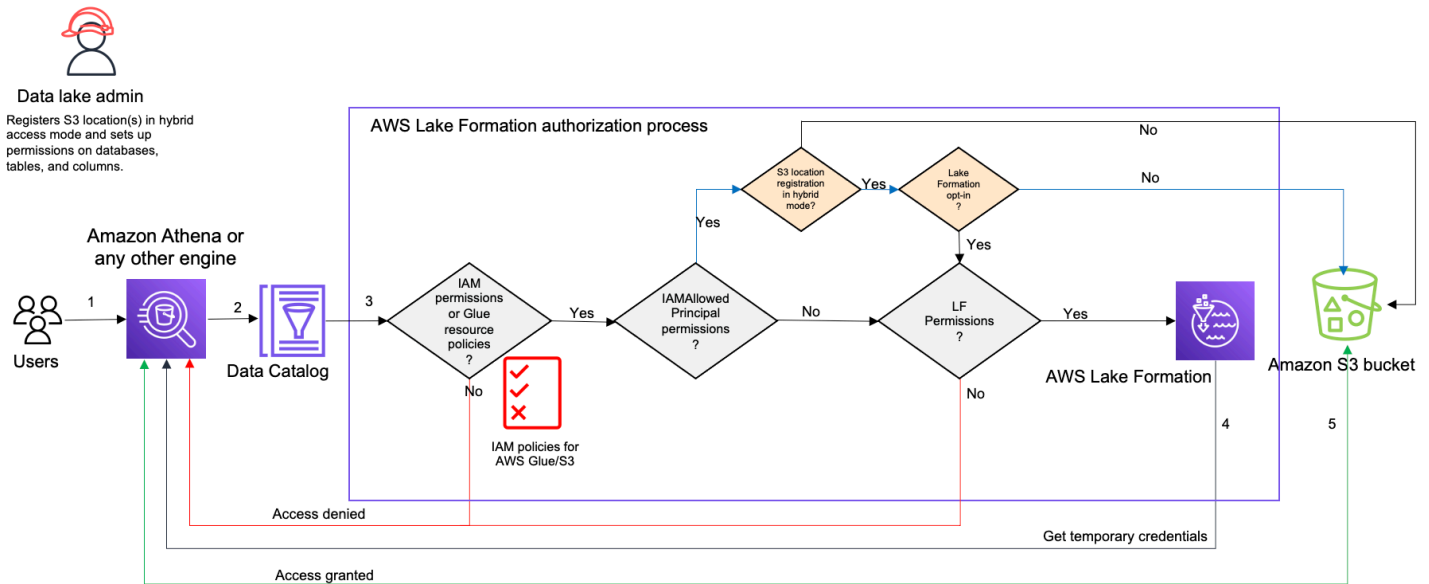
Nesse caso, você pode atualizar o registro de localização do Amazon S3 para o modo de acesso híbrido e compartilhar os dados do Amazon S3 e os metadados do Catálogo de dados usando as políticas de bucket do Amazon S3 e as políticas de recursos do catálogo de dados com as entidades principais na conta do consumidor. Você precisa conceder novamente as permissões existentes do Lake Formation e optar pelas entidades principais antes de atualizar o registro de localização do Amazon S3. Além disso, você precisa conceder permissão Super nas tabelas do banco de dados ao grupo IAMAllowedPrincipals.

## Tópicos

- [Como funciona o modo de acesso híbrido](#)
- [Configurando o modo de acesso híbrido - cenários comuns](#)
- [Removendo entidades principais e recursos do modo de acesso híbrido](#)
- [Visualizando entidades principais e recursos no modo de acesso híbrido](#)
- [Recursos adicionais do](#)

## Como funciona o modo de acesso híbrido

O diagrama a seguir mostra como a autorização do Lake Formation funciona no modo de acesso híbrido quando você consulta os recursos do Catálogo de dados.



Antes de acessar os dados em seu data lake, um administrador de data lake ou um usuário com permissões administrativas configura políticas de usuário individuais da tabela do Catálogo de dados para permitir ou negar o acesso às tabelas em seu Catálogo de dados. Em seguida, uma entidade principal que tem as permissões para realizar a operação `RegisterResource` registra a localização da tabela no Amazon S3 com o Lake Formation no modo de acesso híbrido. O administrador concede permissões do Lake Formation a usuários específicos nos bancos de dados e tabelas do Catálogo de dados e os autoriza a usar as permissões do Lake Formation para esses bancos de dados e tabelas no modo de acesso híbrido.

1. Envia uma consulta - Um principal envia uma consulta ou um script de ETL usando um serviço integrado, como Amazon Athena, Amazon EMR ou Amazon Redshift AWS Glue Spectrum.
2. Solicita dados - O mecanismo analítico integrado identifica a tabela que está sendo solicitada e envia a solicitação de metadados para o Catálogo de dados (`getTable`, `getDatabase`).
3. Verifica as permissões - O Catálogo de dados verifica as permissões de acesso da entidade principal consultora com o Lake Formation.
  - a. Se a tabela não tiver permissões de grupo `IAMAllowedPrincipals` anexadas, as permissões do Lake Formation serão aplicadas.

- b. Se a entidade principal tiver optado por usar as permissões do Lake Formation no modo de acesso híbrido e a tabela tiver permissões do grupo `IAMAllowedPrincipals` anexadas, as permissões do Lake Formation serão aplicadas. O mecanismo de consulta aplica os filtros recebidos do Lake Formation e retorna os dados ao usuário.
  - c. Se a localização da tabela não estiver registrada no Lake Formation e a entidade principal não tiver optado por usar as permissões do Lake Formation no modo de acesso híbrido, o Catálogo de dados verificará se a tabela tem permissões do grupo `IAMAllowedPrincipals` anexadas a ela. Se essa permissão existir na tabela, todas as entidades principais da conta receberão permissões `Super` ou `All` na tabela.
4. Obter credenciais – O Catálogo de dados verifica e informa ao mecanismo se a localização da tabela está registrada no Lake Formation ou não. Se os dados subjacentes estiverem registrados no Lake Formation, o mecanismo analítico solicitará ao Lake Formation credenciais temporárias para acessar os dados no bucket do Amazon S3.
  5. Obter dados – Se a entidade principal estiver autorizada a acessar os dados da tabela, o Lake Formation fornecerá acesso temporário ao mecanismo analítico integrado. Usando o acesso temporário, o mecanismo analítico busca os dados do Amazon S3 e executa a filtragem necessária, como filtragem por coluna, linha ou célula. Quando o mecanismo termina de executar o trabalho, ele retorna os resultados para o usuário. Esse processo chamado de fornecimento de credenciais. Para obter mais informações, consulte [Integração com o Lake Formation](#).
  6.  
Se a localização dos dados da tabela não estiver registrada no Lake Formation, a segunda chamada do mecanismo analítico será feita diretamente para o Amazon S3. A política de bucket do Amazon S3 e a política de usuário do IAM em questão são avaliadas para acesso aos dados. Sempre que você usar as políticas do IAM, siga as práticas recomendadas do IAM. Para obter mais informações, consulte [Práticas recomendadas de segurança no IAM no Guia do usuário do IAM](#).

## Configurando o modo de acesso híbrido - cenários comuns

Assim como nas permissões do Lake Formation, você geralmente tem dois tipos de cenários nos quais pode usar o modo de acesso híbrido para gerenciar o acesso aos dados: fornecer acesso aos principais dentro de um Conta da AWS e fornecer acesso a um externo Conta da AWS ou principal.

Esta seção fornece instruções para configurar o modo de acesso híbrido nos seguintes cenários:

## Gerencie permissões no modo de acesso híbrido em um Conta da AWS

- [Convertendo um AWS Glue recurso em um recurso híbrido](#) — No momento, você está fornecendo acesso às tabelas em um banco de dados para todos os diretores da sua conta usando as permissões do IAM para o Amazon S3 AWS Glue , mas deseja adotar o Lake Formation para gerenciar as permissões de forma incremental.
- [Convertendo um recurso do Lake Formation em um recurso híbrido](#) — No momento, você está usando o Lake Formation para gerenciar o acesso às tabelas em um banco de dados para todos os diretores da sua conta, mas deseja usar o Lake Formation somente para diretores específicos. Você deseja fornecer acesso a novos diretores usando as permissões do IAM para o Amazon S3 no mesmo banco de dados AWS Glue e tabelas.

## Gerencie permissões no modo de acesso híbrido entre Conta da AWS s

- [Compartilhamento de um AWS Glue recurso usando o modo de acesso híbrido](#)— No momento, você não está usando o Lake Formation para gerenciar as permissões de uma tabela, mas deseja aplicar as permissões do Lake Formation para fornecer acesso aos diretores em outra conta.
- [Compartilhando um recurso do Lake Formation usando o modo de acesso híbrido](#)— Você está usando o Lake Formation para gerenciar o acesso a uma tabela, mas deseja fornecer acesso para diretores em outra conta usando permissões do IAM para o Amazon S3 no mesmo banco de dados AWS Glue e tabelas.

## Configurando o modo de acesso híbrido – Etapas de alto nível

1. Registre a localização dos dados do Amazon S3 com o Lake Formation selecionando o Modo de acesso híbrido.
2. As entidades principais devem ter permissão `DATA_LOCATION` em um local de data lake para criar tabelas ou bancos de dados do Catálogo de dados que apontem para esse local.
3. Defina a Configuração da versão entre contas para a versão 4.
4. Conceda permissões refinadas a usuários ou perfis específicos do IAM em bancos de dados e tabelas. Ao mesmo tempo, certifique-se de definir permissões `Super` ou `All` para o grupo `IAMAllowedPrincipals` no banco de dados e para todas ou para as tabelas selecionadas no banco de dados.



5. Opte pelas entidades principais e recursos. Outros diretores da conta podem continuar acessando os bancos de dados e tabelas usando as políticas de permissão do IAM AWS Glue e as ações do Amazon S3.
6. Opcionalmente, limpe as políticas de permissão do IAM para o Amazon S3 para as entidades principais que optaram por usar as permissões do Lake Formation.

## Pré-requisitos para configurar o modo de acesso híbrido

A seguir estão os pré-requisitos para configurar o modo de acesso híbrido:

### Note

Recomendamos que um administrador do Lake Formation registre a localização do Amazon S3 no modo de acesso híbrido e opte por entidades principais e recursos.

1. Conceda permissão de localização de dados (DATA\_LOCATION\_ACCESS) para criar recursos do Catálogo de dados que apontem para os locais do Amazon S3. As permissões de localização de dados controlam a capacidade de criar bancos de dados e tabelas do Catálogo de dados que apontam para os locais específicos do Amazon S3.
2. Para compartilhar recursos do Catálogo de dados com outra conta no modo de acesso híbrido (sem remover as permissões do grupo IAMAllowedPrincipals do recurso), você precisa atualizar as Configurações da versão entre contas para a versão 4. Para atualizar a versão usando o console do Lake Formation, escolha Versão 4 em Configurações de versão entre contas na página Configurações do Catálogo de dados.

Você também pode usar o `put-data-lake-settings` AWS CLI comando para definir o `CROSS_ACCOUNT_VERSION` parâmetro para a versão 4:

```
aws lakeformation put-data-lake-settings --region us-east-1 --data-lake-settings
file://settings
{
  "DataLakeAdmins": [
    {
      "DataLakePrincipalIdentifier": "arn:aws:iam::<111122223333>:user/<user-name>"
    }
  ],
  "CreateDatabaseDefaultPermissions": [],
```

```
"CreateTableDefaultPermissions": [],  
  "Parameters": {  
    "CROSS_ACCOUNT_VERSION": "4"  
  }  
}
```

3.

Para conceder permissões entre contas no modo de acesso híbrido, o concedente deve ter as permissões necessárias do IAM e os AWS Glue serviços. A política AWS gerenciada `AWSLakeFormationCrossAccountManager` concede as permissões necessárias.

Para permitir o compartilhamento de dados entre contas no modo de acesso híbrido, atualizamos a política gerenciada `AWSLakeFormationCrossAccountManager` adicionando duas novas permissões do IAM:

- RAM: `ListResourceSharePermissions`
- RAM: `AssociateResourceSharePermission`

#### Note

Se você não estiver usando a política AWS gerenciada para a função de concedente, adicione as políticas acima às suas políticas personalizadas.

## Convertendo um AWS Glue recurso em um recurso híbrido

Siga estas etapas para registrar uma localização do Amazon S3 no modo de acesso híbrido e integrar novos usuários do Lake Formation sem interromper o acesso aos dados dos usuários existentes do Catálogo de dados.

Descrição do cenário - O local dos dados não está registrado no Lake Formation, e o acesso dos usuários ao banco de dados e às tabelas do catálogo de dados é determinado pelas políticas de permissões do IAM para o Amazon S3 e ações AWS Glue .

Por padrão, o grupo `IAMAllowedPrincipals` tem permissões `Super` em todas as tabelas do banco de dados.

## Para ativar o modo de acesso híbrido para um local de dados que não está registrado no Lake Formation

1. Registre um local do Amazon S3 para ativar o modo de acesso híbrido.

### Console

1. Faça login no [console do Lake Formation](#) como administrador do data lake.
2. No painel de navegação, escolha Localizações do data lake em Administração.
3. Escolha Registrar localizações.

## Register location

### Amazon S3 location

Register an Amazon S3 path as the storage location for your data lake.

#### Amazon S3 path

Choose an Amazon S3 path for your data lake.

*e.g.: s3://bucket/prefix/*

Browse

#### Review location permissions - strongly recommended

Registering the selected location may result in your users gaining access to data already at that location. Before registering a location, we recommend that you review existing location permissions on resources in that location.

Review location permissions

#### IAM role

To add or update data, Lake Formation needs read/write access to the chosen Amazon S3 path. Choose a role that you know has permission to do this, or choose the **AWSServiceRoleForLakeFormationDataAccess** service-linked role. When you register the first Amazon S3 path, the service-linked role and a new inline policy are created on your behalf. Lake Formation adds the first path to the inline policy and attaches it to the service-linked role. When you register subsequent paths, Lake Formation adds the path to the existing policy.

AWSServiceRoleForLakeFormationDataAccess

 Do not select the service linked role if you plan to use EMR.


Enable Data Catalog Federation

Checking this box will allow Lake Formation to assume a role to access tables in a federated database.

### Permission mode

Select the permission mode you want to use to manage access.

Hybrid access mode - *new*

Lake Formation permissions can co-exist with IAM permission policies for AWS Glue and S3 actions to manage access. [Learn more](#) 

Lake Formation

Only Lake Formation permissions are enforced.

Cancel

Register location

4. Na janela Registrar localização, escolha o caminho do Amazon S3 que você deseja registrar no Lake Formation.
5. Para a perfil do IAM, escolha o perfil vinculada ao serviço **AWSServiceRoleForLakeFormationDataAccess** (o padrão) ou um perfil do IAM

personalizado que atenda aos requisitos em [Requisitos para funções usadas para registrar locais](#).

- Escolha o Modo de acesso híbrido para aplicar políticas refinadas de controle de acesso do Lake Formation às entidades principais e bancos de dados e tabelas do Catálogo de dados que apontam para a localização registrada.

Escolha Lake Formation para permitir que o Lake Formation autorize solicitações de acesso ao local registrado.

- Escolha Registrar local.

## AWS CLI

Veja a seguir um exemplo para registrar uma localização de dados no Lake Formation HybridAccessEnabled com:true/false. O valor padrão do parâmetro HybridAccessEnabled é falso. Substitua o caminho, o nome da função e o ID da AWS conta do Amazon S3 por valores válidos.

```
aws lakeformation register-resource --cli-input-json file:file path
json:
  {
    "ResourceArn": "arn:aws:s3:::s3-path",
    "UseServiceLinkedRole": false,
    "RoleArn": "arn:aws:iam::<123456789012>:role/<role-name>",
    "HybridAccessEnabled": true
  }
```

- Conceda permissões e opte pelas entidades principais para usar as permissões do Lake Formation para recursos no modo de acesso híbrido

Antes de optar por entidades principais e recursos no modo de acesso híbrido, verifique se existem concessões Super ou permissões All para o grupo IAMAllowedPrincipals nos bancos de dados e tabelas que têm localização registrada no Lake Formation no modo de acesso híbrido.

### Note

Você não pode conceder ao grupo IAMAllowedPrincipals permissão para All tables em um banco de dados. Você precisa selecionar cada tabela separadamente no menu suspenso e conceder permissões. Além disso, ao criar novas tabelas no banco

de dados, você pode optar por usá-las Use only IAM access control for new tables in new databases nas Configurações do Catálogo de Dados. Essa opção concede permissão Super ao grupo IAMAllowedPrincipals automaticamente quando você cria novas tabelas no banco de dados.

## Console

1. No console do Lake Formation, em Catálogo de dados, escolha Bancos de dados ou Tabelas.
2. Selecione um banco de dados ou uma tabela na lista e escolha Conceder no menu Ações.
3. Escolha entidades principais para conceder permissões no banco de dados, tabelas e colunas usando o método de recurso nomeado ou LF-Tags.

Como alternativa, escolha Permissões do data lake, selecione as entidades principais para conceder permissões na lista e escolha Conceder.

Para obter mais detalhes sobre a concessão de permissões de dados, consulte [Concedendo e revogando permissões nos recursos do Catálogo de dados](#).

### Note

Se você estiver concedendo a permissão Criar tabela a uma entidade principal, também precisará conceder permissões de localização de dados (DATA\_LOCATION\_ACCESS) à entidade principal. Essa permissão não é necessária para atualizar tabelas.

Para ter mais informações, consulte [Concedendo permissões de localização de dados](#).


4. Quando você usa o Método de recurso nomeado para conceder permissões, a opção de optar por entidades principais e recursos está disponível na seção inferior da página Conceder permissão de dados.

Escolha Tornar as permissões do Lake Formation efetivas imediatamente para habilitar as permissões do Lake Formation para as entidades principais e recursos.

### Hybrid access mode - *new*

In hybrid access mode, Lake Formation and IAM policies for AWS Glue and S3 work together.

**Make Lake Formation permissions effective immediately**  
 Lake Formation permissions are enforced for databases, tables, and principals.

 **You might get access denied.**  
 If the checkbox is selected, your Lake Formation permissions are enforced. Make sure that you've completed the required setup for Lake Formation permissions to work. If the checkbox is clear, you can go to [hybrid access mode](#) to add resources and principals. [Learn more](#)

Cancel
Grant

## 5. Selecione Conceder.

Quando você opta pela entidade principal A na tabela A que está apontando para um local de dados, ela permite que a entidade principal A tenha acesso ao local dessa tabela usando as permissões do Lake Formation se o local dos dados estiver registrado no modo híbrido.

## AWS CLI

A seguir está um exemplo de como optar por uma entidade principal e uma tabela no modo de acesso híbrido. Substitua o nome da função, o ID da conta AWS, o nome do banco de dados e o nome da tabela por valores válidos.

```
aws lakeformation create-lake-formation-opt-in --cli-input-json file://file path
json:
{
  "Principal": {
    "DataLakePrincipalIdentifier":
    "arn:aws:iam::<123456789012>:role/<hybrid-access-role>"
  },
  "Resource": {
    "Table": {
      "CatalogId": "<123456789012>",
      "DatabaseName": "<hybrid_test>",
      "Name": "<hybrid_test_table>"
    }
  }
}
```

```
}  
}
```

- a. (Optional) Se você escolher LF-Tags para conceder permissões, você pode optar por entidades principais para usar as permissões do Lake Formation em uma etapa separada. Você pode fazer isso escolhendo o Modo de acesso híbrido em Permissões na barra de navegação esquerda.
- b. Na seção inferior da página do Modo de acesso híbrido, escolha Adicionar para adicionar recursos e entidades principais ao modo de acesso híbrido.
- c. Na página Adicionar recursos e entidades principais, escolha os bancos de dados e tabelas registrados no modo de acesso híbrido. Escolha entidades para optar por usar as permissões do Lake Formation no modo de acesso híbrido.

Você pode escolher `All tables` em um banco de dados para conceder acesso.



# Add resources and principals

Choose databases, tables, and principals to add in hybrid access mode. Lake Formation permissions will be enforced.

[Learn more](#)

## Resources

### Databases

Select one or more databases.

Choose databases ▼

Load more

test X

### Tables - optional

Select one or more tables.

Choose tables ▼

All tables X

## Principals

### IAM users and roles

Add one or more IAM users or roles.

Choose IAM principals to add ▼

datalake\_user X  
User

### AWS account, AWS organization, or IAM principal outside of this account

Enter one or more AWS account IDs, AWS organization IDs, or IAM principal ARNs. Press Enter after each ID or ARN.

Choose AWS account, AWS organization ID, or IAM principal ARN



### You might get access denied

Lake Formation permissions are enforced after you add databases, tables, and principals in hybrid access mode. Make sure that you've completed the required setup for Lake Formation for the permissions to work.

[Learn more](#)

Cancel

Add

## Convertendo um recurso do Lake Formation em um recurso híbrido

Nos casos em que você está usando atualmente as permissões do Lake Formation para seus bancos de dados e tabelas do Catálogo de dados, você pode editar as propriedades de registro de localização para ativar o modo de acesso híbrido. Isso permite que você forneça aos novos diretores acesso aos mesmos recursos usando políticas de permissão do IAM para o Amazon S3 AWS Glue e ações sem interromper as permissões existentes do Lake Formation.

Descrição do cenário - As etapas a seguir pressupõem que você tenha um local de dados registrado no Lake Formation e tenha configurado permissões para entidades principais em bancos de dados, tabelas ou colunas que apontam para esse local. Se o local foi registrado com uma função vinculada ao serviço, você não poderá atualizar os parâmetros de localização e ativar o modo de acesso híbrido. Por padrão, o grupo `IAMAllowedPrincipals` tem Super permissões no banco de dados e em todas as suas tabelas.

### Important

Não atualize um registro de localização para o modo de acesso híbrido sem optar pelos diretores que estão acessando os dados nesse local.

## Ativando o modo de acesso híbrido para um local de dados registrado no Lake Formation

1.

### Warning

Não recomendamos converter um local de dados gerenciado do Lake Formation para o modo de acesso híbrido para evitar a interrupção das políticas de permissão de outros usuários ou cargas de trabalho existentes.

Opte pelas entidades principais existentes que tenham permissões do Lake Formation.

1. Liste e analise as permissões que você concedeu às entidades principais em bancos de dados e tabelas. Para ter mais informações, consulte [Visualizar permissões de banco de dados e tabelas no Lake Formation](#).
2. Escolha o Modo de acesso híbrido em Permissões na barra de navegação esquerda e escolha Adicionar.

3. Na página Adicionar entidades principais e recursos, escolha os bancos de dados e tabelas do local de dados do Amazon S3 que você deseja usar no modo de acesso híbrido. Escolha as entidades principais que já possuem permissões do Lake Formation.
  4. Escolha Adicionar para ativar as entidades principais para usar as permissões do Lake Formation no modo de acesso híbrido.
2. Atualize o registro de bucket/prefixo do Amazon S3 escolhendo a opção de Modo de acesso híbrido.

## Console

1. Faça login no console do Lake Formation como administrador do data lake.
2. No painel de navegação, em Registrar e ingerir, escolha Locais do data lake.
3. Selecione um local e, no menu Ações, escolha Editar.
4. Escolha o Modo de acesso híbrido.
5. Escolha Salvar.
6. Em Catálogo de dados, selecione o banco de dados ou a tabela e conceda permissões Super ou All para o grupo virtual chamado IAMAllowedPrincipals.
7. Verifique se o acesso dos usuários existentes do Lake Formation não foi interrompido quando você atualizou as propriedades de registro de localização. Faça login no console do Athena como entidade principal do Lake Formation e execute um exemplo de consulta em uma tabela que está apontando para o local atualizado.

Da mesma forma, verifique o acesso dos AWS Glue usuários que estão usando as políticas de permissões do IAM para acessar o banco de dados e as tabelas.

## AWS CLI

Veja a seguir um exemplo para registrar uma localização de dados no Lake Formation HybridAccessEnabled com:true/false. O valor padrão do parâmetro HybridAccessEnabled é falso. Substitua o caminho, o nome da função e o ID da AWS conta do Amazon S3 por valores válidos.

```
aws lakeformation update-resource --cli-input-json file://file path
json:
{
  "ResourceArn": "arn:aws:s3:::<s3-path>",
```

```
"RoleArn": "arn:aws:iam::<123456789012>:role/<test>",  
"HybridAccessEnabled": true  
}
```

## Compartilhamento de um AWS Glue recurso usando o modo de acesso híbrido

Compartilhe dados com outra pessoa Conta da AWS ou com um diretor de outra pessoa Conta da AWS aplicando as permissões do Lake Formation sem interromper o acesso baseado em IAM dos usuários existentes do Catálogo de Dados.

Descrição do cenário - A conta do produtor tem um banco de dados do catálogo de dados que tem acesso controlado usando as principais políticas do IAM para Amazon S3 e AWS Glue ações. A localização dos dados do banco de dados não está registrada no Lake Formation. O IAMAllowedPrincipals grupo, por padrão, tem Super permissões no banco de dados e em todas as suas tabelas.

Conceder permissões entre contas do Lake Formation no modo de acesso híbrido

### 1. Configuração da conta de produtor

1. Faça login no console do Lake Formation usando uma função que tenha permissão do IAM `lakeformation:PutDataLakeSettings`.
2. Acesse as Configurações do Catálogo de dados e escolha `Version 4` para as Configurações da versão entre contas.

Se você estiver usando a versão 1 ou 2, consulte as instruções [Como atualizar as configurações da versão de compartilhamento de dados entre contas](#) sobre como atualizar para a versão 3.

Não são necessárias alterações na política de permissão ao atualizar da versão 3 para a 4.

3. Registre a localização do Amazon S3 do banco de dados ou tabela que você planeja compartilhar no modo de acesso híbrido.
4. Verifique se a permissão Super para o grupo IAMAllowedPrincipals existe nos bancos de dados e tabelas nos quais você registrou a localização dos dados no modo de acesso híbrido na etapa acima.
5. Conceda permissões do Lake Formation para AWS organizações, unidades organizacionais (OUs) ou diretamente com um diretor do IAM em outra conta.

6. Se você estiver concedendo permissões diretamente a uma entidade principal do IAM, opte pela entidade principal da conta de consumidor para aplicar as permissões do Lake Formation no modo de acesso híbrido, ativando a opção Tornar as permissões do Lake Formation efetivas imediatamente.

Se você estiver concedendo permissões entre contas para outra AWS conta, ao optar pela conta, as permissões do Lake Formation serão aplicadas somente para os administradores dessa conta. O administrador do data lake da conta do destinatário precisa reduzir as permissões em cascata e optar pelas entidades principais da conta para aplicar as permissões do Lake Formation aos recursos compartilhados que estão no modo de acesso híbrido.

Se você escolher a opção Recursos correspondidos por LF-Tags para conceder permissões entre contas, você precisa primeiro concluir a etapa de concessão de permissões. Você pode optar por incluir entidades principais e recursos no modo de acesso híbrido como uma etapa separada, escolhendo o Modo de acesso híbrido em Permissões na barra de navegação esquerda do console do Lake Formation. Em seguida, escolha Adicionar para adicionar os recursos e as entidades principais aos quais você deseja aplicar as permissões do Lake Formation.

## 2. Configuração de conta de consumidor

1. Faça login no console do Lake Formation em <https://console.aws.amazon.com/lakeformation/> como administrador do data lake.
2. Acesse <https://console.aws.amazon.com/ram> e aceite o convite de compartilhamento de recursos. A guia Compartilhado comigo no AWS RAM console exibe o banco de dados e as tabelas compartilhadas com sua conta.
3. Crie um link de recurso para o banco de dados e/ou tabela compartilhada no Lake Formation.
4. Conceda a permissão Describe no link do recurso e a permissão Grant on target (no recurso compartilhado original) às entidades principais do IAM em sua conta (de consumidor).
5. Conceda permissões do Lake Formation no banco de dados ou na tabela compartilhada com você às entidades principais da sua conta. Opte pelas entidades principais e recursos para aplicar as permissões do Lake Formation no modo de acesso híbrido, ativando a opção Tornar as permissões do Lake Formation efetivas imediatamente.
6. Teste as permissões da entidade principal do Lake Formation executando exemplos de consultas do Athena. Teste o acesso existente de seus AWS Glue usuários com as principais políticas do IAM para Amazon S3 e AWS Glue ações.

(Opcional) Remova a política de bucket do Amazon S3 para acesso a dados e políticas de entidades principais do IAM para AWS Glue e acesso a dados do Amazon S3 para as entidades principais que você configurou para usar permissões do Lake Formation.

## Compartilhando um recurso do Lake Formation usando o modo de acesso híbrido

Permita que novos usuários do Catálogo de dados em uma conta externa acessem bancos de dados e tabelas do Catálogo de dados usando políticas baseadas no IAM sem interromper as permissões de compartilhamento entre contas existentes do Lake Formation.

Descrição do cenário - A conta de produtor tem banco de dados e tabelas gerenciados pelo Lake Formation que são compartilhados com uma conta externa (consumidor) no nível da conta ou no nível de entidade principal do IAM. A localização dos dados do banco de dados é registrada no Lake Formation. O grupo `IAMAllowedPrincipals` não tem permissões `Super` no banco de dados e em suas tabelas.

Conceder acesso entre contas a novos usuários do Catálogo de dados por meio de políticas baseadas em IAM sem interromper as permissões existentes do Lake Formation


### 1. Configuração da conta de produtor

1. Faça login no console do Lake Formation usando uma função que `lakeformation:PutDataLakeSettings`.
2. Em Configurações do Catálogo de dados, escolha `Version 4` para as Configurações da versão entre contas.

Se você estiver usando a versão 1 ou 2, consulte as instruções [Como atualizar as configurações da versão de compartilhamento de dados entre contas](#) sobre como atualizar para a versão 3.

Não são necessárias alterações na política de permissão para atualizar da versão 3 para a 4.

3. Liste as permissões que você concedeu às entidades principais em bancos de dados e tabelas. Para ter mais informações, consulte [Visualizar permissões de banco de dados e tabelas no Lake Formation](#).
4. Reconceda as permissões existentes entre contas do Lake Formation optando por entidades principais e recursos.

 Note

Antes de atualizar um registro de localização de dados para o modo de acesso híbrido para conceder permissões entre contas, você precisa conceder novamente pelo menos um compartilhamento de dados entre contas por conta. Essa etapa é necessária para atualizar as permissões AWS RAM gerenciadas anexadas ao compartilhamento AWS RAM de recursos.

Em julho de 2023, o Lake Formation atualizou as permissões AWS RAM gerenciadas usadas para compartilhar bancos de dados e tabelas:

- `arn:aws:ram::aws:permission/AWSRAMLFEnabledGlueAllTablesReadWriteForDatabase` (política de compartilhamento em nível de banco de dados)
- `arn:aws:ram::aws:permission/AWSRAMLFEnabledGlueTableReadWrite` (política de compartilhamento em nível de tabela)

As concessões de permissão entre contas feitas antes de julho de 2023 não têm essas AWS RAM permissões atualizadas.

Se você concedeu permissões entre contas diretamente às entidades principais, precisará devolvê-las individualmente às entidades principais. Se você pular essa etapa, as entidades principais que acessam o recurso compartilhado podem receber um erro de combinação ilegal.

5. Acesse <https://console.aws.amazon.com/ram>.
6. A guia Compartilhado por mim no AWS RAM console exibe os nomes do banco de dados e da tabela que você compartilhou com uma conta externa ou principal.

Certifique-se de que as permissões anexadas ao recurso compartilhado tenham o ARN correto.

7. Verifique se os recursos no AWS RAM compartilhamento estão no Associated status. Se o status for exibido como Associating, espere até que eles entrem no status Associated. Se o status for Failed, pare e entre em contato com a equipe de serviço do Lake Formation.
8. Escolha o Modo de acesso híbrido em Permissões na barra de navegação esquerda e escolha Adicionar.
9. A página Adicionar entidades principais e recursos mostra os bancos de dados e/ou tabelas e as entidades principais que têm acesso. Você pode fazer as atualizações necessárias adicionando ou removendo entidades principais e recursos.

- 10 Escolha as entidades principais com permissões do Lake Formation para o banco de dados e as tabelas que você deseja alterar para o modo de acesso híbrido. Escolha os bancos de dados e tabelas.
  - 11 Escolha Adicionar para optar pelas entidades principais para aplicar as permissões do Lake Formation no modo de acesso híbrido.
  - 12 Conceda a permissão Super ao grupo virtual IAMAllowedPrincipals em seu banco de dados e nas tabelas selecionadas.
  - 13 Edite o registro Lake Formation da localização do Amazon S3 para o modo de acesso híbrido.
  - 14 Conceda permissões para os AWS Glue usuários na conta externa (consumidor) usando políticas de permissão do IAM para ações do Amazon S3 AWS Glue .
2. Configuração de conta de consumidor

1. Faça login no console do Lake Formation em <https://console.aws.amazon.com/lakeformation/> como administrador do data lake.
2. Acesse <https://console.aws.amazon.com/ram> e aceite o convite de compartilhamento de recursos. A guia Recursos compartilhados comigo na AWS RAM página exibe os nomes do banco de dados e das tabelas que são compartilhados com sua conta.

Para o AWS RAM compartilhamento, certifique-se de que a permissão anexada tenha o ARN correto do convite compartilhado AWS RAM . Verifique se os recursos no AWS RAM compartilhamento estão no Associated status. Se o status for exibido como Associating, espere até que eles entrem no status Associated. Se o status for Failed, pare e entre em contato com a equipe de serviço do Lake Formation.

3. Crie um link de recurso para o banco de dados e/ou tabela compartilhada no Lake Formation.
4. Conceda a permissão Describe no link do recurso e a permissão Grant on target (no recurso compartilhado original) às entidades principais do IAM em sua conta (de consumidor).
5. Em seguida, configure as permissões do Lake Formation para as entidades principais da sua conta no banco de dados ou na tabela compartilhada.

Na barra de navegação esquerda, em Permissões, escolha Modo de acesso híbrido.

6. Escolha Adicionar na seção inferior da página do Modo de acesso híbrido para optar pelas entidades principais e pelo banco de dados ou tabela compartilhados com você na conta de produtor.
7. Conceda permissões para os AWS Glue usuários em sua conta usando políticas de permissão do IAM para ações do Amazon S3 AWS Glue .



8. Teste as permissões e AWS Glue permissões do Lake Formation dos usuários executando exemplos de consultas separadas na tabela usando o Athena

(Opcional) Limpe as políticas de permissão do IAM para o Amazon S3 para as entidades principais que estão no modo de acesso híbrido.

## Removendo entidades principais e recursos do modo de acesso híbrido

Siga estas etapas para remover bancos de dados, tabelas e entidades principais do modo de acesso híbrido.

### Console

1. Faça login no console do Lake Formation em <https://console.aws.amazon.com/lakeformation/>.
2. Em Permissões, escolha o Modo de acesso híbrido.
3. Na página do Modo de acesso híbrido, marque a caixa de seleção ao lado do nome do banco de dados ou da tabela e escolha Remove.
4. Uma mensagem de aviso solicita que você confirme a ação. Escolha Remover.

O Lake Formation não impõe mais permissões para esses recursos, e o acesso a esse recurso será controlado usando IAM e AWS Glue permissões. Isso pode fazer com que o usuário não tenha mais acesso a esse recurso se não tiver as permissões apropriadas do IAM.

### AWS CLI

O exemplo a seguir mostra como remover um recurso do modo de acesso híbrido.

```
aws lakeformation delete-lake-formation-opt-in --cli-input-json file://file path

json:
{
  "Principal": {
    "DataLakePrincipalIdentifier": "arn:aws:iam::<123456789012>:role/role name"
  },
  "Resource": {
    "Table": {
      "CatalogId": "<123456789012>",
      "DatabaseName": "<database name>",
      "Name": "<table name>"
    }
  }
}
```

```
}  
  }  
}
```

## Visualizando entidades principais e recursos no modo de acesso híbrido

Siga estas etapas para visualizar bancos de dados, tabelas e entidades principais no modo de acesso híbrido.

### Console

1. Faça login no console do Lake Formation em <https://console.aws.amazon.com/lakeformation/>.
2. Em Permissões, escolha o Modo de acesso híbrido.
3. A página do Modo de acesso híbrido mostra os recursos e entidades principais que estão atualmente no modo de acesso híbrido.

### AWS CLI

O exemplo a seguir mostra como listar todas as entidades principais e recursos opcionais que estão no modo de acesso híbrido.

```
aws lakeformation list-lake-formation-opt-ins
```

O exemplo a seguir mostra como listar o opt-in para um par específico de entidade principal e recurso.

```
aws lakeformation list-lake-formation-opt-ins --cli-input-json file://file path  
  
json:  
{  
  "Principal": {  
    "DataLakePrincipalIdentifier": "arn:aws:iam::<account-id>:role/<role name>"  
  },  
  "Resource": {  
    "Table": {
```

```
    "CatalogId": "<account-id>",
    "DatabaseName": "<database name>",
    "Name": "<table name>"
  }
}
```

## Recursos adicionais do

Na postagem do blog a seguir, mostraremos as instruções para integrar as permissões do Lake Formation no modo de acesso híbrido para usuários selecionados, enquanto o banco de dados já está acessível a outros usuários por meio das permissões do IAM e do Amazon S3. Analisaremos as instruções para configurar o modo de acesso híbrido em uma AWS conta e entre duas contas.

- [Apresentando o modo de acesso híbrido AWS Glue Data Catalog para proteger o acesso usando as políticas do Lake Formation e do IAM e do Amazon S3.](#)

## Criando tabelas e bancos de dados do Catálogo de dados

AWS Lake Formation usa o Catálogo de dados AWS Glue para armazenar metadados sobre data lakes, fontes de dados, transformações e destinos. Os metadados sobre fontes e destinos de dados estão na forma de bancos de dados e tabelas. As tabelas armazenam informações sobre os dados subjacentes, incluindo informações de esquema, informações de partição e localização dos dados. Bancos de dados são coleções de tabelas. O Catálogo de dados também contém links de recursos, que são links para bancos de dados e tabelas compartilhados em contas externas e são usados para acesso entre contas aos dados no data lake.

Cada conta AWS tem um Catálogo de dados por região AWS.

### Tópicos

- [Criação de um banco de dados](#)
- [Criar tabelas](#)
- [Trabalhar com visualizações](#)

## Criação de um banco de dados

As tabelas de metadados no Catálogo de dados são armazenadas nos bancos de dados. Você pode criar quantos bancos de dados precisar e conceder permissões diferentes do Lake Formation em cada banco de dados.

Os bancos de dados podem ter uma propriedade de localização opcional. Esse local geralmente está dentro de um local do Amazon Simple Storage Service (Amazon S3) registrado no Lake Formation. Quando você especifica um local, as entidades principais não precisam de permissões de localização de dados para criar tabelas do Catálogo de dados que apontem para locais dentro do local do banco de dados. Para ter mais informações, consulte [Underlying data access control](#).

Para criar um banco de dados usando o console do Lake Formation, você deve estar conectado como administrador do data lake ou criador do banco de dados. Um criador de banco de dados é uma entidade principal que recebeu a permissão `CREATE_DATABASE` do Lake Formation. Você pode ver uma lista de criadores de banco de dados na página Funções e tarefas administrativas do console do Lake Formation. Para ver essa lista, você precisa ter a permissão `lakeformation:ListPermissions` do IAM e estar conectado como administrador do data lake ou como criador de banco de dados com a opção de concessão na permissão `CREATE_DATABASE`.

Para criar um banco de dados

1. Abra o console AWS Lake Formation em <https://console.aws.amazon.com/lakeformation/> e faça login como administrador do data lake ou criador do banco de dados.
2. No painel de navegação, em Catálogo de dados, escolha Bancos de dados.
3. Selecione Criar banco de dados.
4. Na caixa de diálogo Criar banco de dados, insira o nome do banco de dados, a localização opcional e a descrição opcional.
5. Opcionalmente, selecione Usar somente controle de acesso do IAM para novas tabelas nesse banco de dados.

Para obter mais informações sobre esta opção, consulte [the section called “Alterando as configurações padrão do seu data lake”](#).

6. Selecione Criar banco de dados.

## Criar tabelas

As tabelas de metadados AWS Lake Formation contêm informações sobre dados no data lake, incluindo informações de esquema, informações de partição e localização dos dados. Essas tabelas são armazenadas no Catálogo de dados AWS Glue. Você os usa para acessar dados subjacentes no data lake e gerenciar esses dados com as permissões do Lake Formation. As tabelas são armazenadas nos bancos de dados no Catálogo de dados.

Há várias maneiras de criar tabelas do Catálogo de dados:

- Execute um crawler em AWS Glue. Consulte [Definição de crawlers](#) no Guia do desenvolvedor AWS Glue.
- Crie e execute um fluxo de trabalho. Consulte [the section called “Importação de dados usando fluxos de trabalho”](#).
- Crie uma tabela manualmente usando o console do Lake Formation, a API AWS Glue ou a AWS Command Line Interface (AWS CLI).
- Crie uma tabela usando o Amazon Athena.
- Crie um link de recurso para uma tabela em uma conta externa. Consulte [the section called “Criação de links de recursos”](#).

### Criando tabelas no Apache Iceberg

AWS Lake Formationsuporta a criação de tabelas Apache Iceberg que usam o formato de dados Apache Parquet no AWS Glue Data Catalog com dados residentes no Amazon S3. Uma tabela no Catálogo de dados é a definição de metadados que representa os dados em um armazenamento de dados. Por padrão, o Lake Formation cria tabelas do Iceberg v2. Para saber a diferença entre as tabelas da v1 e v2, consulte [Alterações de versão do formato](#) na documentação do Apache Iceberg.

[Apache Iceberg](#) é um formato de tabela aberta para conjuntos de dados analíticos muito grandes. O Iceberg permite mudanças fáceis em seu esquema, também conhecido como evolução do esquema, o que significa que os usuários podem adicionar, renomear ou remover colunas de uma tabela de dados sem interromper os dados subjacentes. O Iceberg também fornece suporte para controle de versão de dados, o que permite que os usuários acompanhem as alterações nos dados ao longo do tempo. Isso ativa o atributo de viagem no tempo, que permite que os usuários acessem e consultem versões históricas dos dados e analisem as alterações nos dados entre atualizações e exclusões.

Você pode usar o console do Lake Formation ou a operação `CreateTable` na API AWS Glue para criar uma tabela Iceberg no Catálogo de dados. Para obter mais informações, consulte [CreateTable action \(Python: create\\_table\)](#).

Ao criar uma tabela do Iceberg no Catálogo de dados, você deve especificar o formato da tabela e o caminho do arquivo de metadados no Amazon S3 para poder realizar leituras e gravações.

Você pode usar o Lake Formation para proteger sua tabela do Iceberg usando permissões de controle de acesso refinadas ao registrar a localização de dados do Amazon S3 com AWS Lake Formation. Para dados de origem no Amazon S3 e metadados que não estão registrados no Lake Formation, o acesso é determinado pelas políticas de permissões do IAM para ações do Amazon S3 e do AWS Glue. Para ter mais informações, consulte [Gerenciando permissões do Lake Formation](#).

#### Note

O Catálogo de dados não oferece suporte à criação de partições e à adição de propriedades da tabela do Iceberg.

## Tópicos

- [Pré-requisitos](#)
- [Criando uma tabela no Iceberg](#)

## Pré-requisitos

Para criar tabelas Iceberg no Catálogo de dados e configurar as permissões de acesso aos dados do Lake Formation, você precisa preencher os seguintes requisitos:

1. Permissões necessárias para criar tabelas do Iceberg sem os dados registrados no Lake Formation.

Além das permissões necessárias para criar uma tabela no Catálogo de dados, o criador da tabela precisa as seguintes permissões:

- `s3:PutObject` no recurso `arn:aws:s3:::{bucketName}`
- `s3:GetObject` no recurso `arn:aws:s3:::{bucketName}`
- `s3:DeleteObject` no recurso `arn:aws:s3:::{bucketName}`

## 2. Permissões necessárias para criar tabelas do Iceberg com dados registrados no Lake Formation:

Para usar o Lake Formation para gerenciar e proteger os dados em seu data lake, registre sua localização no Amazon S3 que tenha os dados para tabelas com o Lake Formation. Isso é para que a Lake Formation possa fornecer credenciais para serviços analíticos AWS como Athena, Redshift Spectrum e Amazon EMR para acessar dados. Para obter mais informações sobre como registrar um local do Amazon S3, consulte [Adicionar uma localização do Amazon S3 ao seu data lake](#).

Uma entidade principal que lê e grava os dados subjacentes registrados no Lake Formation exige as seguintes permissões:

- `lakeformation:GetDataAccess`
- `DATA_LOCATION_ACCESS`

Uma entidade principal que tem permissões de localização de dados em um local também tem permissões de localização em todos os locais secundários.

Para obter mais informações sobre permissões de localização de dados, consulte [Controle de acesso a dados subjacente](#).

Para permitir a compactação, o serviço precisa assumir um perfil do IAM que tenha permissões para atualizar tabelas no catálogo de dados. Para obter detalhes, consulte [Pré-requisitos de otimização de tabelas](#)

### Criando uma tabela no Iceberg

Você pode criar tabelas do Iceberg v1 e v2 usando o console Lake Formation ou AWS Command Line Interface conforme documentado nesta página. Você também pode criar tabelas Iceberg usando o console AWS Glue ou Crawler do AWS Glue. Para obter mais informações, consulte [Catálogo de dados e crawlers](#) no Guia do desenvolvedor AWS Glue.

### Para criar uma tabela no Iceberg

#### Console

1. Faça login no AWS Management Console e abra o console do Lake Formation em <https://console.aws.amazon.com/lakeformation/>.

2. Em Catálogo de dados, escolha Tabelas e use o botão Criar tabela para especificar os seguintes atributos:
  - Nome da tabela: insira um nome para a tabela. Se você estiver usando o Athena para acessar tabelas, use essas [dicas de nomenclatura](#) no Guia do usuário do Amazon Athena.
  - Banco de dados: escolha um banco de dados existente ou crie um novo.
  - Descrição: descrição da tabela. Você pode escrever uma descrição para ajudá-lo a entender o conteúdo da tabela.
  - Formato da tabela: para Formato da tabela, escolha Apache Iceberg.

**Table format**  
Data Catalog managed tables support data compaction for Iceberg table type. [Learn more](#)

Standard AWS Glue table (default)  
Create a standard AWS Glue table.

Apache Iceberg table - New  
Create an Iceberg table that supports automatic data compaction.

Enable compaction  
Enable compaction for open table formats to optimize storage and improve query performance. [View pricing](#)

**IAM role**  
To run compaction, the IAM role assumed by the job should have necessary permissions. [Learn more](#)

Choose an IAM role

View

Create new IAM role

- Ativar compactação: escolha Ativar compactação para compactar objetos pequenos do Amazon S3 na tabela em objetos maiores.
- Perfil do IAM: para executar a compactação, o serviço assume um perfil do IAM em seu nome. Você pode escolher um perfil do IAM usando o menu suspenso. Certifique-se de que a função tenha as permissões necessárias para habilitar a compactação.

Para saber mais sobre as permissões necessárias, consulte [Pré-requisitos de otimização de tabelas](#).

- Localização: especifique o caminho para a pasta no Amazon S3 que armazena a tabela de metadados. O Iceberg precisa de um arquivo de metadados e de um local no Catálogo de dados para poder realizar leituras e gravações.
- Esquema: escolha Adicionar colunas para adicionar colunas e tipos de dados das colunas. Você tem a opção de criar uma tabela vazia e atualizar o esquema posteriormente.



O Catálogo de dados oferece suporte aos tipos de dados do Hive. Para obter mais informações, consulte [Tipos de dados do Hive](#).

O Iceberg permite que você evolua o esquema e a partição depois de criar a tabela. Você pode usar as [consultas do Athena](#) para atualizar o esquema da tabela e as consultas do [Spark](#) para atualizar as partições.

## AWS CLI

```
aws glue create-table \  
  --database-name iceberg-db \  
  --region us-west-2 \  
  --open-table-format-input '{  
    "IcebergInput": {  
      "MetadataOperation": "CREATE",  
      "Version": "2"  
    }  
  }' \  
  --table-input '{"Name": "test-iceberg-input-demo",  
    "TableType": "EXTERNAL_TABLE",  
    "StorageDescriptor": {  
      "Columns": [  
        {"Name": "col1", "Type": "int"},  
        {"Name": "col2", "Type": "int"},  
        {"Name": "col3", "Type": "string"}  
      ],  
      "Location": "s3://DOC_EXAMPLE_BUCKET_ICEBERG/"  
    }  
  }'
```

## Otimizar tabelas Iceberg

Os data lakes do Amazon S3 usando formatos de tabela aberta, como o Apache Iceberg, armazenam os dados como objetos do Amazon S3. Ter milhares de pequenos objetos Amazon S3 em uma tabela de data lake aumenta a sobrecarga de metadados nas tabelas Iceberg e afeta o desempenho de leitura. Para melhor desempenho de leitura por serviços de análise AWS, como Amazon Athena e Amazon EMR e trabalhos de ETL AWS Glue, AWS Glue Data Catalog fornece compactação gerenciada (um processo que compacta pequenos objetos do Amazon S3 em objetos maiores) para tabelas Iceberg no Catálogo de dados. Você pode usar o console, o console AWS

Glue ou a API, AWS CLI ou AWS do Lake Formation para ativar ou desativar a compactação de tabelas individuais do Iceberg que estão no Catálogo de dados.

O otimizador de tabelas monitora constantemente as partições da tabela e inicia o processo de compactação quando o limite é excedido para o número de arquivos e tamanhos de arquivo. No Catálogo de dados, o valor limite padrão para iniciar a compactação é definido como 384 MB, enquanto na biblioteca Iceberg o limite para compactação é de aproximadamente 75% do tamanho do arquivo de destino. O Catálogo de dados executa a compactação sem interferir nas consultas simultâneas. O Catálogo de dados oferece suporte à compactação de dados somente para tabelas no formato Parquet.

Para conhecer tipos de dados, formatos de compactação e limitações compatíveis, consulte [Formatos e limitações compatíveis para compactação gerenciada de dados](#).

## Tópicos

- [Pré-requisitos de otimização de tabelas](#)
- [Permitir compactação](#)
- [Desabilitar compactação](#)
- [Visualizar detalhes de compactação](#)
- [Visualizar métricas do Amazon CloudWatch](#)
- [Excluindo um otimizador](#)

## Pré-requisitos de otimização de tabelas

O otimizador de tabela assume as permissões do perfil do (IAM) AWS Identity and Access Management que você especificou ao habilitar a compactação para uma tabela. O perfil do IAM deve ter as permissões para ler dados e atualizar metadados no Catálogo de dados. Você pode criar um perfil do IAM e anexar as políticas em linha a seguir:

- Adicione a seguinte política em linha que concede ao Amazon S3 permissões de leitura/gravação no local para dados que não estão registrados no Lake Formation. Essa política também inclui permissões para atualizar a tabela no Catálogo de dados e permitir que AWS Glue adicione logs em logs Amazon CloudWatch e a publicação de métricas. Para dados de origem no Amazon S3 que não estão registrados no Lake Formation, o acesso é determinado pelas políticas de permissões do IAM para o Amazon S3 e ações AWS Glue.

Nas políticas em linha a seguir, substitua `bucket-name` pelo nome do bucket do Amazon S3, `aws-account-id` e `region` por um número de conta da AWS e por uma região do catálogo de dados válidos, `database_name` pelo nome do seu banco de dados e `table_name` pelo nome da tabela.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "s3:PutObject",
        "s3:GetObject",
        "s3:DeleteObject"
      ],
      "Resource": [
        "arn:aws:s3:::<bucket-name>/*"
      ]
    },
    {
      "Effect": "Allow",
      "Action": [
        "s3:ListBucket"
      ],
      "Resource": [
        "arn:aws:s3:::<bucket-name>"
      ]
    },
    {
      "Effect": "Allow",
      "Action": [
        "glue:UpdateTable",
        "glue:GetTable"
      ],
      "Resource": [
        "arn:aws:glue:<region>:<aws-account-id>:table/<database-name>/<table-
name>",
        "arn:aws:glue:<region>:<aws-account-id>:database/<database-name>",
        "arn:aws:glue:<region>:<aws-account-id>:catalog"
      ]
    }
  ]
}
```

```

        "Effect": "Allow",
        "Action": [
            "logs:CreateLogGroup",
            "logs:CreateLogStream",
            "logs:PutLogEvents"
        ],
        "Resource": "arn:aws:logs:<region>:<aws-account-id>:log-group:/aws-glue/
iceberg-compaction/logs:*"
    }
]
}

```

- Use a política a seguir para habilitar a compactação de dados registrados no Lake Formation.

Se o perfil de compactação não tiver permissões de grupo IAM\_ALLOWED\_PRINCIPALS concedidas na tabela, o perfil exigirá as permissões ALTER, DESCRIBE, INSERT e DELETE do Lake Formation na tabela.

Para obter mais informações sobre o registro de um bucket do Amazon S3 no Lake Formation, consulte [Adicionar uma localização do Amazon S3 ao seu data lake](#).

```

{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "lakeformation:GetDataAccess"
      ],
      "Resource": "*"
    },
    {
      "Effect": "Allow",
      "Action": [
        "glue:UpdateTable",
        "glue:GetTable"
      ],
      "Resource": [
        "arn:aws:glue:<region>:<aws-account-id>:table/<databaseName>/<tableName>",
        "arn:aws:glue:<region>:<aws-account-id>:database/<database-name>",
        "arn:aws:glue:<region>:<aws-account-id>:catalog"
      ]
    }
  ]
}

```

```

    ]
  },
  {
    "Effect": "Allow",
    "Action": [
      "logs:CreateLogGroup",
      "logs:CreateLogStream",
      "logs:PutLogEvents"
    ],
    "Resource": "arn:aws:logs:<region>:<aws-account-id>:log-group:/aws-glue/iceberg-compaction/logs:*"
  }
]
}

```

- (Opcional) Para compactar tabelas Iceberg com dados em buckets do Amazon S3 criptografados [usando criptografia do lado do servidor](#), a função de compactação exige permissões para descriptografar objetos do Amazon S3 e gerar uma nova chave de dados para gravar objetos nos buckets criptografados. Adicione a seguinte política à chave desejada AWS KMS. Oferecemos suporte somente à criptografia em nível de bucket.

```

{
  "Effect": "Allow",
  "Principal": {
    "AWS": "arn:aws:iam::<aws-account-id>:role/<compaction-role-name>"
  },
  "Action": [
    "kms:Decrypt",
    "kms:GenerateDataKey"
  ],
  "Resource": "*"
}

```

- (Opcional) Para a localização dos dados registrada no Lake Formation, a função usada para registrar a localização exige permissões para descriptografar objetos do Amazon S3 e gerar uma nova chave de dados para gravar objetos nos buckets criptografados. Para ter mais informações, consulte [Registrando uma localização criptografada do Amazon S3](#).
- (Opcional) Se a chave do AWS KMS estiver armazenada em uma conta da AWS diferente, você precisará incluir as permissões a seguir no perfil de compactação.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "kms:Decrypt",
        "kms:GenerateDataKey"
      ],
      "Resource": ["arn:aws:kms:<REGION>:<KEY_OWNER_ACCOUNT_ID>:key/<KEY_ID>" ]
    }
  ]
}
```

- A função que você usa para executar a compactação deve ter a permissão `iam:PassRole` da função.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "iam:PassRole"
      ],
      "Resource": [
        "arn:aws:iam::<account-id>:role/<compaction-role-name>"
      ]
    }
  ]
}
```

- Adicione a política de confiança a seguir à função para que o serviço AWS Glue assumo o perfil do IAM para executar o processo de compactação.

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

```
        "Effect": "Allow",
        "Principal": {
            "Service": "glue.amazonaws.com"
        },
        "Action": "sts:AssumeRole"
    }
]
}
```

## Permitir compactação

Você pode usar o console, o console AWS Glue, AWS CLI, ou a API AWS do Lake Formation para habilitar a compactação de suas tabelas do Apache Iceberg no Catálogo de dados. Para novas tabelas, você pode escolher o Apache Iceberg como formato de tabela e ativar a compactação ao criar a tabela. A compactação está desabilitada por padrão para novas tabelas.

### Console

Para habilitar a compactação

1. Abra o console do Lake Formation em <https://console.aws.amazon.com/lakeformation/> e faça login como administrador do data lake, criador da tabela ou usuário que tenha recebido as permissões `glue:UpdateTable` e `lakeformation:GetDataAccess` na tabela.
2. No painel de navegação, em Catálogo de dados, escolha Tabelas.
3. Na página Tabelas, escolha uma tabela em formato de tabela aberta para a qual você deseja habilitar a compactação e, em seguida, no menu Ações, escolha Habilitar compactação.
4. Você também pode ativar a compactação selecionando a tabela e abrindo a página de Detalhes da tabela. Escolha a guia Otimização de tabela na seção inferior da página e escolha Ativar compactação.

The screenshot displays the AWS Lake Formation console interface for a table named 'icebergtable1'. The left sidebar shows navigation options like Dashboard, Data Catalog, and Permissions. The main content area is divided into sections: 'Table details' and 'Compaction history (0)'. The 'Table details' section includes fields for Database (icebergdemo), Table format (Apache Iceberg), and Location (s3://lmmr-iceberg-demo-shyamrnt-nrt/iceberg/icebergdemo.db/icebergtable1). The 'Compaction history' section shows a message: 'No compaction run. No compaction run to display.' and an 'Enable compaction' button.

5. Em seguida, selecione uma função do IAM existente no menu suspenso com as permissões mostradas na seção [Pré-requisitos de otimização de tabelas](#).

Quando você escolhe a opção Criar um novo perfil do IAM, o serviço cria uma função personalizada com as permissões necessárias para executar a compactação.

The screenshot shows the 'Enable compaction' dialog box. The title is 'Enable compaction' and the subtitle is 'Enable compaction for managed tables in Glue Data Catalog to optimize storage and improve query performances. View pricing'. The 'IAM role' section contains the text: 'This IAM role will run the compaction job on your behalf. Learn more'. Below this, there is a section for 'IAM role' with the text: 'To run compaction, the IAM role assumed by the job should have necessary permissions. Learn more'. A dropdown menu is set to 'Admin', and there are 'View' and 'Create new IAM role' buttons. At the bottom right, there are 'Cancel' and 'Enable compaction' buttons.

Siga as etapas abaixo para atualizar um perfil do IAM existente:

- Para atualizar a política de permissões para o perfil do IAM, no console do IAM, acesse a função do IAM que está sendo usada para executar a compactação.
- Na seção Adicionar permissões, escolha Criar política. Na janela recém-aberta do navegador, crie uma nova política para usar com sua função.
- Na página Criar política, escolha a guia JSON. Copie o código JSON mostrado nos Pré-requisitos no campo do editor de políticas.



## AWS CLI

O exemplo a seguir mostra como habilitar a compactação. Substitua o ID da conta por um ID de conta válido AWS. Substitua o nome do banco de dados e o nome da tabela pelo nome real da tabela do Iceberg e pelo nome do banco de dados. Substitua o `roleArn` pelo nome do recurso (ARN) AWS do perfil do IAM e o nome do perfil do IAM que tem as permissões necessárias para executar a compactação.

```
aws glue create-table-optimizer \  
  --catalog-id 123456789012 \  
  --database-name iceberg_db \  
  --table-name iceberg_table \  
  --table-optimizer-configuration  
'{"roleArn":"arn:aws:iam::123456789012:role/compaction_role", "enabled":'true'}' \  
  --type compaction
```

## AWS API

Chame a operação `CreateTableOptimizer` para ativar a compactação de uma tabela.

Depois de ativar a compactação, a guia Otimização de tabela mostra os seguintes detalhes da compactação (após aproximadamente 15 a 20 minutos):

### Horário de início

A época em que o processo de compactação começou no Lake Formation. O valor é um timestamp no horário UTC.

### Horário de término

A hora em que o processo de compactação terminou no Catálogo de dados. O valor é um timestamp no horário UTC.

### Status

O status de execução da compactação. Os valores são sucesso ou falha.

### Arquivos compactados

Número total de arquivos compactados.

## Bytes compactados

Número total de bytes compactados.

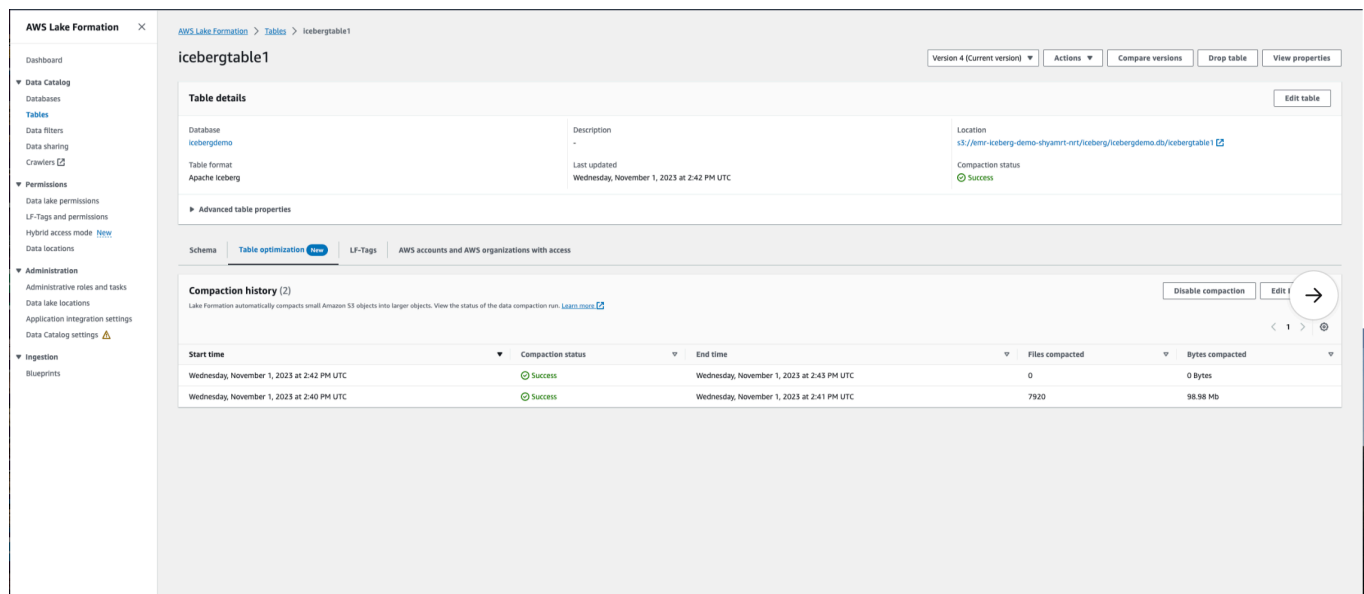
## Desabilitar compactação

Você pode desativar a compactação automática para uma tabela específica do Apache Iceberg usando o console AWS Glue ou AWS CLI.

## Console

1. Escolha Catálogo de dados e escolha Tabelas. Na lista de tabelas, escolha a tabela em formato de tabela aberta que você deseja desativar a compactação.
2. Você pode escolher uma tabela Iceberg e escolher Desativar compactação em Ações.

Você também pode desativar a compactação da tabela escolhendo Desativar compactação na seção inferior da página de Detalhes das tabelas.



The screenshot displays the AWS Lake Formation console interface for a table named 'icebergtable1'. The interface includes a sidebar with navigation options like 'Data Catalog', 'Permissions', and 'Administration'. The main content area shows 'Table details' for 'icebergtable1', including its database, format, location, and last updated time. Below this, there is a 'Compaction history' section with a table listing compaction runs. A 'Disable compaction' button is located in the top right corner of the table details section.

Start time	Compaction status	End time	Files compacted	Bytes compacted
Wednesday, November 1, 2023 at 2:42 PM UTC	Success	Wednesday, November 1, 2023 at 2:43 PM UTC	0	0 Bytes
Wednesday, November 1, 2023 at 2:40 PM UTC	Success	Wednesday, November 1, 2023 at 2:41 PM UTC	7920	98.98 Mb

3. Escolha Desativar compactação na mensagem de confirmação. Você poderá reabilitar a compactação mais tarde.

Após a confirmação, a compactação é desativada e o status de compactação da tabela volta para Off.

## AWS CLI

No exemplo a seguir, substitua o ID da conta por um ID de conta válido AWS. Substitua o nome do banco de dados e o nome da tabela pelo nome real da tabela do Iceberg e pelo nome do banco de dados. Substitua o `roleArn` pelo nome do recurso (ARN) AWS do perfil do IAM e o nome real do perfil do IAM que tem as permissões necessárias para executar a compactação.

```
aws glue update-table-optimizer \  
  --catalog-id 123456789012 \  
  --database-name iceberg_db \  
  --table-name iceberg_table \  
  --table-optimizer-configuration  
'{"roleArn":"arn:aws:iam::123456789012:role/compaction_role", "enabled":'false'}'\  
  --type compaction
```

## AWS API

Chame a `UpdateTableOptimizer` operação para desativar a compactação de uma tabela específica.

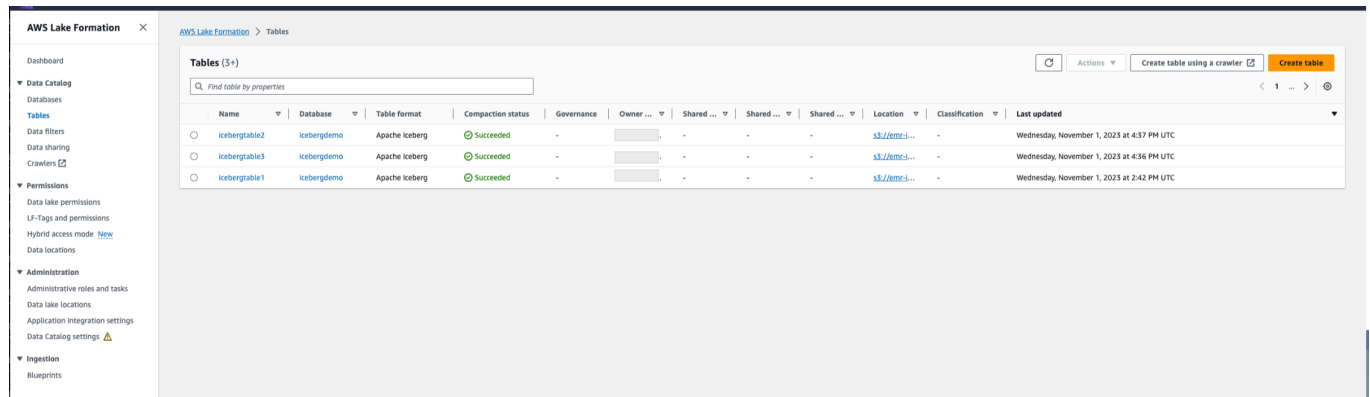
### Visualizar detalhes de compactação

Você pode visualizar o status de compactação do Apache Iceberg no console do Lake Formation, AWS CLI ou usando as operações da API AWS.

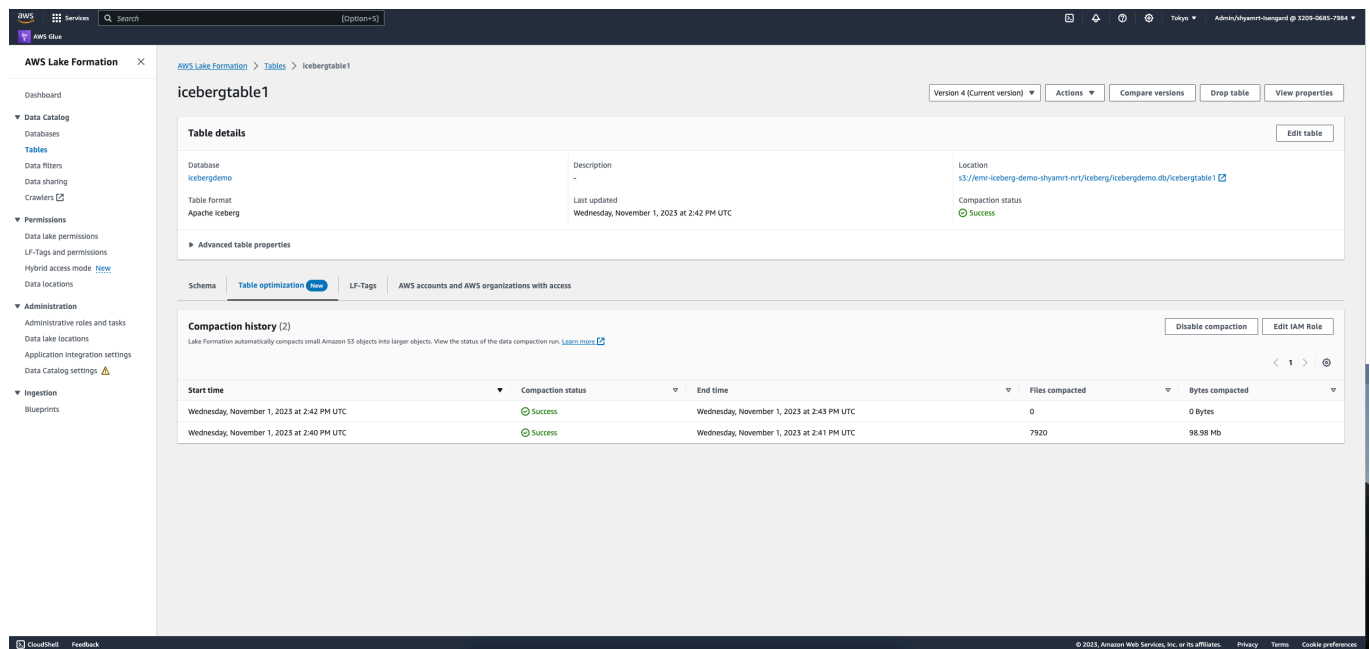
### Console

Para visualizar o status de compactação das tabelas do Iceberg (console)

- Você pode visualizar o status de compactação das tabelas Iceberg no console do Lake Formation escolhendo Tabelas em Catálogo de dados. O campo Status da compactação mostra o status da execução da compactação. Você pode exibir o formato da tabela e o status da compactação usando as preferências da tabela.



- Para ver o histórico de execução de compactação de uma tabela específica, escolha Tabelas em AWS Glue Data Catalog e escolha uma tabela para ver os detalhes da tabela. A guia Otimização da tabela mostra o histórico de compactação da tabela.



## AWS CLI

Você pode visualizar os detalhes de compactação usando AWS CLI.

Nos exemplos a seguir, substitua o ID da conta por um ID de conta válido AWS, o nome do banco de dados e o nome da tabela pelo nome real da tabela Iceberg.

- Para obter os detalhes da última execução de compactação para uma tabela

```
aws get-table-optimizer \
  --catalog-id 123456789012 \
```

```
--database-name iceberg_db \  
--table-name iceberg_table \  
--type compaction
```

- Use o exemplo a seguir para recuperar o histórico de um otimizador para uma tabela específica.

```
aws list-table-optimizer-runs \  
  --catalog-id 123456789012 \  
  --database-name iceberg_db \  
  --table-name iceberg_table \  
  --type compaction
```

- O exemplo a seguir mostra como recuperar a execução de compactação e os detalhes de configuração de vários otimizadores. Você pode especificar no máximo 20 otimizadores.

```
aws glue batch-get-table-optimizer \  
  --entries '[{"catalogId":"123456789012", "databaseName":"iceberg_db",  
  "tableName":"iceberg_table", "type":"compaction"}]'
```

## AWS API

- Use a operação `GetTableOptimizer` para recuperar os detalhes da última execução de um otimizador.
- Use a operação `ListTableOptimizerRuns` para recuperar o histórico de um determinado otimizador em uma tabela específica. Você pode especificar 20 otimizadores em uma única chamada de API.
- Use a operação `BatchGetTableOptimizer` para recuperar detalhes de configuração de vários otimizadores em sua conta. Esta operação não oferece suporte a chamadas entre contas.

## Visualizar métricas do Amazon CloudWatch

Depois de executar a compactação com sucesso, o serviço cria métricas Amazon CloudWatch sobre o desempenho do trabalho de compactação. Você pode acessar as CloudWatch Métricas e escolher

Métricas, Todas as métricas. Você pode filtrar métricas pelo namespace específico (por exemplo AWS Glue), nome da tabela ou nome do banco de dados.

Para obter mais informações, consulte [Visualizar métricas disponíveis](#) no Guia do usuário Amazon CloudWatch.

- Número de bytes compactados
- Número de arquivos compactados
- Número de DPU alocado para o trabalho
- Duração do trabalho (horas)

### Excluindo um otimizador

Você pode excluir um otimizador e os metadados associados à tabela usando nossa operação de API AWS CLI ou AWS.

Execute o comando AWS CLI a seguir para excluir o histórico de compactação de uma tabela.

```
aws glue delete-table-optimizer \  
  --catalog-id 123456789012 \  
  --database-name iceberg_db \  
  --table-name iceberg_table \  
  --type compaction
```

Use a operação `DeleteTableOptimizer` para excluir um otimizador para uma tabela.

### Procurando por tabelas

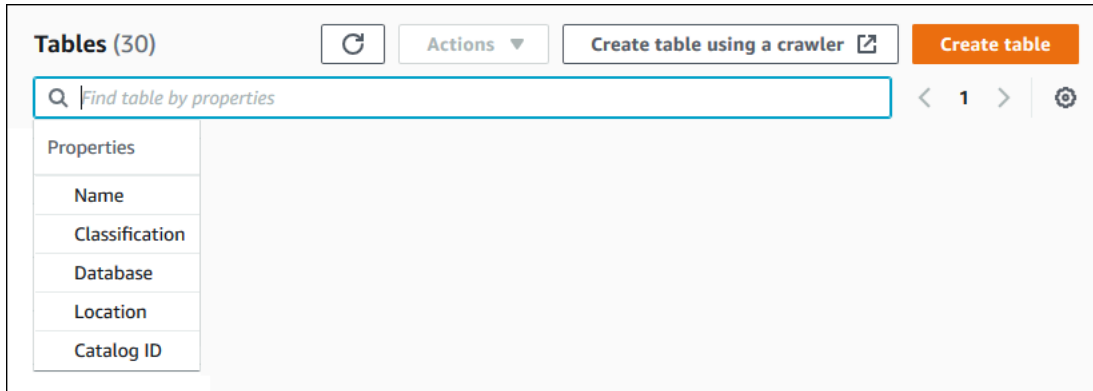
Você pode usar o console AWS Lake Formation para pesquisar tabelas do Catálogo de dados por nome, localização, banco de dados contendo e muito mais. Os resultados da pesquisa mostram somente as tabelas nas quais você tem permissões do Lake Formation.

Para procurar tabelas (console)

1. Faça login no AWS Management Console e abra o console do Lake Formation em <https://console.aws.amazon.com/lakeformation/>.
2. No painel de navegação, selecione Tabelas.

3. Posicione o cursor no campo de pesquisa na parte superior da página. O campo tem o texto do espaço reservado Localizar tabela por propriedades.

O menu Propriedades é exibido, mostrando as várias propriedades da tabela pelas quais pesquisar.



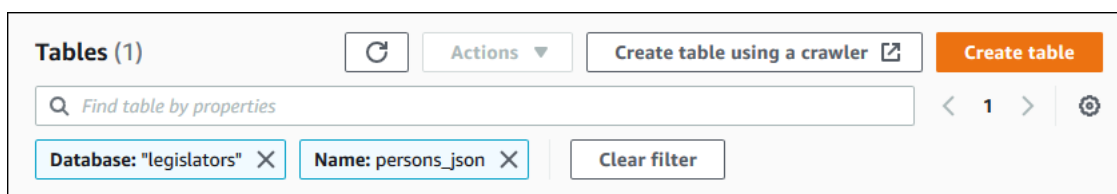
4. Execute um destes procedimentos:

- Pesquise contendo banco de dados.
  1. Escolha Banco de dados no menu Propriedades e, em seguida, escolha um banco de dados no menu Bancos de dados exibido ou digite o nome do banco de dados e pressione Enter.

As tabelas nas quais você tem permissões no banco de dados são listadas.

2. (Opcional) Para restringir a lista a uma única tabela no banco de dados, posicione o cursor no campo de pesquisa novamente, escolha Nome no menu Propriedades e escolha um nome de tabela no menu Tabelas exibido ou digite um nome de tabela e pressione Enter.

A tabela única é listada e tanto o nome do banco de dados quanto o nome da tabela aparecem como blocos sob o campo de pesquisa.



Para ajustar o filtro, feche um dos ladrilhos ou escolha Limpar filtro.

- Pesquise por outras propriedades.
  1. Escolha uma propriedade de pesquisa no menu Propriedades.

Para pesquisar por ID da conta AWS, escolha ID do catálogo no menu Propriedades, insira uma ID de conta válida AWS (por exemplo, 111122223333) e pressione Enter.

Para pesquisar por localização, escolha Localização no menu Propriedades e selecione uma localização no menu Localizações que aparece. Todas as tabelas na localização raiz da localização selecionada (por exemplo, Amazon S3) são retornadas.

## Compartilhamento de tabelas e bancos de dados do Catálogo de dados entre contas AWS

Você pode compartilhar recursos do Catálogo de dados (bancos de dados e tabelas) com contas externas AWS concedendo permissões do Lake Formation sobre os recursos às contas externas. Os usuários podem então executar consultas e trabalhos que unem e consultam tabelas em várias contas. Com algumas restrições, quando você compartilha um recurso do Catálogo de dados com outra conta, as entidades principais dessa conta podem operar nesse recurso como se o recurso estivesse em seu Catálogo de dados.

Você não compartilha recursos com entidades principais específicas em contas externas AWS – você compartilha os recursos com uma conta AWS ou organização. Ao compartilhar um recurso com uma organização AWS, você está compartilhando o recurso com todas as contas em todos os níveis dessa organização. O administrador do data lake em cada conta externa deve então conceder permissões sobre os recursos compartilhados às entidades principais da conta.

Para obter mais informações, consulte [Concedendo e revogando permissões nos recursos do Catálogo de dados](#) e [Compartilhamento de dados entre contas no Lake Formation](#).

 Consulte também:

- [Acessar e visualizar tabelas e bancos de dados compartilhados do Catálogo de Dados](#)
- [Pré-requisitos](#)

## Trabalhar com visualizações

Esse recurso está em prévia de release e sujeito a alterações. Para obter mais informações, consulte a seção Beta e pré-visualizações no documento de [Termos de serviço da AWS](#).



No AWS Glue Data Catalog, visualização é uma tabela virtual na qual o conteúdo é definido por uma consulta que faz referência a uma ou mais tabelas. É possível criar uma visualização que faça referência a até dez tabelas usando editores SQL para Amazon Athena, Amazon Redshift ou Amazon EMR. As tabelas de referência subjacentes de uma visualização podem pertencer ao mesmo banco de dados ou a bancos de dados diferentes na mesma Conta da AWS.

SQL é uma linguagem de programação usada para consultar tabelas, e cada mecanismo analítico da AWS usa a própria variação de SQL ou dialeto de SQL. O catálogo de dados aceita a criação de visualizações usando diferentes dialetos de SQL, desde que cada dialeto faça referência ao mesmo conjunto de tabelas, colunas e tipos de dados. Ao definir um esquema de visualização comum e um objeto de metadados que você pode consultar em vários mecanismos, as visualizações do catálogo de dados permitem usar visualizações uniformes em todo o data lake.

Ao gerenciar visualizações no catálogo de dados, é possível usar o AWS Lake Formation para conceder permissões refinadas por meio do método de recurso nomeado ou usando tags do LF e compartilhá-las entre Contas da AWS, AWS Organizations e unidades organizacionais. Também é possível compartilhar visualizações do catálogo de dados entre Regiões da AWS. Isso permite que os usuários forneçam acesso aos dados entre Regiões da AWS sem duplicar a fonte de dados.

Para obter mais informações sobre compartilhamento de dados entre contas e acesso a dados entre regiões, consulte:

- [Compartilhamento de dados entre contas no Lake Formation](#)
- [Acessar tabelas entre regiões](#)

É possível usar as visualizações do catálogo de dados para:

- Criar e gerenciar permissões em um único esquema de visualização. Isso ajuda a evitar o risco de permissões inconsistentes em visualizações duplicadas criadas em vários mecanismos.
- Conceda permissões aos usuários em uma visualização que faz referência a várias tabelas sem conceder permissões diretamente nas tabelas de referência subjacentes.

Para conhecer as limitações, consulte [Considerações e limitações das visualizações do catálogo de dados](#).

## Tópicos

- [Pré-requisitos para criar visualizações](#)

- [Criar visualizações](#)
- [Conceder permissões nas visualizações do catálogo de dados](#)

## Pré-requisitos para criar visualizações

- Para criar visualizações no catálogo de dados, é necessário registrar os locais de dados subjacentes do Amazon S3 das tabelas de referência no Lake Formation.

Para obter detalhes sobre o registro de dados no Lake Formation, consulte [Adicionar uma localização do Amazon S3 ao seu data lake](#).

- O programador de visualizações deve ser um perfil do IAM. Outras identidades do IAM não podem criar visualizações do catálogo de dados.
- O perfil do IAM que define a visualização deve ter as seguintes permissões:
  - Permissão SELECT completa do Lake Formation com a opção Grantable em todas as tabelas de referência.
  - Uma política de confiança para que o Lake Formation e os serviços do AWS Glue assumam o perfil.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "DataCatalogViewDefinerAssumeRole1",
      "Effect": "Allow",
      "Principal": {
        "Service": [
          "glue.amazonaws.com",
          "lakeformation.amazonaws.com"
        ]
      },
      "Action": "sts:AssumeRole"
    }
  ]
}
```

- A permissão iam:PassRole para o AWS Glue e o Lake Formation.

```
{
```

```

"Version": "2012-10-17",
"Statement": [
  {
    "Sid": "DataCatalogViewDefinerPassRole1",
    "Action": [
      "iam:PassRole"
    ],
    "Effect": "Allow",
    "Resource": "*",
    "Condition": {
      "StringEquals": {
        "iam:PassedToService": [
          "glue.amazonaws.com",
          "lakeformation.amazonaws.com"
        ]
      }
    }
  }
]
}

```

- AWS Glue e permissões do Lake Formation.

```

{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "Glue:GetDatabase",
        "Glue:GetDatabases",
        "Glue:CreateTable",
        "Glue:GetTable",
        "Glue:UpdateTable",
        "Glue>DeleteTable",
        "Glue:GetTables",
        "Glue:SearchTables",
        "Glue:BatchGetPartition",
        "Glue:GetPartitions",
        "Glue:GetPartition",
        "Glue:GetTableVersion",
        "Glue:GetTableVersions",
        "lakeFormation:GetDataAccess",
        "lakeFormation:GetTemporaryTableCredentials",

```

```

        "lakeFormation:GetTemporaryGlueTableCredentials",
        "lakeFormation:GetTemporaryUserCredentialsWithSAML"
    ],
    "Resource": "*"
  }
]
}

```

- Não será possível criar visualizações se o banco de dados sob o qual a visualização está sendo criada tiver uma permissão Super ou ALL concedida ao grupo IAMAllowedPrincipals. Para revogar a permissão Super do grupo IAMAllowedPrincipals em um banco de dados, consulte [Etapa 4: mude seus armazenamentos de dados para o modelo de permissões do Lake Formation](#).

Se as configurações de data lake existentes não permitirem a definição de um grupo CreateTableDefaultPermissions ou IAMAllowedPrincipals vazio, será possível criar um banco de dados e codificar a configuração do data lake usando a estrutura a seguir.

```

{
  "DataLakeSettings": {
    "DataLakeAdmins": [
      {
        "DataLakePrincipalIdentifier":
"arn:aws:iam::<AccountId>:user/<Username>"
      }
    ],
    "CreateTableDefaultPermissions": [
      {
        "Principal": {
          "DataLakePrincipalIdentifier": "IAM_ALLOWED_PRINCIPALS"
        },
        "Permissions": []
      }
    ]
  }
}

```

## Criar visualizações

É possível usar editores SQL para Athena, Amazon Redshift ou Amazon EMR para criar visualizações no AWS Glue Data Catalog.

Para obter mais informações sobre a sintaxe para criar e gerenciar visualizações do catálogo de dados, consulte:

- [Uso das visualizações do AWS Glue Data Catalog](#) no Guia do usuário do Amazon Athena.
- [Creating views in the AWS Glue Data Catalog](#) no Guia do desenvolvedor do banco de dados do Amazon Redshift.
- [Working with AWS Glue Data Catalog views](#) no Guia de gerenciamento do Amazon EMR.

Depois de criar uma visualização do catálogo de dados, os detalhes da visualização no console do Lake Formation.

1. Selecione Visualizações no catálogo de dados no console do Lake Formation.
2. Uma lista das visualizações disponíveis é exibida na página de visualizações.
3. Selecione uma visualização na lista e a página de detalhes mostrará os atributos da visualização.

[AWS Lake Formation](#) > [Views](#) > europe\_players

## europe\_players

Version 1 (Current version) ▾ Actions ▾

### Details

Name europe_players	Database views_demo_database	Definer role admin <a href="#">↗</a>
Last updated November 22, 2023 at 10:41 PM UTC	Status ✔ Ready	Description -

Schema | **SQL definitions** | LF-Tags | Cross-account access | Underlying tables

### SQL definitions (2)

Add SQL definition ▾

List of available SQL definitions in different engines. Choose an engine from the list to add or edit the definition.

Engine name ▲	Version ▾	Status ▾	SQL statement	Edit definition <a href="#">↗</a>
Athena	3	✔ Ready	View	<a href="#">Amazon Athena</a>
Redshift	1.0	✔ Ready	View	<a href="#">Amazon Redshift</a>

## Esquema

Selecione uma linha Column e selecione Editar tags do LF para atualizar os valores das tags ou atribuir novas tags do LF.

## Definições de SQL

É possível ver uma lista das definições de SQL disponíveis. Selecione Adicionar definição de SQL e escolha um mecanismo de consulta para adicionar uma definição de SQL. Selecione um mecanismo de consulta (Athena ou Amazon Redshift) na coluna Edit definition para atualizar as definições de SQL.

## Tags do LF

Selecione Editar tags do LF para editar valores para uma tag ou atribuir novas tags. É possível usar tags do LF para conceder permissões nas visualizações.

## Acesso entre contas

É possível ver uma lista de Contas da AWS, organizações e unidades organizacionais (UOs) que você compartilhou na visualização do catálogo de dados.

## Tabelas subjacentes

As tabelas subjacentes referenciadas na definição de SQL usada para criar a exibição são mostradas nessa guia.

## Conceder permissões nas visualizações do catálogo de dados

Depois de criar visualizações, é possível conceder permissões de data lake em visualizações a entidades principais entre Contas da AWS, organizações e unidades organizacionais. Para obter mais informações sobre como conceder permissões, consulte [Conceder permissões em visualizações usando o método de recurso nomeado](#).

## Importação de dados usando fluxos de trabalho no Lake Formation

Com o AWS Lake Formation, você pode importar seus dados usando fluxos de trabalho. Um fluxo de trabalho define a fonte de dados e o cronograma para importar dados para o seu data lake. É um contêiner para crawlers do AWS Glue, trabalhos e acionadores usados para orquestrar os processos de carregamento e atualização do data lake.

## Tópicos

- [Plantas e fluxos de trabalho em Lake Formation](#)
- [Criação de um fluxo de trabalho](#)
- [Executar um fluxo de trabalho](#)

## Plantas e fluxos de trabalho em Lake Formation

Um fluxo de trabalho encapsula uma atividade complexa de vários trabalhos de vários trabalhos de vários trabalhos de extração, transformação e carregamento (ETL). Os fluxos de trabalho geram crawlers do AWS Glue, trabalhos e gatilhos para orquestrar o carregamento e a atualização dos dados. O Lake Formation executa e rastreia um fluxo de trabalho como uma entidade única. Você pode configurar um fluxo de trabalho para ser executado sob demanda ou de acordo com um cronograma.

Os fluxos de trabalho que você cria no Lake Formation são visíveis no console do AWS Glue como um gráfico acíclico direcionado (DAG). Cada nó do DAG é uma tarefa, um crawler ou um gatilho. Para monitorar o progresso e solucionar problemas, você pode acompanhar o status de cada nó no fluxo de trabalho.

Quando um fluxo de trabalho do Lake Formation é concluído, o usuário que executou o fluxo de trabalho recebe a permissão SELECT do Lake Formation nas tabelas do Catálogo de Dados que o fluxo de trabalho cria.

Você também pode criar fluxos de trabalho no AWS Glue. No entanto, como o Lake Formation permite que você crie um fluxo de trabalho a partir de um esquema, criar fluxos de trabalho é muito mais simples e automatizado no Lake Formation. Lake Formation fornece os seguintes tipos de esquemas:

- Instantâneo do banco de dados: carrega ou recarrega dados de todas as tabelas no data lake a partir de uma fonte JDBC. Você pode excluir alguns dados da fonte com base em um padrão de exclusão.
- Banco de dados incremental: carrega somente novos dados no data lake a partir de uma fonte JDBC, com base em marcadores previamente definidos. Você especifica as tabelas individuais no banco de dados de origem do JDBC a serem incluídas. Para cada tabela, você escolhe as colunas dos favoritos e a ordem de classificação dos favoritos para acompanhar os dados que foram carregados anteriormente. Na primeira vez que você executa um esquema de banco de dados incremental em um conjunto de tabelas, o fluxo de trabalho carrega todos os dados das tabelas e define marcadores para a próxima execução do esquema de banco de dados incremental.

Portanto, você pode usar um esquema de banco de dados incremental em vez do esquema de instantâneo do banco de dados para carregar todos os dados, desde que você especifique cada tabela na fonte de dados como um parâmetro.

- Arquivo de log: carrega dados em massa de fontes de arquivos de log, incluindo AWS CloudTrail e logs do Elastic Load Balancing e do Application Load Balancer.

Use a tabela a seguir para ajudar a decidir se deve usar um snapshot de banco de dados ou um esquema de banco de dados incremental.

Use o instantâneo do banco de dados quando...	Use o banco de dados incremental quando...
<ul style="list-style-type: none"> <li>• A evolução do esquema é flexível. (As colunas são renomeadas, as colunas anteriores são excluídas e novas colunas são adicionadas em seu lugar.)</li> <li>• É necessária uma consistência completa entre a origem e destino.</li> </ul>	<ul style="list-style-type: none"> <li>• A evolução do esquema é incremental. (Há somente adição sucessiva de colunas.)</li> <li>• Somente novas linhas são adicionadas; as linhas anteriores não são atualizadas.</li> </ul>

### Note

Os usuários não podem editar plantas e fluxos de trabalho criados pelo Lake Formation.

## Criação de um fluxo de trabalho

Antes de começar, verifique se você concedeu as permissões de dados e as permissões de localização de dados necessárias para a função `LakeFormationWorkflowRole`. Isso é para que o fluxo de trabalho possa criar tabelas de metadados no catálogo de dados e gravar dados em locais de destino no Amazon S3. Para obter mais informações, consulte [\(Opcional\) Crie uma função do IAM para fluxos de trabalho](#) e [Visão geral das permissões do Lake Formation](#).

Para criar um fluxo de trabalho a partir de um esquema

1. Abra o console do AWS Lake Formation em <https://console.aws.amazon.com/lakeformation/>. Faça login como administrador do data lake ou como usuário com permissões de engenheiro de




dados. Para obter mais informações, consulte [Referência de personas e permissões do IAM do Lake Formation](#).

2. No painel de navegação, selecione esquemas e escolha Usar esquema.
3. Na página Usar um esquema, escolha um quadro para selecionar o tipo de esquema.
4. Em Fonte de importação, especifique a fonte de dados.

Se você estiver importando de uma fonte JDBC, especifique o seguinte:

- Conexão de banco de dados: escolha uma conexão na lista. Crie conexões adicionais usando o console do AWS Glue. O nome de usuário JDBC e a senha na conexão determinam os objetos do banco de dados aos quais o fluxo de trabalho tem acesso.
- Caminho dos dados de origem: insira *<banco de dados>/<esquema>/<tabela>* or *<banco de dados>/<tabela>*, dependendo do produto do banco de dados. Oracle Database e MySQL não oferecem suporte a esquema no caminho. Você pode substituir o caractere percentual (%) por *<esquema>* ou *<tabela>*. Por exemplo, para um banco de dados Oracle com um identificador de sistema (SID) orcl, informe orcl/% para importar todas as tabelas às quais o usuário nomeado na conexão tem acesso.

 Important

Este campo diferencia letras maiúsculas de minúsculas. O fluxo de trabalho falhará se houver uma incompatibilidade de maiúsculas e minúsculas em qualquer um dos componentes.

Se você especificar um banco de dados MySQL, o ETL do AWS Glue usa o driver JDBC Mysql5 por padrão, portanto, o MySQL8 não é suportado nativamente. Você pode editar o script de trabalho ETL para usar um parâmetro `customJdbcDriverS3Path` conforme descrito em [Valores JDBC ConnectionType](#) no Guia do desenvolvedor AWS Glue para usar um driver JDBC diferente que suporte MySQL8.

Se você estiver importando de um arquivo de log, certifique-se de que a função especificada para o fluxo de trabalho (a “função do fluxo de trabalho”) tenha as permissões necessárias do IAM para acessar a fonte de dados. Por exemplo, para importar logs do AWS CloudTrail, o usuário deve ter as permissões `cloudtrail:DescribeTrails` e

`cloudtrail:LookupEvents` para ver a lista de logs do CloudTrail ao criar o fluxo de trabalho, e a função do fluxo de trabalho deve ter permissões na localização do CloudTrail no Amazon S3.

5. Faça um dos seguintes procedimentos:

- Para o tipo de esquema de instantâneo do banco de dados, identifique opcionalmente um subconjunto de dados a serem importados especificando um ou mais padrões de exclusão. Esses padrões de exclusão são padrões `glob` no estilo Unix. Elas são armazenadas como uma propriedade das tabelas criadas pelo fluxo de trabalho.

Para obter detalhes sobre os padrões de exclusão disponíveis, consulte [Incluir e excluir padrões](#) no Guia do desenvolvedor AWS Glue.


- Para o tipo de esquema de banco de dados incremental, especifique os seguintes campos. Adicione uma linha para cada tabela a ser importada.

Nome da tabela

Tabela a ser importada. Deve estar em letras minúsculas.

Teclas de marcadores

Lista delimitada por vírgula dos nomes das colunas que definem as teclas dos favoritos. Se estiver em branco, a chave primária será usada para determinar novos dados. As maiúsculas e minúsculas de cada coluna devem corresponder às maiúsculas e minúsculas, conforme definido na fonte de dados.

 Note

A chave primária se qualifica como a chave de bookmark padrão apenas se estiver aumentando ou diminuindo sequencialmente (sem lacunas). Se você quiser usar a chave primária como chave de marcador e ela tiver lacunas, você deverá nomear a coluna da chave primária como chave de marcador.

Pedido de favoritos

Quando você escolhe Ascendente, as linhas com valores maiores que os marcados são identificadas como novas. Quando você escolhe Descendente, as linhas com valores menores que os viores marcados são identificadas como novas.

## Esquema de particionamento

(Opcional) Lista de colunas-chave de particionamento, delimitada por barras (/). Exemplo: year/month/day.

**Incremental data**  
Enter tables in the data source to import along with bookmark columns to determine previously imported data.

Table name	Bookmark keys	Bookmark order	Partitioning scheme - optional	
<input type="text" value="Enter a table name"/>	<input type="text" value="Enter a bookmark"/> <small>Comma-delimited list of bookmark columns.</small>	<input type="text" value="Choose a sort. ▼"/>	<input type="text" value="Type partitioning"/>	<input type="button" value="Remove"/>
<input type="button" value="Add"/>				

Para obter mais informações, consulte [Rastreamento de dados processados usando marcadores de trabalho](#) no Guia do desenvolvedor do AWS Glue.

6. Em Importar destino, especifique o banco de dados de destino, a localização de destino do Amazon S3 e o formato dos dados.

Certifique-se de que a função do fluxo de trabalho tenha as permissões necessárias do Lake Formation no banco de dados e no local de destino do Amazon S3.

### Note

Atualmente, os esquemas não oferecem suporte à criptografia de dados no destino.

7. Escolha uma frequência de importação.

Você pode especificar uma expressão cron com a opção Personalizada.

8. Em Opções de importação:

- a. Insira um nome de fluxo de trabalho.
- b. Para a função, escolha a função LakeFormationWorkflowRole, que você criou em [\(Opcional\) Crie uma função do IAM para fluxos de trabalho](#).
- c. Opcionalmente, especifique um prefixo de tabela. O prefixo é anexado aos nomes das tabelas do Catálogo de Dados que o fluxo de trabalho cria.

9. Selecione Criar e aguarde até que o console informe que o fluxo de trabalho foi criado com sucesso.

**i** Tip

Você recebeu a seguinte mensagem de erro?

```
User: arn:aws:iam::<account-id>:user/<username> is not authorized
to perform: iam:PassRole on resource:arn:aws:iam::<account-
id>:role/<rolename>...
```

Nesse caso, verifique se você substituiu `<account-id>` por um número de conta AWS válido em todas as políticas.

**i** Consulte também:

- [Plantas e fluxos de trabalho em Lake Formation](#)

## Executar um fluxo de trabalho

Você pode executar um fluxo de trabalho usando o console do Lake Formation, o console do AWS Glue, a interface de linha de comando (AWS CLI) do AWS Glue ou a API.

Para executar um fluxo de trabalho (console do Lake Formation)

1. Abra o console do AWS Lake Formation em <https://console.aws.amazon.com/lakeformation/>. Faça login como administrador do data lake ou como usuário com permissões de engenheiro de dados. Para obter mais informações, consulte [Referência de personas e permissões do IAM do Lake Formation](#).
2. No painel de navegação, escolha Esquemas.
3. Na página Esquemas, selecione o fluxo de trabalho. Em seguida, no menu Ações, escolha Iniciar.
4. À medida que o fluxo de trabalho é executado, você pode ver o progresso na coluna de Status da última execução. Selecione o botão de atualizar ocasionalmente.

O status vai de EM EXECUÇÃO para Descoberta, para Importação e para CONCLUÍDO.

Quando o fluxo de trabalho for concluído:

- O Catálogo de Dados tem novas tabelas de metadados.

- Seus dados são ingeridos no data lake.

Se o fluxo de trabalho falhar, faça o seguinte:

- a. Selecione o fluxo de trabalho. Escolha Ações e Exibir gráfico.

O fluxo de trabalho é aberto no console do AWS Glue.

- b. Certifique-se de que o fluxo de trabalho esteja selecionado e escolha a guia Histórico.
- c. Em Histórico, selecione a execução mais recente e selecione Exibir informações da execução.
- d. Selecione um trabalho ou crawler com falha no gráfico dinâmico (runtime) e revise a mensagem de erro. Os nós com falha são vermelhos ou amarelos.

 Consulte também:

- [Plantas e fluxos de trabalho em Lake Formation](#)

# Gerenciando permissões do Lake Formation

O Lake Formation fornece controles de acesso central para dados em seu data lake. Você pode definir regras baseadas em políticas de segurança para seus usuários e aplicativos por função no Lake Formation, e a integração com AWS Identity and Access Management autentica esses usuários e funções. Depois que as regras são definidas, o Lake Formation aplica seus controles de acesso em granularidade em nível de tabela e coluna para usuários do Amazon Redshift Spectrum e do Amazon Athena.

## Tópicos

- [Concedendo permissões de localização de dados](#)
- [Concedendo e revogando permissões nos recursos do Catálogo de dados](#)
- [Exemplo de cenário de permissões](#)
- [Filtragem de dados e segurança em nível de célula no Lake Formation](#)
- [Visualizar permissões de banco de dados e tabelas no Lake Formation](#)
- [Revogando a permissão usando o console Lake Formation](#)
- [Compartilhamento de dados entre contas no Lake Formation](#)
- [Acessar e visualizar tabelas e bancos de dados compartilhados do Catálogo de Dados](#)
- [Criação de links de recursos](#)
- [Acessar tabelas entre regiões](#)

## Concedendo permissões de localização de dados

As permissões de localização de dados AWS Lake Formation permitem que as entidades principais criem e alterem recursos do Catálogo de dados que apontam para locais registrados designados do Amazon S3. As permissões de localização de dados funcionam em conjunto com as permissões de dados do Lake Formation para proteger as informações em seu data lake.

O Lake Formation não usa o serviço AWS Resource Access Manager (AWS RAM) para conceder permissões de localização de dados, então você não precisa aceitar convites de compartilhamento de recursos para obter permissões de localização de dados.

Você pode conceder permissões de localização de dados usando o console do Lake Formation, a API ou AWS Command Line Interface (AWS CLI).

**Note**

Para que uma concessão seja bem-sucedida, você deve primeiro registrar a localização dos dados no Lake Formation.

**Consulte também:**

- [Underlying data access control](#)

**Tópicos**

- [Concessão de permissões de localização de dados \(mesma conta\)](#)
- [Concessão de permissões de localização de dados \(conta externa\)](#)
- [Conceder permissões em um local de dados compartilhado com sua conta](#)

## Concessão de permissões de localização de dados (mesma conta)

Siga estas etapas para conceder permissões de localização de dados às entidades principais da sua conta AWS. Você pode conceder permissões usando o console do Lake Formation, a API ou o AWS Command Line Interface (AWS CLI).

Para conceder permissões de localização de dados (mesma conta, console)

1. Abra o console do AWS Lake Formation em <https://console.aws.amazon.com/lakeformation/>. Faça login como administrador do data lake ou como entidade principal que concedeu permissões no local de dados desejado.
2. No painel de navegação, escolha Locais de dados.
3. Selecione Conceder.
4. Na caixa de diálogo Conceder permissões, verifique se o quadro Minha conta está selecionado. Em seguida, forneça as seguintes informações:
  - Para usuários e funções do IAM, escolha um ou mais entidades principais.
  - Para usuários e grupos do SAML e do Amazon QuickSight, insira um ou mais nomes dos recursos da Amazon (ARNs) para usuários ou grupos federados por meio do SAML ou ARNs para usuários ou grupos do Amazon QuickSight.

Insira um ARN por vez e pressione Enter após cada ARN. Para obter informações sobre como construir os ARNs, consulte [Lake Formation concede e revoga comandos AWS CLI](#).

- Para Locais de armazenamento, escolha Browse e escolha um local de armazenamento do Amazon Simple Storage Service (Amazon S3). O local deve ser registrado no Lake Formation. Escolha Procurar novamente para adicionar outro local. Você também pode digitar o local, mas certifique-se de preceder o local com `s3://`.
- Em Local da conta registrada, insira o ID da conta AWS em que o local está registrado. O padrão é o ID da sua conta. Em um cenário de várias contas, os administradores de data lake em uma conta de destinatário podem especificar a conta do proprietário aqui ao conceder a permissão de localização de dados a outras entidades principais na conta do destinatário.
- (Opcional) Para permitir que os diretores selecionados concedam permissões de localização de dados no local selecionado, selecione Concedível.

**Grant permissions** ✕

Add access permissions for specific storage locations.

**My account**  
User or role from this AWS account.

**External account**  
AWS account or AWS organization outside of my account.

**IAM users and roles**  
Add one or more IAM users or roles.

Choose IAM principals to add ▼

**datalake\_user** ✕  
User

**SAML and Amazon QuickSight users and groups**  
Enter a SAML user or group ARN or Amazon QuickSight ARN. Press Enter to add additional ARNs.

Ex: `arn:aws:iam:<AccountId>:saml-provider/<SamlProviderName>`

**Storage locations**  
Choose one or more data lake locations.

`s3://retail/transactions/2020q1` Browse

**Registered account location**  
The account where this storage location is registered in AWS Lake Formation.

`123456789012`

Grantable

Cancel Grant

## 5. Selecione Conceder.



## Para conceder permissões de localização de dados (mesma conta, AWS CLI)

- Execute um comando `grant-permissions` e conceda `DATA_LOCATION_ACCESS` à entidade principal, especificando o caminho do Amazon S3 como o recurso.

### Example

O exemplo a seguir concede permissões de localização de dados em `s3://retail` ao usuário `datalake_user1`.

```
aws lakeformation grant-permissions --principal
DataLakePrincipalIdentifier=arn:aws:iam::<account-id>:user/datalake_user1
--permissions "DATA_LOCATION_ACCESS" --resource '{ "DataLocation":
{"ResourceArn":"arn:aws:s3:::retail"} }'
```

### Example

O exemplo a seguir concede permissões de localização de dados em `s3://retail` ao grupo `ALLIAMPrincipals`.

```
aws lakeformation grant-permissions --principal
DataLakePrincipalIdentifier=111122223333:IAMPrincipals --
permissions "DATA_LOCATION_ACCESS" --resource '{ "DataLocation":
{"ResourceArn":"arn:aws:s3:::retail", "CatalogId": "111122223333"} }'
```

 Consulte também:

- [Referência de permissões do Lake Formation](#)

## Concessão de permissões de localização de dados (conta externa)

Siga estas etapas para conceder permissões de localização de dados a uma conta AWS ou organização externa.

Você pode conceder permissões usando o console do Lake Formation, a API ou o AWS Command Line Interface (AWS CLI).

### Antes de começar

Certifique-se de que todos os pré-requisitos de acesso entre contas sejam atendidos. Para obter mais informações, consulte [Pré-requisitos](#).

Para conceder permissões de localização de dados (conta externa, console)

1. Abra o console do AWS Lake Formation em <https://console.aws.amazon.com/lakeformation/>. Faça login como administrador do data lake.
2. No painel de navegação, escolha Locais de dados e escolha Conceder.
3. Na caixa de diálogo Conceder permissões, escolha o quadro Conta externa.
4. Forneça as informações a seguir:
  - Para ID da conta AWS ou ID da organização AWS, insira números de conta AWS, IDs de organização ou IDs de unidade organizacional válidos.

Pressione Enter após cada ID.

O ID da organização consiste em "o-" seguido por 10 a 32 letras minúsculas ou dígitos.

Um ID de unidade organizacional consiste em "ou-" seguido de 4 a 32 letras minúsculas ou dígitos (o ID da raiz que contém a UO). Essa string é seguida por um segundo "-" (hífen) e 8 a 32 letras minúsculas ou dígitos adicionais.

- Em Locais de armazenamento, escolha Procurar e escolha um local de armazenamento do Amazon Simple Storage Service (Amazon S3). O local deve ser registrado no Lake Formation.

5. Selecione Concedível.
6. Selecione Conceder.

Para conceder permissões de localização de dados (conta externa, AWS CLI)

- Para conceder permissões a uma conta externa AWS, digite um comando semelhante ao seguinte:

```
aws lakeformation grant-permissions --principal
  DataLakePrincipalIdentifier=111122223333 --permissions "DATA_LOCATION_ACCESS"
  --permissions-with-grant-option "DATA_LOCATION_ACCESS" --resource
  '{ "DataLocation": {"CatalogId":"123456789012","ResourceArn":"arn:aws:s3::retail/
  transactions/2020q1"}}'
```

Esse comando concede a `DATA_LOCATION_ACCESS` a opção de concessão à conta 1111-2222-3333 no local `s3://retail/transactions/2020q1` do Amazon S3, que pertence à conta 1234-5678-9012.

Para conceder permissões a uma organização, digite um comando semelhante ao seguinte:

```
aws lakeformation grant-permissions --principal
  DataLakePrincipalIdentifier=arn:aws:organizations::111122223333:organization/
```

```
o-abcdefghijkl --permissions "DATA_LOCATION_ACCESS" --permissions-  
with-grant-option "DATA_LOCATION_ACCESS" --resource '{"DataLocation":  
  {"CatalogId":"123456789012","ResourceArn":"arn:aws:s3::retail/  
transactions/2020q1"}}'
```

Este comando concede a DATA\_LOCATION\_ACCESS a opção de concessão à organização o-abcdefghijkl no local do Amazon S3 s3://retail/transactions/2020q1, que pertence à conta 1234-5678-9012.

Para conceder permissões a uma entidade principal em uma conta externa AWS, digite um comando semelhante ao seguinte.

```
aws lakeformation grant-permissions --principal  
DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1  
--permissions "DATA_LOCATION_ACCESS" --resource '{ "DataLocation":  
{"ResourceArn":"arn:aws:s3::retail/transactions/2020q1", "CatalogId":  
"123456789012"}}'
```

Esse comando é concedido DATA\_LOCATION\_ACCESS a uma entidade principal na conta 1111-2222-3333 na localização s3://retail/transactions/2020q1 do Amazon S3, que pertence à conta 1234-5678-9012.

## Example

O exemplo a seguir concede permissões de localização de dados a s3://retail para um grupo ALLIAMPrincipals em uma conta externa.

```
aws lakeformation grant-permissions --principal  
DataLakePrincipalIdentifier=111122223333:IAMPrincipals --  
permissions "DATA_LOCATION_ACCESS" --resource '{ "DataLocation":  
{"ResourceArn":"arn:aws:s3::retail", "CatalogId": "123456789012"}}'
```

### Consulte também:

- [Referência de permissões do Lake Formation](#)

## Conceder permissões em um local de dados compartilhado com sua conta

Depois que um recurso do Catálogo de Dados é compartilhado com sua conta AWS, como administrador do data lake, você pode conceder permissões sobre o recurso a outras entidades principais da sua conta. Se a permissão ALTER for concedida em uma tabela compartilhada e a tabela apontar para um local registrado no Amazon S3, você também deverá conceder permissões de localização de dados no local. Da mesma forma, se a permissão CREATE\_TABLE ou ALTER for concedida em um banco de dados compartilhado e o banco de dados tiver uma propriedade de localização que aponte para um local registrado, você também deverá conceder permissões de localização de dados no local.

Para conceder permissões de localização de dados em um local compartilhado a uma entidade principal em sua conta, sua conta deve ter recebido a permissão DATA\_LOCATION\_ACCESS no local com a opção de concessão. Ao conceder DATA\_LOCATION\_ACCESS a outra entidade principal em sua conta, você deve incluir a ID do Catálogo de dados (ID da conta AWS) da conta do proprietário. A conta do proprietário é a conta que registrou o local.

Você pode usar o console AWS Lake Formation, a API ou o AWS Command Line Interface (AWS CLI) para conceder permissões de localização de dados.

Para conceder permissões em um local de dados compartilhado com sua conta (console)

- Siga as etapas em [Concessão de permissões de localização de dados \(mesma conta\)](#).

Para Locais de armazenamento, você deve digitar os locais. Em Localização da conta registrada, insira o ID da conta AWS do proprietário.

Para conceder permissões em um local de dados compartilhado com sua conta (AWS CLI)

- Digite um dos comandos a seguir para conceder permissões a um usuário ou a uma função.

```
aws lakeformation grant-permissions --principal
DataLakePrincipalIdentifier=arn:aws:iam::<account-id>:user/<user-name>
--permissions "DATA_LOCATION_ACCESS" --resource '{ "DataLocation":
{"CatalogId":"<owner-account-ID>","ResourceArn":"arn:aws:s3:::<s3-location>"}}'
aws lakeformation grant-permissions --principal
DataLakePrincipalIdentifier=arn:aws:iam::<account-id>:role/<role-name>
--permissions "DATA_LOCATION_ACCESS" --resource '{ "DataLocation":
{"CatalogId":"<owner-account-ID>","ResourceArn":"arn:aws:s3:::<s3-location>"}}'
```

# Concedendo e revogando permissões nos recursos do Catálogo de dados

É possível conceder Permissões de data lake a entidades principais no AWS Lake Formation para que elas possam criar e gerenciar recursos do catálogo de dados e acessar os dados subjacentes. É possível conceder Permissões de data lake em bancos de dados, tabelas e visualizações. Ao conceder permissões em tabelas, é possível limitar o acesso a colunas ou linhas específicas da tabela para um controle de acesso ainda mais refinado.

É possível conceder permissões em tabelas e visualizações individuais ou, com uma única operação de concessão, você pode conceder permissões em todas as tabelas e visualizações em um banco de dados. Se você conceder permissões em todas as tabelas em um banco de dados, estará concedendo implicitamente a permissão DESCRIBE no banco de dados. Em seguida, o banco de dados aparece na página Bancos de dados no console e é retornado pela operação da API `GetDatabases`.

Você pode conceder permissões usando o método de recurso nomeado ou o método de controle de acesso baseado em tags do Lake Formation (LF-TBAC).

É possível conceder permissões a entidades principais na mesma Conta da AWS ou a contas ou organizações externas. Quando você concede para contas ou organizações externas, você está compartilhando recursos que possui com essas contas ou organizações. As entidades principais dessas contas ou organizações podem então acessar os recursos do Catálogo de dados que você possui e os dados subjacentes.

## Note

No momento, o método LF-TBAC aceita a concessão de permissões entre contas a entidades principais, Contas da AWS, organizações e unidades organizacionais (UOs) do IAM.

Ao conceder permissões a contas ou organizações externas, você deve incluir a opção de concessão. Somente o administrador do data lake na conta externa pode acessar os recursos compartilhados até que o administrador conceda permissões sobre os recursos compartilhados a outras entidades principais na conta externa.

Você pode conceder permissões do Catálogo de dados usando o console AWS Lake Formation, a API ou o AWS Command Line Interface (AWS CLI).

### Note

Ao excluir um recurso do catálogo de dados, todas as permissões associadas ao recurso se tornam inválidas. Recriar o mesmo recurso com o mesmo nome não recuperará as permissões do Lake Formation. Os usuários precisarão configurar novas permissões novamente.

### Consulte também:

- [Compartilhamento de tabelas e bancos de dados do Catálogo de dados entre contas AWS](#)
- [Controle de acesso a metadados](#)
- [Referência de permissões do Lake Formation](#)

## Permissões do IAM necessárias para conceder ou revogar as permissões do Lake Formation

Todos as entidades principais, incluindo o administrador do data lake, precisam das seguintes permissões AWS Identity and Access Management (IAM) para conceder ou revogar as permissões do catálogo de dados AWS Lake Formation ou as permissões de localização de dados com a API Lake Formation ou com AWS CLI:

- `lakeformation:GrantPermissions`
- `lakeformation:BatchGrantPermissions`
- `lakeformation:RevokePermissions`
- `lakeformation:BatchRevokePermissions`
- `glue:GetTable` ou `glue:GetDatabase` para uma tabela ou um banco de dados para o qual você está concedendo permissões com o método de recurso nomeado.

**Note**

Os administradores do data lake possuem permissões implícitas do Lake Formation para conceder e revogar permissões do Lake Formation. Mas eles ainda precisam das permissões do IAM nas operações de concessão e revogação da API do Lake Formation.

Os perfis do IAM com política `AWSLakeFormationDataAdmin` gerenciada por AWS não podem adicionar novos administradores de data lake porque essa política contém uma negação explícita para a operação da API Lake Formation, `PutDataLakeSetting`.

A política do IAM a seguir é recomendada para entidades principais que não são administradores de data lake e que desejam conceder ou revogar permissões usando o console do Lake Formation.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "lakeformation:ListPermissions",
        "lakeformation:GrantPermissions",
        "lakeformation:BatchGrantPermissions",
        "lakeformation:RevokePermissions",
        "lakeformation:BatchRevokePermissions",
        "glue:GetDatabases",
        "glue:SearchTables",
        "glue:GetTables",
        "glue:GetDatabase",
        "glue:GetTable",
        "iam:ListUsers",
        "iam:ListRoles",
        "sso-directory:DescribeUser",
        "sso-directory:DescribeGroup",
        "sso:DescribeInstance"
      ],
      "Resource": "*"
    }
  ]
}
```



Todas as permissões `glue:` e `iam:` desta política estão disponíveis na política AWS gerenciada por `AWSGlueConsoleFullAccess`.

Para conceder permissões usando o controle de acesso baseado em tags do Lake Formation (LF-TBAC), as entidades principais precisam de permissões adicionais do IAM. Para obter mais informações, consulte [Considerações e práticas recomendadas de controle de acesso com base em tags do Lake Formation](#) e [Referência de personas e permissões do IAM do Lake Formation](#).

## Permissões do entre contas

Os usuários que desejam conceder permissões entre contas do Lake Formation usando o método de recurso nomeado também devem ter as permissões na política `AWSLakeFormationCrossAccountManager` gerenciada por AWS.

Os administradores do data lake precisam dessas mesmas permissões para conceder permissões entre contas, além da permissão AWS Resource Access Manager (AWS RAM) para permitir a concessão de permissões às organizações. Para obter mais informações, consulte [Permissões de administrador do data lake](#).

## O usuário administrador

Uma entidade principal com permissões administrativas – por exemplo, com a política `AdministratorAccess` gerenciada por AWS – tem permissões para conceder permissões do Lake Formation e criar administradores de data lake. Para negar a um usuário ou função o acesso às operações de administrador do Lake Formation, anexe ou adicione à política uma declaração `Deny` para as operações da API do administrador.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Action": [
        "lakeformation:GetDataLakeSettings",
        "lakeformation:PutDataLakeSettings"
      ],
      "Effect": "Deny",
      "Resource": [
        "*"
      ]
    }
  ]
}
```

```
}
```

### Important

Para evitar que os usuários se adicionem como administradores com um script de extração, transformação e carregamento (ETL), certifique-se de que todos os usuários e funções não administradores tenham acesso negado a essas operações de API. A política `AWSLakeFormationDataAdmin` gerenciada por AWS contém uma negação explícita da operação da API do Lake Formation, `PutDataLakeSetting` que impede que os usuários adicionem novos administradores de data lake.

## Conceder permissões de data lake usando o método de recurso nomeado

É possível usar o método de recurso nomeado para conceder permissões do Lake Formation em tabelas, visualizações e bancos de dados específicos do catálogo de dados. Você pode conceder permissões usando o console AWS Lake Formation, a API ou o AWS Command Line Interface (AWS CLI).

### Tópicos

- [Conceder permissões de banco de dados usando o método de recurso nomeado](#)
- [Conceder permissões de tabela usando o método de recurso nomeado](#)
- [Conceder permissões em visualizações usando o método de recurso nomeado](#)

## Conceder permissões de banco de dados usando o método de recurso nomeado


As etapas a seguir explicam como conceder permissões de banco de dados usando o método de recurso nomeado.

### Console

Use a página Conceder permissões de data lake no console do Lake Formation. A página está dividida nas seguintes seções:

- Entidades principais: usuários e perfis do IAM, usuários e grupos do Centro de Identidade do IAM, usuários e grupos do SAML, contas da AWS, organizações ou unidades organizacionais aos quais conceder permissões.

- Tags do LF ou recursos do catálogo: bancos de dados, tabelas, visualizações ou links de recursos nos quais conceder permissões.
- Permissões – As permissões do Lake Formation devem ser concedidas.

 Note


Para conceder permissões em um link de recurso de banco de dados, consulte [Como conceder permissões de links de recursos](#).

1. Abra a página Conceder permissões de data lake.

Abra o console AWS Lake Formation em <https://console.aws.amazon.com/lakeformation/> e faça login como administrador do data lake, criador do banco de dados ou usuário do IAM com permissões concedidas no banco de dados.

Execute um destes procedimentos:

- No painel de navegação, em Permissões, escolha Permissões do data lake. Em seguida, escolha Conceder.
- No painel de navegação, selecione Bancos de dados, em Catálogo de dados. Depois, na página Bancos de dados, selecione um banco de dados e, no menu Ações, em Permissões, escolha Conceder.

 Note

Você pode conceder permissões em um banco de dados por meio de seu link de recurso. Para fazer isso, na página Bancos de dados, escolha um link de recurso e, no menu Ações, escolha Conceder no destino. Para obter mais informações, consulte [Como os links de recursos funcionam no Lake Formation](#).

2. Na seção Entidades principais, selecione um tipo de entidade principal e especifique as entidades principais às quais conceder permissões.

[AWS Lake Formation](#) > Grant permissions

## Grant data lake permissions

### Principals

Choose the principals to grant permissions.

**IAM users and roles**  
Users or roles from this AWS account.

**IAM Identity Center - new**  
Users and groups configured in IAM Identity Center.

**SAML users and groups**  
SAML users and group or QuickSight ARNs.

**External accounts**  
AWS account, AWS organization or IAM principal outside of this account

### Users and groups (3)

Choose users and groups to grant permissions.

Remove

Add

< 1 > ⚙

<input type="checkbox"/>	Name <a href="#">↗</a>	Type
<input type="checkbox"/>	<a href="#">DataStewards</a>	Group
<input type="checkbox"/>	<a href="#">user1</a>	User
<input type="checkbox"/>	<a href="#">user2</a>	User

### Usuários e perfis do IAM

Escolha um ou mais usuários ou perfis na lista de usuários e perfis do IAM.


### IAM Identity Center

Selecione um ou mais usuários ou grupos na lista Usuários e grupos. Selecione Adicionar para adicionar mais usuários ou grupos.

### Usuários e grupos SAML

Para usuários e grupos do SAML e do Amazon QuickSight, insira um ou mais nomes dos recursos da Amazon (ARNs) para usuários ou grupos federados por meio do SAML ou ARNs para usuários ou grupos do Amazon QuickSight. Pressione Enter após cada ARN.

Para obter informações sobre como construir os ARNs, consulte [Lake Formation concede e revoga comandos AWS CLI](#).

 Note

A integração do Lake Formation com o Amazon QuickSight é compatível somente com o Amazon QuickSight Enterprise Edition.

## Contas externas

Em Conta da AWS, AWS Organization ou Entidade principal do IAM, insira um ou mais IDs de contas da AWS, IDs de organizações, IDs de unidades organizacionais ou o ARN do usuário ou do perfil do IAM. Pressione Enter após cada ID.

O ID da organização consiste em "o-" seguido por 10 a 32 letras minúsculas ou dígitos.

Uma ID de unidade organizacional começa com "ou-" seguida de 4 a 32 letras minúsculas ou dígitos (o ID da raiz que contém a UO). Essa sequência é seguida por um segundo travessão "-" e 8 a 32 letras minúsculas ou dígitos adicionais.

3. Na seção Tags do LF ou recursos do catálogo, selecione Recursos do catálogo de dados nomeados.

### LF-Tags or catalog resources

Resources matched by LF-Tags (recommended)  
Manage permissions indirectly for resources or data matched by a specific set of LF-Tags.

Named data catalog resources  
Manager permissions for specific databases or tables, in addition to fine-grained data access.

**Databases**  
Select one or more databases.

Choose databases ▼

Load more

retail ✕

**Tables - optional**  
Select one or more tables.

Choose tables ▼


Load more

4. Escolha um ou mais bancos de dados na lista Banco de dados. Também é possível escolher uma ou mais Tabelas e/ou Filtros de dados.
5. Na seção Permissões, selecione permissões e permissões concedidas. Em Permissões do banco de dados, selecione uma ou mais permissões para conceder.

### Database permissions

Database permissions  
Choose specific access permissions to grant.

<input type="checkbox"/> Create table <input type="checkbox"/> Alter <input type="checkbox"/> Drop  <input type="checkbox"/> Describe	<input type="checkbox"/> Super <small>This permission is the union of all the individual permissions to the left, and supersedes them.</small>
<p style="margin: 0;">Grantable permissions Choose the permission that may be granted to others.</p> <input type="checkbox"/> Create table <input type="checkbox"/> Alter <input type="checkbox"/> Drop  <input type="checkbox"/> Describe	<input type="checkbox"/> Super <small>This permission allows the principal to grant any of the permissions to the left, and supersedes those grantable permissions.</small>

 Note

Depois de conceder Create Table ou Alter em um banco de dados que tenha uma propriedade de localização que aponta para um local registrado, certifique-se também de conceder permissões de localização de dados no local às entidades principais. Para obter mais informações, consulte [Concedendo permissões de localização de dados](#).

6. (Opcional) Em Permissões concedidas, selecione as permissões que o beneficiário da concessão pode conceder a outras entidades principais em sua conta AWS. Essa opção não é compatível quando você está concedendo permissões a uma entidade principal do IAM a partir de uma conta externa.
7. Selecione Conceder.

## AWS CLI

Você pode conceder permissões de banco de dados usando o método de recurso nomeado e o AWS Command Line Interface (AWS CLI).

## Para conceder permissões de banco de dados usando o AWS CLI

- Execute um comando `grant-permissions` e especifique um banco de dados ou o Catálogo de dados como recurso, dependendo da permissão concedida.

Nos exemplos a seguir, substitua `<account-id>` por um ID de conta AWS válido.

### Example – Concessão para criar um banco de dados

Este exemplo concede `CREATE_DATABASE` ao usuário `datalake_user1`. Como o recurso no qual essa permissão é concedida é o Catálogo de dados, o comando especifica uma estrutura `CatalogResource` vazia como parâmetro `resource`.

```
aws lakeformation grant-permissions --principal
  DataLakePrincipalIdentifier=arn:aws:iam::<account-id>:user/datalake_user1 --
permissions "CREATE_DATABASE" --resource '{ "Catalog": {} }'
```

### Example – Concessão para criar tabelas em um banco de dados designado

O próximo exemplo concede `CREATE_TABLE` no banco de dados `retail` ao usuário `datalake_user1`.

```
aws lakeformation grant-permissions --principal
  DataLakePrincipalIdentifier=arn:aws:iam::<account-id>:user/datalake_user1 --
permissions "CREATE_TABLE" --resource '{ "Database": {"Name": "retail"} }'
```

### Example – Conceder para uma conta externa AWS com a opção Conceder

O próximo exemplo concede `CREATE_TABLE` com a opção de concessão no banco de dados à conta externa `retail 1111-2222-3333`.

```
aws lakeformation grant-permissions --principal
  DataLakePrincipalIdentifier=111122223333 --permissions "CREATE_TABLE"
--permissions-with-grant-option "CREATE_TABLE" --resource '{ "Database":
{"Name": "retail"} }'
```

### Example – Concessão a uma organização

O próximo exemplo concede a `ALTER` a opção de concessão no banco de dados `issues` à organização `o-abcdefghijkl`.

```
aws lakeformation grant-permissions --principal
DataLakePrincipalIdentifier=arn:aws:organizations::111122223333:organization/
o-abcdefghijkl --permissions "ALTER" --permissions-with-grant-option "ALTER" --
resource '{ "Database": {"Name":"issues"}}'
```

### Example - Conceder **ALLIAMPrincipals** na mesma conta

O próximo exemplo concede permissão `CREATE_TABLE` no banco de dados `retail` a todas as entidades principais na mesma conta. Essa opção permite que cada entidade principal da conta crie uma tabela no banco de dados e crie um link de recurso de tabela, permitindo que mecanismos de consulta integrados acessem bancos de dados e tabelas compartilhados. Essa opção é especialmente útil quando uma entidade principal recebe uma concessão entre contas e não tem permissão para criar links de recursos. Nesse cenário, o administrador do data lake pode criar um banco de dados de espaço reservado e conceder permissão `CREATE_TABLE` ao grupo `ALLIAMPrincipal`, permitindo que cada entidade principal do IAM na conta crie links de recursos no banco de dados de espaço reservado.

```
aws lakeformation grant-permissions --principal
DataLakePrincipalIdentifier=111122223333:IAMPrincipals
--permissions "CREATE_TABLE" --resource '{ "Database":
{"Name":"temp","CatalogId":"111122223333"}}'
```

### Example - Conceder a **ALLIAMPrincipals** em uma conta externa

O próximo exemplo concede `CREATE_TABLE` no banco de dados `retail` a todas as entidades principais em uma conta externa. Essa opção permite que cada entidade principal da conta crie uma tabela no banco de dados.

```
aws lakeformation grant-permissions --principal
DataLakePrincipalIdentifier=111122223333:IAMPrincipals
--permissions "CREATE_TABLE" --resource '{ "Database":
{"Name":"retail","CatalogId":"123456789012"}}'
```

#### Note

Depois de conceder `CREATE_TABLE` ou `ALTER` em um banco de dados que tenha uma propriedade de localização que aponta para um local registrado, certifique-se também de



conceder permissões de localização de dados no local às entidades principais. Para obter mais informações, consulte [Concedendo permissões de localização de dados](#).

### Consulte também

- [Referência de permissões do Lake Formation](#)
- [Conceder permissões em um banco de dados ou tabela compartilhada com sua conta](#)
- [Acessar e visualizar tabelas e bancos de dados compartilhados do Catálogo de Dados](#)

## Conceder permissões de tabela usando o método de recurso nomeado

É possível usar o console do Lake Formation ou a AWS CLI para conceder permissões do Lake Formation nas tabelas do catálogo de dados. Você pode conceder permissões em tabelas individuais ou, com uma única operação de concessão, você pode conceder permissões em todas as tabelas em um banco de dados.

Se você conceder permissões em todas as tabelas em um banco de dados, estará concedendo implicitamente a permissão DESCRIBE no banco de dados. Em seguida, o banco de dados aparece na página Bancos de dados no console e é retornado pela operação da API GetDatabases.

Ao escolher SELECT como a permissão a ser concedida, você tem a opção de aplicar um filtro de coluna, filtro de linha ou filtro de célula.

### Console

As etapas a seguir explicam como conceder permissões de tabela usando o método de recurso nomeado e a página Conceder permissões de data lake no console do Lake Formation. A página está dividida nas seguintes seções:

- Entidades principais: os usuários, perfis, contas AWS, organizações ou unidades organizacionais a quem conceder permissões.
- LF-Tags ou recursos do catálogo – Os bancos de dados, tabelas ou links de recursos nos quais conceder permissões.
- Permissões – As permissões do Lake Formation devem ser concedidas.

**Note**

Para conceder permissões em um link de recurso de tabela, consulte [Como conceder permissões de links de recursos](#).

1. Abra a página Conceder permissões de data lake.

Abra o console AWS Lake Formation em <https://console.aws.amazon.com/lakeformation/> e faça login como administrador do data lake, criador da tabela ou usuário que tenha recebido permissões na tabela com a opção de concessão.

Execute um destes procedimentos:

- No painel de navegação, selecione Permissões de data lake, em Permissões. Em seguida, escolha Conceder.
- No painel de navegação, selecione Tabelas. Em seguida, na página Tabelas, escolha uma tabela e, no menu Ações, em Permissões, escolha Conceder.

**Note**

Você pode conceder permissões em uma tabela por meio de seu link de recurso. Para fazer isso, na página Tabelas, escolha um link de recurso e, no menu Ações, escolha Conceder no destino. Para obter mais informações, consulte [Como os links de recursos funcionam no Lake Formation](#).

2. Na seção Entidades principais, selecione um tipo de entidade principal e especifique as entidades principais às quais conceder permissões.

[AWS Lake Formation](#) > Grant permissions

## Grant data lake permissions

### Principals

Choose the principals to grant permissions.

<input type="radio"/> <b>IAM users and roles</b> Users or roles from this AWS account.	<input checked="" type="radio"/> <b>IAM Identity Center - new</b> Users and groups configured in IAM Identity Center.	<input type="radio"/> <b>SAML users and groups</b> SAML users and group or QuickSight ARNs.	<input type="radio"/> <b>External accounts</b> AWS account, AWS organization or IAM principal outside of this account
---	--	--	--

### Users and groups (3)

Choose users and groups to grant permissions.

Remove

Add

&lt;

1

&gt;



<input type="checkbox"/>	Name <a href="#">↗</a>	Type
<input type="checkbox"/>	<a href="#">DataStewards</a>	Group
<input type="checkbox"/>	<a href="#">user1</a>	User
<input type="checkbox"/>	<a href="#">user2</a>	User

### Usuários e perfis do IAM

Escolha um ou mais usuários ou perfis na lista de usuários e perfis do IAM.

### IAM Identity Center

Selecione um ou mais usuários ou grupos na lista Usuários e grupos.

### Usuários e grupos SAML

Para usuários e grupos do SAML e do Amazon QuickSight, insira um ou mais nomes dos recursos da Amazon (ARNs) para usuários ou grupos federados por meio do SAML ou ARNs para usuários ou grupos do Amazon QuickSight. Pressione Enter após cada ARN.

Para obter informações sobre como construir os ARNs, consulte [Lake Formation concede e revoga comandos AWS CLI](#).

**Note**

A integração do Lake Formation com o Amazon QuickSight é compatível somente com o Amazon QuickSight Enterprise Edition.

## Contas externas

Em Conta da AWS, AWS Organization ou Entidade principal do IAM, insira um ou mais IDs válidos de Conta da AWS, IDs de organizações, IDs de unidades organizacionais ou o ARN do usuário ou do perfil do IAM. Pressione Enter após cada ID.

O ID da organização consiste em "o-" seguido por 10 a 32 letras minúsculas ou dígitos.

Uma ID de unidade organizacional começa com "ou-" seguida de 4 a 32 letras minúsculas ou dígitos (o ID da raiz que contém a UO). Essa sequência é seguida por um segundo caractere "-" e de 8 a 32 letras minúsculas ou dígitos adicionais.

- Na seção LF-Tags ou recursos do catálogo, escolha um banco de dados. Em seguida, escolha uma ou mais tabelas ou Todas as tabelas.

### LF-Tags or catalog resources

Resources matched by LF-Tags (recommended)  
Manage permissions indirectly for resources or data matched by a specific set of LF-Tags.

Named data catalog resources  
Manager permissions for specific databases or tables, in addition to fine-grained data access.

#### Databases

Select one or more databases.

▼

retail ✕

#### Tables - optional

Select one or more tables.

▼

inventory ✕  
No description available

#### 4. Especificar as permissões sem filtragem de dados

Na seção Permissões, selecione as permissões da tabela a serem concedidas e, opcionalmente, selecione as permissões que podem ser concedidas.

### Table and column permissions

**Table permissions**  
Choose specific access permissions to grant.

<input checked="" type="checkbox"/> Alter	<input checked="" type="checkbox"/> Insert	<input type="checkbox"/> Drop	<input type="checkbox"/> Super
<input type="checkbox"/> Delete	<input checked="" type="checkbox"/> Select	<input checked="" type="checkbox"/> Describe	This permission is the union of all the individual permissions to the left, and supersedes them.

**Grantable permissions**  
Choose the permission that may be granted to others.

<input type="checkbox"/> Alter	<input type="checkbox"/> Insert	<input type="checkbox"/> Drop	<input type="checkbox"/> Super
<input type="checkbox"/> Delete	<input type="checkbox"/> Select	<input type="checkbox"/> Describe	This permission allows the principal to grant any of the permissions to the left, and supersedes those grantable permissions.

Se você conceder Selecionar, a seção Permissões de dados aparecerá abaixo da seção Permissões de tabela e coluna, com a opção Acesso a todos os dados selecionada por padrão. Aceite o padrão.

### Data permissions

**All data access**  
Grant access to all data without any restrictions.

**Simple column-based access**  
Grant data access to specific columns only.

**Advanced cell-level filters**  
Grant access to specific columns and/or rows with data filters.

5. Selecione Conceder.
6. Especifique a permissão Selecionar com filtragem de dados

Selecione a permissão Selecionar. Não selecione nenhuma outra permissão.

A seção Permissões de dados aparece abaixo da seção Permissões de tabela e coluna.

7. Execute um destes procedimentos:
  - Aplique somente a filtragem simples de colunas.
    1. Escolha Acesso simples baseado em colunas.

### Table and column permissions

**Table permissions**  
Choose specific access permissions to grant.

<input type="checkbox"/> Alter	<input type="checkbox"/> Insert	<input type="checkbox"/> Drop	<input type="checkbox"/> Super
<input type="checkbox"/> Delete	<input checked="" type="checkbox"/> Select	<input type="checkbox"/> Describe	This permission is the union of all the individual permissions to the left, and supersedes them.

**Grantable permissions**  
Choose the permission that may be granted to others.

<input type="checkbox"/> Alter	<input type="checkbox"/> Insert	<input type="checkbox"/> Drop	<input type="checkbox"/> Super
<input type="checkbox"/> Delete	<input checked="" type="checkbox"/> Select	<input type="checkbox"/> Describe	This permission allows the principal to grant any of the permissions to the left, and supersedes those grantable permissions.

---

### Data permissions

**All data access**  
Grant access to all data without any restrictions.

**Simple column-based access**  
Grant data access to specific columns only.

**Advanced cell-level filters**  
Grant access to specific columns and/or rows with data filters.

**Choose permission filter**  
Choose whether to include or exclude columns.

**Include columns**  
Grant permissions to access specific columns.

**Exclude columns**  
Grant permissions to access all but specific columns.

**Select columns**

Choose one or more columns ▼

**Grantable permissions**  
Choose the permission that may be granted to others.

Select

2. Escolha se deseja incluir ou excluir colunas e, em seguida, escolha as colunas a serem incluídas ou excluídas.

Somente listas de inclusão são suportadas ao conceder permissões a uma conta AWS ou organização externa.

3. (Opcional) Em Permissões concedidas, ative a opção de concessão para a permissão Selecionar.

Se você incluir a opção de concessão, o destinatário da concessão poderá conceder permissões somente nas colunas que você conceder a ele.

**Note**

Você também pode aplicar a filtragem de colunas somente criando um filtro de dados que especifique um filtro de coluna e especifique todas as linhas como filtro de linha. No entanto, isso requer mais etapas.

- Aplique filtragem de coluna, linha ou célula.
  1. Escolha Filtros avançados em nível de célula.

<input type="checkbox"/>	Filter name	Table	Database	Table catalog ID
<input type="checkbox"/>	restrict-pharma	orders	sales	111122223333
<input type="checkbox"/>	no-pharma	orders	sales	111122223333

2. (Opcional) Expanda Visualizar permissões existentes.
3. (Opcional) Escolha Criar novo filtro.
4. (Opcional) Para ver detalhes dos filtros listados, criar novos filtros ou excluir filtros existentes, escolha Gerenciar filtros.

A página Filtros de dados é aberta em uma nova janela do navegador.

Quando terminar de acessar a página Filtros de dados, retorne à página Conceder permissões e, se necessário, atualize a página para ver os novos filtros de dados que você criou.

5. Selecione um ou mais filtros de dados a serem aplicados à concessão.

**Note**

Se não houver filtros de dados na lista, isso significa que nenhum filtro de dados foi criado para a tabela selecionada.

8. Selecione Conceder.

## AWS CLI

Você pode conceder permissões de tabela usando o método de recurso nomeado e o AWS Command Line Interface (AWS CLI).

Para conceder permissões de tabela usando o AWS CLI

- Execute um comando `grant-permissions` e especifique uma tabela como recurso.

Example – Concessão em uma única tabela - sem filtragem

O exemplo a seguir concede SELECT e ALTER ao usuário `datalake_user1` na conta AWS 1111-2222-3333 na tabela `inventory` do banco de dados `retail`.

```
aws lakeformation grant-permissions --principal
DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1 --
permissions "SELECT" "ALTER" --resource '{ "Table": {"DatabaseName":"retail",
"Name":"inventory"}}'
```

**Note**

Se você conceder a permissão ALTER em uma tabela que tem seus dados subjacentes em um local registrado, certifique-se de também conceder permissões de localização de dados no local às entidades principais. Para obter mais informações, consulte [Concedendo permissões de localização de dados](#).

Example – Conceder em todas as tabelas com a opção Conceder - sem filtragem

O próximo exemplo concede SELECT com a opção de concessão em todas as tabelas no banco de dados `retail`.



```
aws lakeformation grant-permissions --principal
DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1 --
permissions "SELECT" --permissions-with-grant-option "SELECT" --resource '{ "Table":
{ "DatabaseName": "retail", "TableWildcard": {} } }'
```

### Example – Grant com filtragem simples de colunas

O próximo exemplo concede a SELECT um subconjunto de colunas na tabela persons. Ele usa filtragem de coluna simples.

```
aws lakeformation grant-permissions --principal
DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1 --
permissions "SELECT" --resource '{ "TableWithColumns": {"DatabaseName": "hr",
"Name": "persons", "ColumnNames": ["family_name", "given_name", "gender"]}}'
```

### Example – Conceda com um filtro de dados

Este exemplo concede SELECT na tabela orders e aplica o filtro de dados restrict-pharma.

```
aws lakeformation grant-permissions --cli-input-json file://grant-params.json
```

A seguir está o conteúdo do arquivo grant-params.json.

```
{
  "Principal": {"DataLakePrincipalIdentifier": "arn:aws:iam::111122223333:user/
datalake_user1"},
  "Resource": {
    "DataCellsFilter": {
      "TableCatalogId": "111122223333",
      "DatabaseName": "sales",
      "TableName": "orders",
      "Name": "restrict-pharma"
    }
  },
  "Permissions": ["SELECT"],
  "PermissionsWithGrantOption": ["SELECT"]
}
```

### Consulte também

- [Visão geral das permissões do Lake Formation](#)
- [Filtragem de dados e segurança em nível de célula no Lake Formation](#)
- [Referência de personas e permissões do IAM do Lake Formation](#)
- [Como conceder permissões de links de recursos](#)
- [Acessar e visualizar tabelas e bancos de dados compartilhados do Catálogo de Dados](#)

## Conceder permissões em visualizações usando o método de recurso nomeado

As etapas a seguir explicam como conceder permissões em visualizações usando o método de recurso nomeado e a página Conceder permissões de data lake. A página está dividida nas seguintes seções:

- Entidades principais: usuários e perfis do IAM, usuários e grupos do Centro de Identidade do IAM, Contas da AWS, organizações ou unidades organizacionais aos quais conceder permissões.
- Tags do LF ou recursos do catálogo: bancos de dados, tabelas, visualizações ou links de recursos nos quais conceder permissões.
- Permissões: as permissões de data lake a serem concedidas.

Abra a página Conceder permissões de data lake

1. Abra o console AWS Lake Formation em <https://console.aws.amazon.com/lakeformation/> e faça login como administrador do data lake, criador do banco de dados ou usuário do IAM com permissões concedidas no banco de dados.
2. Execute um destes procedimentos:
  - No painel de navegação, em Permissões, escolha Permissões do data lake. Em seguida, escolha Conceder.
  - No painel de navegação, selecione Visualizações em Catálogo de dados. Depois, na página Visualizações, selecione uma visualização e, no menu Ações, em Permissões, escolha Conceder.

 Note

É possível conceder permissões em uma visualização por meio do link de recurso. Para fazer isso, na página Visualizações, escolha um link de recurso e, no menu Ações, selecione Conceder no destino. Para obter mais informações, consulte [Como os links de recursos funcionam no Lake Formation](#).

## Especificar a entidade principal

Na seção Entidades principais, escolha um tipo de entidade principal e, em seguida, especifique as entidades principais para conceder permissões.

### Usuários e perfis do IAM

Escolha um ou mais usuários ou perfis na lista de usuários e perfis do IAM.


### IAM Identity Center

Selecione um ou mais usuários ou grupos na lista Usuários e grupos.

### Usuários e grupos SAML

Para usuários e grupos do SAML e do Amazon QuickSight, insira um ou mais nomes dos recursos da Amazon (ARNs) para usuários ou grupos federados por meio do SAML ou ARNs para usuários ou grupos do Amazon QuickSight. Pressione Enter após cada ARN.

Para obter informações sobre como construir os ARNs, consulte [Lake Formation concede e revoga comandos AWS CLI](#).

 Note

A integração do Lake Formation com o Amazon QuickSight é compatível somente com o Amazon QuickSight Enterprise Edition.

## Contas externas

Em Conta da AWS, AWS Organization ou Entidade principal do IAM, insira um ou mais IDs de contas da AWS, IDs de organizações, IDs de unidades organizacionais ou o ARN do usuário ou do perfil do IAM. Pressione Enter após cada ID.

O ID da organização consiste em "o-" seguido por 10 a 32 letras minúsculas ou dígitos.

Uma ID de unidade organizacional começa com "ou-" seguida de 4 a 32 letras minúsculas ou dígitos (o ID da raiz que contém a UO). Essa sequência é seguida por um segundo travessão "-" e 8 a 32 letras minúsculas ou dígitos adicionais.

 Consulte também:

- [Acessar e visualizar tabelas e bancos de dados compartilhados do Catálogo de Dados](#)

## Especificar as visualizações

Na seção Tags do LF ou recursos do catálogo, selecione uma ou mais visualizações nas quais conceder permissões.

1. Escolha Recursos do catálogo de dados nomeado.
2. Selecione uma ou mais visualizações na lista Visualizações. Também é possível escolher um ou mais filtros de bancos de dados, tabelas e/ou dados.

Conceder permissões de data lake a All views em um banco de dados resultará em permissões do favorecido em todas as tabelas e visualizações do banco de dados.

## Especifique as permissões

Na seção Permissões, selecione permissões e permissões concedidas.

## View permissions

View permissions  
Choose specific access permissions to grant.

Select     Describe     Drop

Super  
This permission is the union of all the individual permissions to the left, and supersedes them.

Grantable permissions  
Choose the permission that may be granted to others.

Select     Describe     Drop

Super  
This permission allows the principal to grant any of the permissions to the left, and supersedes those grantable permissions.

Cancel **Grant**

1. Em Visualizar permissões, selecione uma ou mais permissões a serem concedidas.
2. (Opcional) Em Permissões que podem ser concedidas, selecione as permissões que o beneficiário pode conceder a outras entidades principais na Conta da AWS. Essa opção não é compatível quando você está concedendo permissões a uma entidade principal do IAM a partir de uma conta externa.
3. Selecione Conceder.

 Consulte também:

- [Referência de permissões do Lake Formation](#)
- [Conceder permissões em um banco de dados ou tabela compartilhada com sua conta](#)

## Controle de acesso baseado em tags do Lake Formation

O controle de acesso baseado em tags do Lake Formation (LF-TBAC) é uma estratégia de autorização que define permissões com base em atributos. Em Lake Formation, esses atributos são chamados de LF-Tags. Você pode anexar tags LF aos recursos do catálogo de dados e conceder permissões aos diretores do Lake Formation sobre esses recursos usando essas tags LF. O Lake

Formation permite operações nesses recursos quando o valor da tag do principal corresponde ao valor da tag do recurso. O LF-TBAC é útil em ambientes que estão crescendo rapidamente e ajuda em situações em que o gerenciamento de políticas se torna um problema.

O LF-TBAC é o método recomendado para conceder permissões do Lake Formation quando há um grande número de recursos do Catálogo de Dados. O LF-TBAC é mais escalável do que o método de recurso nomeado e requer menos sobrecarga de gerenciamento de permissões.

### Note

As tags IAM não são iguais às LF-Tags. Essas tags não são intercambiáveis. As LF-Tags são usadas para conceder permissões do Lake Formation e as tags IAM são usadas para definir políticas do IAM.

## Como funciona o controle de acesso baseado em tags do Lake Formation

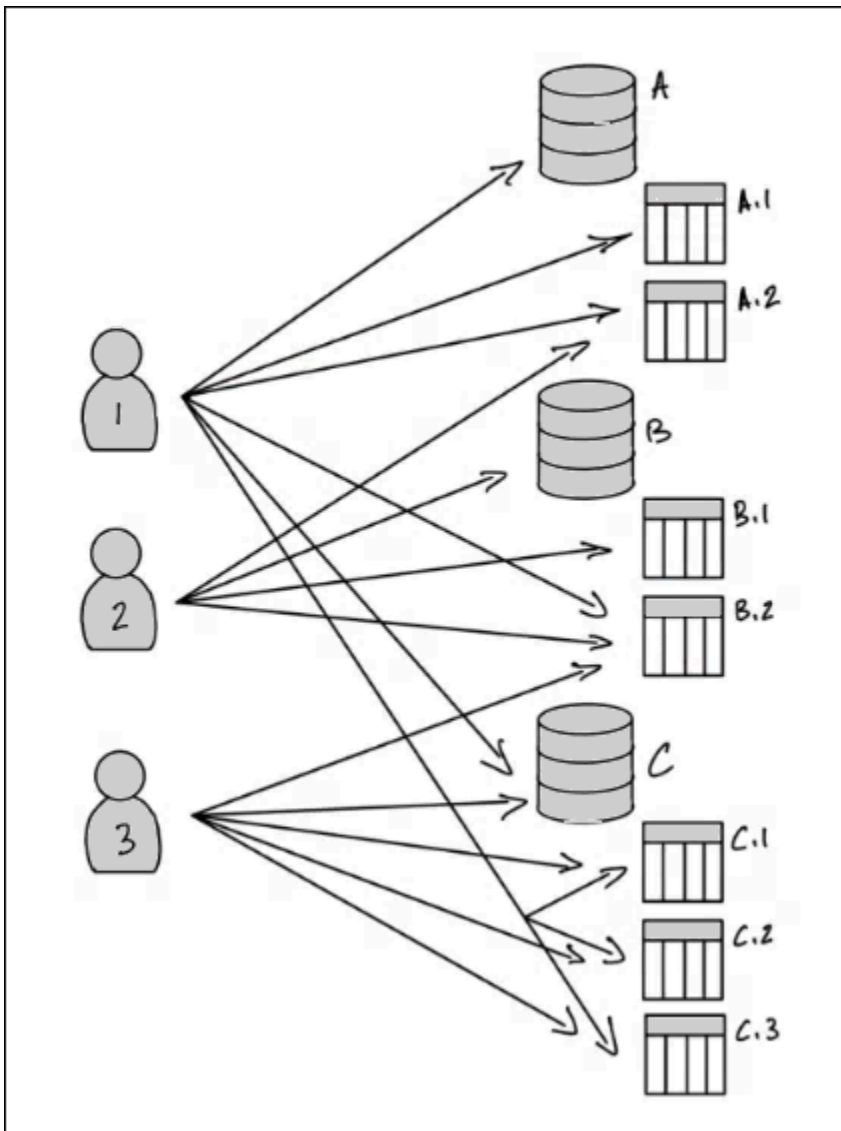
Cada LF-Tag é um par de valores-chave, como `department=sales` ou `classification=restricted`. Uma chave pode ter vários valores definidos, como `department=sales,marketing,engineering,finance`.

Para usar o método LF-TBAC, os administradores e engenheiros de dados do data lake realizam as seguintes tarefas.

Tarefa	Detalhes da tarefa
1. Defina as propriedades e os relacionamentos das LF-Tags.	-
2. Crie os criadores do LF-Tag em Lake Formation.	<a href="#">Adicionando criadores de LF-Tags</a>
3. Crie a LF-TAG em Lake Formation.	<a href="#">Criação de LF-Tags</a>
4. Atribua LF-Tags aos recursos do Catálogo de Dados.	<a href="#">Atribuição de LF-Tags aos recursos do catálogo de dados</a>
5. Conceda permissões a outras entidades principais para atribuir LF-Tags aos	<a href="#">Concedendo, revogando e listando permissões de valor da LF-Tag</a>

Tarefa	Detalhes da tarefa
recursos, opcionalmente com a opção de concessão.	
6. Conceda expressões de LF-Tag às entidades principais, opcionalmente com a opção de concessão.	<a href="#">Conceder permissões de data lake usando o método LF-TBAC</a>
7. (Recomendado) Depois de verificar se as entidades principais têm acesso aos recursos corretos por meio do método LF-TBAC, revogue as permissões concedidas usando o método de recurso nomeado.	-

Considere o caso em que você deve conceder permissões a três entidade principais em três bancos de dados e sete tabelas.



Para obter as permissões indicadas no diagrama anterior usando o método de recurso nomeado, você precisaria fazer 17 concessões, da seguinte forma (em pseudocódigo).

```
GRANT CREATE_TABLE ON Database A TO PRINCIPAL 1
GRANT SELECT, INSERT ON Table A.1 TO PRINCIPAL 1
GRANT SELECT, INSERT ON Table A.2 TO PRINCIPAL 1
GRANT SELECT, INSERT ON Table B.2 TO PRINCIPAL 1
...
GRANT SELECT, INSERT ON Table A.2 TO PRINCIPAL 2
GRANT CREATE_TABLE ON Database B TO PRINCIPAL 2
...
GRANT SELECT, INSERT ON Table C.3 TO PRINCIPAL 3
```

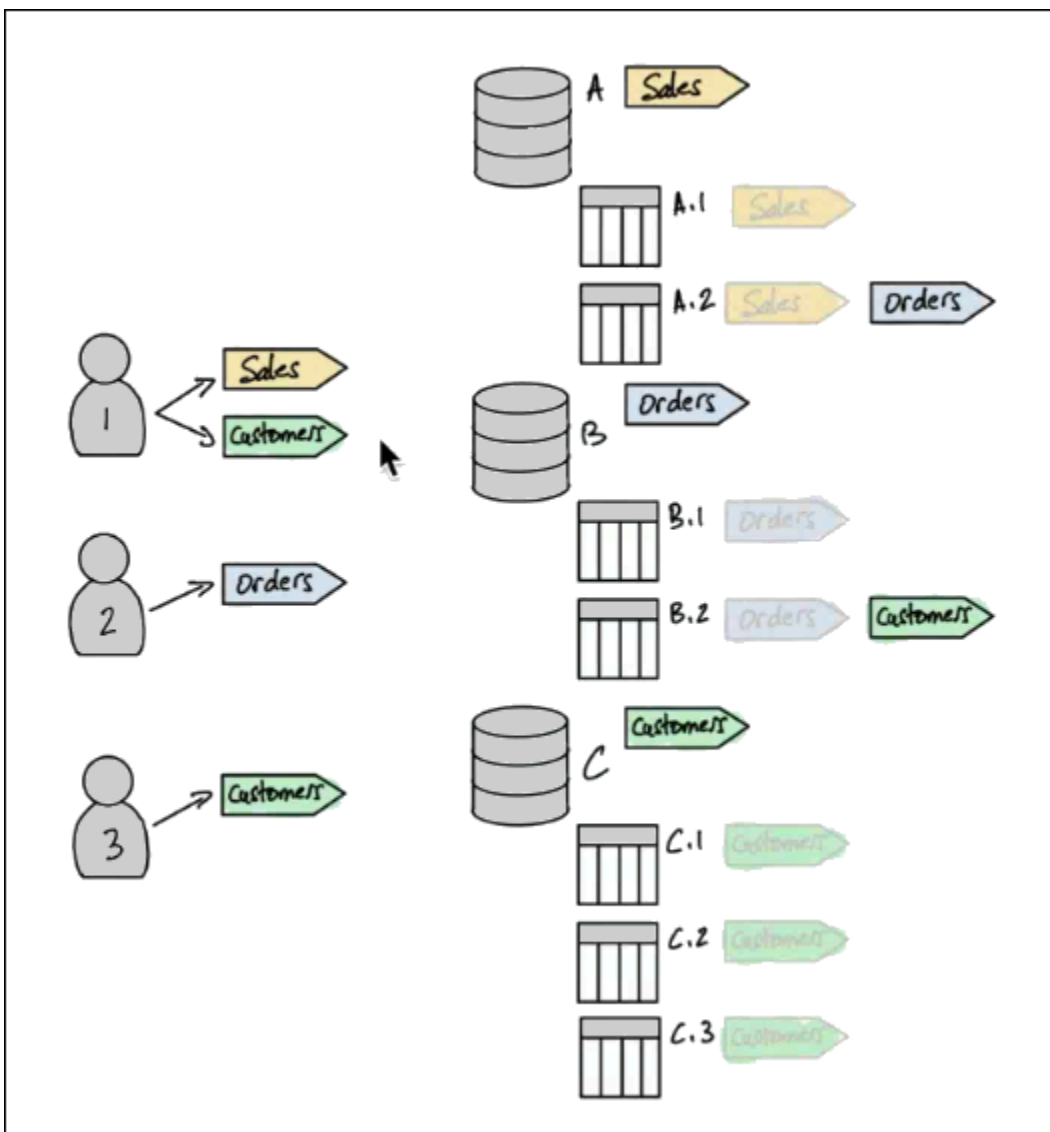


Agora, considere como você concederia permissões usando o LF-TBAC. O diagrama a seguir indica que você atribuiu LF-Tags a bancos de dados e tabelas e concedeu permissões sobre LF-Tags aos principais.

Neste exemplo, as LF-Tags representam áreas do data lake que contêm análises para diferentes módulos de um pacote de aplicativos de planejamento de recursos corporativos (ERP). Você deve controlar o acesso aos dados analíticos dos vários módulos. Todas as LF-Tags têm a chave `module` e os valores possíveis `Sales`, `Orders` e `Customers`. Um exemplo de como uma LF-Tag se parece:

```
module=Sales
```

O diagrama mostra somente os valores da LF-TAG.



## Atribuições de tags aos recursos e herança do Catálogo de Dados

As tabelas herdam as LF-Tags dos bancos de dados e as colunas herdam as LF-Tags das tabelas. Os valores herdados podem ser substituídos. No diagrama anterior, as LF-Tags esmaecidas são herdadas.

Por causa da herança, o administrador do data lake precisa fazer somente as cinco seguintes atribuições de LF-Tag aos recursos (em pseudocódigo).

```
ASSIGN TAGS module=Sales TO database A
ASSIGN TAGS module=Orders TO table A.2
ASSIGN TAGS module=Orders TO database B
ASSIGN TAGS module=Customers TO table B.2
ASSIGN TAGS module=Customers TO database C
```

### Concessões de tags a entidade principais

Depois de atribuir LF-Tags aos bancos de dados e tabelas, o administrador do data lake deve fazer apenas quatro concessões de LF-Tags às entidades principais, da seguinte forma (em pseudocódigo).

```
GRANT TAGS module=Sales TO Principal 1
GRANT TAGS module=Customers TO Principal 1
GRANT TAGS module=Orders TO Principal 2
GRANT TAGS module=Customers TO Principal 3
```

Agora, uma entidade principal com a LF-Tag `module=Sales` pode acessar os recursos do Catálogo de Dados com a LF-Tag `module=Sales` (por exemplo, banco de dados A), uma entidade principal com a LF-Tag `module=Customers` pode acessar recursos com a LF-Tag `module=Customers` e assim por diante.

Os comandos de concessão anteriores estão incompletos. Isso ocorre porque, embora indiquem por meio de LF-Tags os recursos do Catálogo de Dados sobre os quais as entidades principais têm permissões, eles não indicam exatamente quais permissões do Lake Formation (por exemplo `SELECT`, `ALTER`) as entidades principais têm sobre esses recursos. Portanto, os comandos de pseudocódigo a seguir são uma representação mais precisa de como as permissões do Lake Formation são concedidas nos recursos do Catálogo de Dados por meio de LF-Tags.

```
GRANT (CREATE_TABLE ON DATABASES) ON TAGS module=Sales TO Principal 1
GRANT (SELECT, INSERT ON TABLES) ON TAGS module=Sales TO Principal 1
GRANT (CREATE_TABLE ON DATABASES) ON TAGS module=Customers TO Principal 1
```

```
GRANT (SELECT, INSERT ON TABLES) ON TAGS module=Customers TO Principal 1
GRANT (CREATE_TABLE ON DATABASES) ON TAGS module=Orders TO Principal 2
GRANT (SELECT, INSERT ON TABLES) ON TAGS module=Orders TO Principal 2
GRANT (CREATE_TABLE ON DATABASES) ON TAGS module=Customers TO Principal 3
GRANT (SELECT, INSERT ON TABLES) ON TAGS module=Customers TO Principal 3
```

Reunir tudo: permissões resultantes em recursos

Considerando as LF-Tags atribuídas aos bancos de dados e tabelas no diagrama anterior e as LF-Tags concedidas às entidades principais no diagrama, a tabela a seguir lista as permissões do Lake Formation que as entidades principais têm nos bancos de dados e tabelas.

Entidade principal	Permissões concedidas por meio de LF-Tags
Entidade principal 1	<ul style="list-style-type: none"> <li>• CREATE_TABLE no banco de dados A</li> <li>• SELECT, INSERT na tabela A.1</li> <li>• SELECT, INSERT na tabela B.2</li> <li>• CREATE_TABLE no banco de dados C</li> <li>• SELECT, INSERT na tabela C.1</li> <li>• SELECT, INSERT na tabela C.2</li> <li>• SELECT, INSERT na tabela C.3</li> </ul>
Entidade principal 2	<ul style="list-style-type: none"> <li>• SELECT, INSERT na tabela A.2</li> <li>• CREATE_TABLE no banco de dados B</li> <li>• SELECT, INSERT na tabela B.1</li> <li>• SELECT, INSERT na tabela B.2</li> </ul>
Entidade principal 3	<ul style="list-style-type: none"> <li>• SELECT, INSERT na tabela B.2</li> <li>• CREATE_TABLE no banco de dados C</li> <li>• SELECT, INSERT na tabela C.1</li> <li>• SELECT, INSERT na tabela C.2</li> <li>• SELECT, INSERT na tabela C.3</li> </ul>

## Conclusão

Neste exemplo simples, usando cinco operações de atribuição e oito operações de concessão, o administrador do data lake conseguiu especificar 17 permissões. Quando há dezenas de bancos de dados e centenas de tabelas, a vantagem do método LF-TBAC sobre o método de recurso nomeado fica clara. No caso hipotético da necessidade de conceder a todos os principais acesso a todos os recursos, e onde  $n(P)$  é o número de entidades principais e  $n(R)$  o número de recursos:

- Com o método de recurso nomeado, o número de concessões necessárias é  $n(P) \times n(R)$ .
- Com o método LF-TBAC, usando uma única LF-TAG, o total do número de concessões para entidades principais e atribuições de recursos é  $n(P) + n(R)$ .

#### Consulte também

- [Gerenciar LF-Tags para controle de acesso a metadados](#)
- [Conceder permissões de data lake usando o método LF-TBAC](#)

#### Tópicos

- [Gerenciar LF-Tags para controle de acesso a metadados](#)
- [Concedendo, revogando e listando permissões de valor da LF-Tag](#)

## Gerenciar LF-Tags para controle de acesso a metadados

Para usar o método de controle de acesso baseado em tags do Lake Formation (LF-TBAC) para proteger os recursos do Catálogo de Dados (bancos de dados, tabelas e colunas), você cria LF-Tags, as atribui aos recursos e concede permissões de LF-Tag a entidades principais.

Antes de atribuir LF-Tags aos recursos do catálogo de dados ou conceder permissões a entidades principais, você precisa definir LF-Tags. Somente um administrador de data lake ou uma entidade principal com permissões de criador de LF-Tags pode criar LF-Tags.

### Criadores de LF-Tags

O criador do LF-Tag é uma entidade principal não administradora que tem permissões para criar e gerenciar LF-Tags. Os administradores do Data Lake podem adicionar criadores de LF-Tags usando a CLI ou o console do Lake Formation. Os criadores de LF-Tags têm permissões implícitas do Lake Formation para atualizar e excluir LF-Tags, atribuir LF-Tags a recursos e conceder permissões de LF-Tag e permissões de valor de LF-Tag a outras entidades principais.

Com as funções de criador de LF-Tags, os administradores do data lake podem delegar tarefas de gerenciamento de tags, como criar e atualizar valores e chaves de tags, a entidades principais que não são administradoras. Os administradores do Data Lake também podem conceder aos criadores de LF-Tag permissões `Create LF-Tag` concedíveis. Em seguida, o criador da LF-Tag pode conceder a permissão para criar LF-Tags a outras entidades principais.

Você pode conceder dois tipos de permissões nas LF-Tags:

- Permissões LF-Tag: `Create LF-Tag`, `Alter` e `Drop`. Essas permissões são necessárias para criar, atualizar e excluir LF-Tags.

Os administradores do data lake e os criadores de LF-Tags têm implicitamente essas permissões nas LF-Tags que criam e podem concedê-las explicitamente às entidades principais para gerenciar tags no data lake.

- Permissões do par chave-valor do LF-Tag: `Assign`, `Describe` e `Grant with LF-Tag expressions`. Essas permissões são necessárias para atribuir LF-Tags a bancos de dados, tabelas e colunas do Catálogo de Dados e para conceder permissões sobre os recursos a entidades principais usando o controle de acesso baseado em tags do Lake Formation. Os criadores de LF-Tags recebem implicitamente essas permissões ao criar LF-Tags.

Depois de receber a permissão `Create LF-Tag` e criar LF-Tags com sucesso, o criador de LF-Tags pode atribuir LF-Tags a recursos e conceder permissões de LF-Tag (`Create LF-Tag`, `Alter` e `Drop`) a outras entidades principais não administrativas para gerenciar tags no data lake. Você pode gerenciar LF-Tags usando o console do Lake Formation, a API ou a AWS Command Line Interface (AWS CLI).

#### Note

Os administradores do Data Lake têm permissões implícitas do Lake Formation para criar, atualizar e excluir LF-Tags, atribuir LF-Tags a recursos e conceder permissões de LF-Tags às entidades principais.

Para conhecer as práticas recomendadas e as considerações, consulte [Considerações e práticas recomendadas de controle de acesso com base em tags do Lake Formation](#).

#### Tópicos

- [Adicionando criadores de LF-Tags](#)

- [Criação de LF-Tags](#)
- [Atualizar LF-Tags](#)
- [Excluir LF-Tags](#)
- [Listar LF-Tags](#)
- [Atribuição de LF-Tags aos recursos do catálogo de dados](#)
- [Visualização de tags LF-Tags atribuídas a um recurso](#)
- [Visualizando os recursos aos quais uma LF-Tag está atribuída](#)
- [Ciclo de vida de uma LF-tag](#)
- [Comparação do controle de acesso baseado em tags do Lake Formation com o controle de acesso baseado em atributos do IAM](#)

#### Consulte também

- [Concedendo, revogando e listando permissões de valor da LF-Tag](#)
- [Conceder permissões de data lake usando o método LF-TBAC](#)
- [Controle de acesso baseado em tags do Lake Formation](#)

## Adicionando criadores de LF-Tags

Por padrão, os administradores do data lake podem criar, atualizar e excluir LF-Tags, atribuir tags aos recursos do catálogo de dados e conceder permissões de tag aos principais. Se você quer delegar as operações de criação e gerenciamento de tags a entidades principais que não são administradores, o administrador do data lake pode criar funções de criador de LF-Tags e conceder a permissão `Create LF-Tag` do Lake Formation para as funções. Com a permissão `Create LF-Tag` concedida, os criadores de LF-Tags podem delegar tarefas de criação e manutenção de tags a outras entidades principais não administrativas.

#### Note

As concessões de permissão entre contas podem incluir somente permissões `Describe` e `Associate`. Você não pode conceder permissões `Create LF-Tag`, `Drop`, `Alter` e `Grant with LFTag expressions` a entidades principais em uma conta diferente.

## Tópicos

- [Permissões de IAM necessárias para criar LF-tags](#)
- [Adicionar criadores de LF-Tags](#)

### Consulte também

- [Concedendo, revogando e listando permissões de valor da LF-Tag](#)
- [Conceder permissões de data lake usando o método LF-TBAC](#)
- [Controle de acesso baseado em tags do Lake Formation](#)

## Permissões de IAM necessárias para criar LF-tags

Você deve configurar permissões para permitir que uma entidade principal do Lake Formation crie LF-tags. Adicione a seguinte declaração à política de permissões para a entidade principal que precisa ser um criador de LF-Tag.

### Note

Embora os administradores do data lake tenham permissões implícitas do Lake Formation para criar, atualizar e excluir LF-Tags, atribuir LF-Tags aos recursos e conceder LF-Tags às entidades principais, os administradores do data lake também precisam das seguintes permissões do IAM.

Para obter mais informações, consulte [Referência de personas e permissões do IAM do Lake Formation](#).

```
{
  "Sid": "Transformational",
  "Effect": "Allow",
  "Action": [
    "lakeformation:AddLFTagsToResource",
    "lakeformation:RemoveLFTagsFromResource",
    "lakeformation:GetResourceLFTags",
    "lakeformation:ListLFTags",
    "lakeformation:CreateLFTag",
```

```
    "lakeformation:GetLFTag",
    "lakeformation:UpdateLFTag",
    "lakeformation:DeleteLFTag",
    "lakeformation:SearchTablesByLFTags",
    "lakeformation:SearchDatabasesByLFTags"
  ]
}
```

As entidades principais que atribuem LF-Tags aos recursos e concedem LF-Tags a entidades principais devem ter as mesmas permissões, exceto as `CreateLFTag`, `UpdateLFTag` e `DeleteLFTag`.

### Adicionar criadores de LF-Tags

Um criador de LF-Tags pode criar uma LF-Tag, atualizar a chave e os valores da tag, excluir tags, associar tags aos recursos do catálogo de dados e conceder permissões sobre os recursos do catálogo de dados aos principais usando o método LF-TBAC. O criador da LF-Tag também pode conceder essas permissões a entidades principais.

Você pode criar funções de criador de LF-Tags usando o console do AWS Lake Formation, a API ou a AWS Command Line Interface (AWS CLI).

#### console

Para adicionar um criador de LF-Tag

1. Abra o console do Lake Formation em <https://console.aws.amazon.com/lakeformation/>.

Faça login como administrador de data lake.

2. No painel de navegação, em Permissões, selecione permissões e LF-Tags.

Na página Permissões e LF-Tags, escolha a seção Criadores de LF-Tags e selecione Adicionar criadores de LF-Tags.





## Add LF-Tag creators

LF-Tag creators can create and manage LF-Tags. [Learn more](#) 

### LF-Tag creator details

**IAM users and roles**  
Add IAM users or roles.

Choose IAM principals to add 

lf-developer   
User

**Permission**  
Choose the permission to grant.

Create LF-Tag

**Grantable permission**  
Choose the permission that may be granted to others.

Create LF-Tag

Cancel Add

3. Na página Adicionar criadores de LF-Tags, escolha um perfil ou usuário do IAM que tenha as permissões necessárias para criar LF-Tags.
4. Ativar caixa de seleção de permissão Create LF-Tag.
5. (Opcional) Para permitir que as entidades principais selecionadas concedam a permissão Create LF-Tag às entidades principais, escolha a permissão Create LF-Tag concedível.
6. Escolha Adicionar.

## AWS CLI

```
aws lakeformation grant-permissions --cli-input-json file://grantCreate
{
  "Principal": {
    "DataLakePrincipalIdentifier": "arn:aws:iam::123456789012:user/tag-manager"
  },
  "Resource": {
    "Catalog": {}
  },
  "Permissions": [
```

```

    "CreateLFTag"
  ],
  "PermissionsWithGrantOption": [
    "CreateLFTag"
  ]
}

```

A seguir estão as permissões disponíveis para a função de criador do LF-Tag:

Permissão	Descrição
Drop	Uma entidade principal com essa permissão em uma LF-Tag pode excluir uma LF-Tag do data lake. A entidade principal obtém a permissão <code>Describe</code> implícita em todos os valores de tag de um recurso de LF-Tag.
Alter	Uma entidade principal com essa permissão em uma LF-Tag pode adicionar ou remover o valor da tag de uma LF-Tag. A entidade principal obtém a permissão <code>Alter</code> implícita em todos os valores de tag de uma LF-Tag.
Describe	Uma entidade principal com essa permissão em uma LF-Tag pode visualizar a LF-Tag e os valores dela ao atribuir LF-Tags a recursos ou conceder permissões em LF-Tags. Você pode conceder <code>Describe</code> em todos os valores-chave ou em valores específicos.
Associate	Uma entidade principal com essa permissão em uma LF-Tag pode atribuir a LF-Tag a um recurso do Catálogo de Dados. Conceder <code>Associate</code> concede implicitamente <code>Describe</code> .
Grant with LF-Tag expression	Uma entidade principal com essa permissão em uma LF-Tag pode conceder permissões sobre os recursos de um catálogo de dados usando os valores e a chave da LF-Tag. Conceder <code>Grant with LF-Tag expression</code> concede implicitamente <code>Describe</code> .

Essas permissões são concedidas. Uma entidade principal que tenha recebido essas permissões com a opção de concessão pode concedê-las a outras entidades principais.

## Criação de LF-Tags

Todas as LF-Tags devem ser definidas no Lake Formation antes de serem usadas. Uma LF-Tag consiste em uma chave e um ou mais valores possíveis para essa chave.

Depois que o administrador do data lake configurar as permissões necessárias do IAM e as permissões do Lake Formation para a função de criador da LF-Tag, a entidade principal pode criar uma LF-Tag. O criador da LF-Tag obtém permissão implícita para atualizar ou remover qualquer valor da LF-Tag e excluir a LF-Tag.

Você pode criar funções de criador de LF-Tags usando o console do AWS Lake Formation, a API ou a AWS Command Line Interface (AWS CLI).

### Console

Para criar uma LF-Tag

1. Abra o console do Lake Formation em <https://console.aws.amazon.com/lakeformation/>.

Faça login como uma entidade principal com permissões de criador de LF-Tag ou como administrador do data lake.

2. No painel de navegação, em Permissões, selecione LF-Tags.

A página LF-Tags) é exibida.

	Key	Values	Owner account ID	LF-Tag permissions
<input type="radio"/>	LF-Test	lf-businessanalyst, customer	054881201579	<a href="#">View</a>
<input type="radio"/>	module	Customers	054881201579	<a href="#">View</a>

3. Selecione Adicionar LF-tag.
4. Na caixa de diálogo Adicionar LF-Tag, insira uma chave e um ou mais valores.

Cada chave deve ter pelo menos um valor. Para inserir vários valores, insira uma lista delimitada por vírgulas e pressione Enter ou insira um valor por vez e escolha Adicionar após cada um. O número máximo de aliases por usuário é de 1000.

5. Escolha Adicionar Tag.

## AWS CLI

Para criar uma LF-Tag

- Insira um comando `create-lf-tag`.

O exemplo a seguir cria uma LF-Tag com chave `module` e valores `Customers` e `Orders`.

```
aws lakeformation create-lf-tag --tag-key module --tag-values Customers Orders
```

Como criador da tags, a entidade principal obtém a permissão `Alter` sobre essa LF-Tag e pode atualizar ou remover qualquer valor de tag dessa LF-Tag. A entidade principal criadora da LF-Tag também pode conceder a permissão `Alter` a outra entidade principal para atualizar e remover os valores da tag dessa LF-Tag.

## Atualizar LF-Tags

Você atualiza uma LF-Tag na qual tem permissão `Alter` adicionando ou excluindo valores de chave permitidos. Não é possível alterar a chave LF-tag. Para alterar a chave, exclua a LF-Tag e adicione uma com a chave necessária. Além da permissão `Alter`, você também precisa da permissão do IAM `lakeformation:UpdateLFTag` para atualizar os valores.

Quando você exclui um valor de LF-Tag, nenhuma verificação é realizada para verificar a presença desse valor de LF-Tag em qualquer recurso do Catálogo de Dados. Se o valor da LF-Tag excluído estiver associado a um recurso, ele não estará mais visível para o recurso, e todas as entidades principais que receberam permissões nesse par de valores-chave não terão mais as permissões.

Antes de excluir um valor de LF-Tag, você pode, opcionalmente, usar o [comando `remove-lf-tags-from-resource`](#) para remover a LF-Tag dos recursos do Catálogo de Dados que têm o valor que você deseja excluir e, em seguida, remarcar o recurso com os valores que você deseja manter.

Somente administradores do data lake, o criador da LF-Tag e as entidades principais que têm permissões `Alter` na LF-Tag podem atualizar uma LF-Tag.

Você pode atualizar uma LF-Tag usando o console do AWS Lake Formation, a API ou a AWS Command Line Interface (AWS CLI).

## Console

Como atualizar uma política de LF-Tag (console)

1. Abra o console do Lake Formation em <https://console.aws.amazon.com/lakeformation/>.

Faça login como administrador do data lake, criador de LF-Tag ou como uma entidade principal com a permissão `Alter` na LF-Tag.

2. No painel de navegação, em Permissões, selecione LF-Tags.
3. Na página LF-Tags, selecione uma LF-Tag e, em seguida, escolha Editar.
4. Na caixa de diálogo Editar LF-Tag, adicione ou remova valores de LF-Tag.

Para inserir vários Valores, insira uma lista delimitada por vírgulas e pressione Enter ou insira um valor por vez e escolha Adicionar após cada um.

5. Escolha Salvar.

## AWS CLI

Para atualizar uma LF-Tag (AWS CLI)

- Insira um comando `update-lf-tag`. Forneça um ou ambos os argumentos a seguir:
  - `--tag-values-to-add`
  - `--tag-values-to-delete`

### Example

O exemplo a seguir substitui o valor `vp` pelo valor `vice-president` da chave LF-Tag `level`.

```
aws lakeformation update-lf-tag --tag-key level --tag-values-to-add vice-president
--tag-values-to-delete vp
```

## Excluir LF-Tags

Você pode excluir LF-Tags que não estão mais em uso. Nenhuma verificação é realizada quanto à presença da LF-Tag em um recurso do Catálogo de Dados. Se a LF-Tag excluída estiver associada a um recurso, ela não estará mais visível para ele, e todas as entidades principais que receberam permissões nessa LF-Tag não as terão mais.

Antes de excluir uma LF-Tag, você pode, opcionalmente, usar o comando [remove-lf-tags-from-resource](#) para remover a LF-Tag de todos os recursos.

Somente administradores do data lake, o criador da LF-Tag ou uma entidade principal que tenha a permissão `Drop` sobre a LF-Tag podem excluir uma LF-Tag. Além da permissão `Drop`, a entidade principal também precisa da permissão `lakeformation:DeleteLFTag` do IAM para excluir uma LF-Tag.

Você pode excluir uma LF-Tag usando o console do AWS Lake Formation, a API ou a AWS Command Line Interface (AWS CLI).

### Console

Para excluir uma LF-Tag (console)

1. Abra o console do Lake Formation em <https://console.aws.amazon.com/lakeformation/>.

Faça login como administrador de data lake.

2. No painel de navegação, em LF-Tags e Permissões, selecione LF-Tags.
3. Na página LF-Tags, selecione uma LF-Tag e, em seguida, escolha Excluir.
4. Na caixa de diálogo Excluir ambiente de tag?, para confirmar a exclusão, insira o valor da chave LF-Tag no campo designado e escolha Excluir.

### AWS CLI

Para excluir uma LF-Tag (AWS CLI)

- Insira um comando `delete-lf-tag`. Forneça a chave da LF-Tag a ser excluída.

#### Example

O exemplo a seguir exclui a LF-tag com a chave `region`.

```
aws lakeformation delete-lf-tag --tag-key region
```

## Listar LF-Tags

Você pode listar as LF-Tags nas quais você tem as permissões `Describe` ou `Associate`. Os valores listados com cada chave de LF-Tag são os valores sobre os quais você tem permissões.

O criador do LF-Tag tem permissões implícitas para ver os LF-tags que ele criou.

Os administradores do Data Lake podem ver todas as LF-Tags definidas na conta AWS local e todas as LF-Tags para as quais as permissões `Describe` e `Associate` e foram concedidas à conta local por contas externas. O administrador do data lake pode ver todos os valores de todas as LF-Tags.

Você pode listar LF-Tags usando o console do AWS Lake Formation, a API ou a AWS Command Line Interface (AWS CLI).

## Console

Para listar tags (console)

1. Abra o console do Lake Formation em <https://console.aws.amazon.com/lakeformation/>.

Faça login como criador da LF-Tag, como administrador do data lake ou como entidade principal que recebeu permissões em LF-Tags e que tem a permissão do IAM `lakeformation:ListLFTags`.

2. No painel de navegação, em LF-Tags e permissões, selecione LF-Tags.

A página LF-Tags é exibida.

Key	Values	Owner account ID	LF-Tag permissions
LF-Test	lf-businessanalyst, customer	054881201579	View
module	Customers	054881201579	View

Verifique a coluna ID da conta do proprietário para determinar as LF-Tags que foram compartilhadas com sua conta a partir de uma conta externa.

## AWS CLI

Para listar LF-Tags (AWS CLI)

- Execute o comando a seguir como administrador de data lake ou como entidade principal que tenha recebido permissões em LF-Tags e que tenha a permissão do IAM `lakeformation:ListLFTags`.

```
aws lakeformation list-lf-tags
```

A saída é semelhante à seguinte.

```
{
  "LFTags": [
    {
      "CatalogId": "111122223333",
      "TagKey": "level",
      "TagValues": [
        "director",
        "vp",
        "c-level"
      ]
    },
    {
      "CatalogId": "111122223333",
      "TagKey": "module",
      "TagValues": [
        "Orders",
        "Sales",
        "Customers"
      ]
    }
  ]
}
```

Para ver também as LF-Tags que foram concedidas por contas externas, inclua a opção de comando `--resource-share-type ALL`.



```
aws lakeformation list-lf-tags --resource-share-type ALL
```

A saída é semelhante à seguinte. Observe a chave `NextToken`, que indica que há mais para listar.

```
{
  "LFTags": [
    {
      "CatalogId": "111122223333",
      "TagKey": "level",
      "TagValues": [
        "director",
        "vp",
        "c-level"
      ]
    },
    {
      "CatalogId": "111122223333",
      "TagKey": "module",
      "TagValues": [
        "Orders",
        "Sales",
        "Customers"
      ]
    }
  ],
  "NextToken": "eyJleHBpcmF0aWw...ZXh0Ijpo0cnV1fQ=="
}
```

Repita o comando e adicione o argumento `--next-token` para ver todas as LF-Tags e LF-Tags locais restantes que foram concedidas por contas externas. As LF-Tags de contas externas estão sempre em uma página separada.

```
aws lakeformation list-lf-tags --resource-share-type ALL
--next-token eyJleHBpcmF0aWw...ZXh0Ijpo0cnV1fQ==
```

```
{
  "LFTags": [
    {
      "CatalogId": "123456789012",
```

```
        "TagKey": "region",
        "TagValues": [
            "central",
            "south"
        ]
    }
]
}
```

## API

Você pode usar os SDKs disponíveis para o Lake Formation para listar as tags que o solicitante tem permissão para visualizar.

```
import boto3

client = boto3.client('lakeformation')
...

response = client.list_lf_tags(
    CatalogId='string',
    ResourceShareType='ALL',
    MaxResults=50'
)
```

Este comando retorna um objeto dict com a seguinte estrutura:

```
{
  'LFTags': [
    {
      'CatalogId': 'string',
      'TagKey': 'string',
      'TagValues': [
        'string',
      ]
    },
  ],
  'NextToken': 'string'
}
```

Para mais informações sobre as permissões necessárias, consulte [Referência de personas e permissões do IAM do Lake Formation](#).

### Atribuição de LF-Tags aos recursos do catálogo de dados

Você pode atribuir LF-tags a recursos do Catálogo de dados (bancos de dados, tabelas e colunas) para controlar o acesso a esses recursos. Somente entidades principais que recebem LF-Tags correspondentes (e as que recebem acesso com o método de recurso nomeado) podem acessar os recursos.

Se uma tabela herdar uma LF-Tag de um banco de dados ou uma coluna herdar uma LF-Tag de uma tabela, você poderá substituir o valor herdado atribuindo um novo valor à chave LF-Tag.

Número máximo de tags que você pode atribuir a um recurso é de 50.

### Tópicos

- [Requisitos para gerenciar tags atribuídas aos recursos](#)
- [Atribuir LF-Tags a uma coluna da tabela](#)
- [Atribuir LF-Tags aos recursos do catálogo de dados](#)
- [Atualização de LF-Tags para um recurso](#)
- [Remoção de LF-Tag de um recurso](#)

### Requisitos para gerenciar tags atribuídas aos recursos

Para atribuir uma LF-Tag a um recurso do Catálogo de Dados, você deve:

- Tenha a permissão ASSOCIATE do Lake Formation na LF-Tag.
- Tenha a permissão `lakeformation:AddLFTagsToResource` do IAM.
- Tenha a permissão `glue:getDatabase` em um banco de dados Glue.
- Seja o proprietário do recurso (criador), tenha a permissão do Lake Formation Super no recurso com a opção GRANT ou tenha as seguintes permissões com a opção GRANT:
  - Para bancos de dados na mesma conta AWS: DESCRIBE, CREATE\_TABLE, ALTER e DROP
  - Para bancos de dados em uma conta externa: DESCRIBE, CREATE\_TABLE e ALTER
  - Para tabelas (e colunas): DESCRIBE, ALTER, DROP, INSERT, SELECT e DELETE

Além disso, a LF-Tag e o recurso ao qual ela está sendo atribuída devem estar na mesma conta AWS.

Para remover uma LF-Tag de um recurso do catálogo de dados, você deve atender a esses requisitos e também ter a permissão do IAM `lakeformation:RemoveLFTagsFromResource`.

### Atribuir LF-Tags a uma coluna da tabela

Para atribuir LF-Tags a uma coluna da tabela (console)

1. Abra o console do Lake Formation em <https://console.aws.amazon.com/lakeformation/>.


Faça login como um usuário que atende aos requisitos listados acima.

2. No painel de navegação, selecione Tabelas.
3. Escolha um nome de tabela (não o botão de opção ao lado do nome da tabela).
4. Na página de detalhes da tabela, na seção Esquema, escolha Editar esquema.
5. Na página Editar esquema, selecione uma ou mais colunas e escolha Editar tags.

#### Note

Se você pretende adicionar ou excluir colunas e salvar uma nova versão, faça isso primeiro. Em seguida, edite as LF-Tags.

A caixa de diálogo Editar LF-Tags é exibida e exibe todas as LF-Tags herdadas da tabela.

**Edit LF-Tags: product\_id** [Learn More](#) 

**LF-Tags**

After they are associated with catalog resources, LF-Tags allow you to create scalable permissions.

Inherited keys	Values
<input type="text" value="level"/>	<input type="text" value="director (inherited)"/>
<input type="text" value="module"/>	<input type="text" value="Orders (inherited)"/>

[Assign new LF-Tag](#)

You can add 50 more tags.

- (Opcional) Na lista Valores ao lado do campo Chaves herdadas, escolha um valor para substituir o valor herdado.
- (Opcional) Escolha Atribuir nova LF-Tag. Em seguida, em Chaves atribuídas, escolha uma chave e, em Valores, escolha um valor para a chave.

### Edit LF-Tags: product\_id [Learn More](#) ✕

**LF-Tags**  
After they are associated with catalog resources, LF-Tags allow you to create scalable permissions.

Inherited keys	Values
<input type="text" value="level"/>	director (inherited) ▼
<input type="text" value="module"/>	Orders (inherited) ▼

---

Assigned keys	Values	
<input type="text" value="environment"/> ✕	Production ▲	<input type="button" value="Remove"/>
<input type="button" value="Assign new LF-Tag"/>	Production	
	Development	

You can add 49 more tags.

8. (Opcional) Selecione Adicionar nova LF-Tag novamente para adicionar outra LF-Tag.
9. Escolha Salvar.

Atribuir LF-Tags aos recursos do catálogo de dados

Console

Para atribuir LF-Tags a um banco de dados ou tabela do Catálogo de Dados

1. Abra o console do Lake Formation em <https://console.aws.amazon.com/lakeformation/>.

Faça login como um usuário que atende aos requisitos listados anteriormente.

2. No painel de navegação, em Catálogo de dados, faça o seguinte:
  - Para atribuir LF-Tags aos bancos de dados, escolha Bancos de dados.
  - Para atribuir LF-Tags às tabelas, escolha Tabelas.

- Escolha um banco de dados ou tabela e, no menu Ações, escolha Editar tags.

A caixa de diálogo Editar LF-Tags: **resource-name** é exibida.

Se uma tabela herdar LF-Tags do banco de dados que a contém, a janela exibirá as LF-Tags herdadas. Caso contrário, ele exibirá o texto “Não há LF-Tags herdadas associadas ao recurso.”

### Edit LF-Tags: inventory [Learn More](#)

✕

LF-Tags

After they are associated with catalog resources, LF-Tags allow you to create scalable permissions.

Inherited keys	Values
<input type="text" value="level"/>	<input type="text" value="director (inherited)"/>

---

Assigned keys	Values	
<input type="text" value="module"/> ✕	<input type="text" value="Enter LF-Tag value"/> ▲	Remove
<input type="button" value="Assign new LF-Tag"/>	<div style="border: 1px solid #ccc; padding: 2px;">Orders</div> <div style="border: 1px solid #ccc; padding: 2px;">Sales</div> <div style="border: 1px solid #ccc; padding: 2px;">Customers</div>	
You can add 49 more tags.		

- (Opcional) Se uma tabela tiver LF-Tags herdadas, na lista Valores ao lado do campo Chaves herdadas, você poderá escolher um valor para substituir o valor herdado.
- Para atribuir novas LF-Tags, execute estas etapas:
  - Escolha Atribuir nova LF-Tag.
  - No campo Chaves atribuídas, escolha uma chave LF-Tag e, no campo Valores, escolha um valor.
  - (Opcional) Escolha Atribuir nova LF-Tag novamente para atribuir uma LF-Tag adicional.
- Escolha Salvar.

## AWS CLI

Para atribuir LF-Tags a um recurso do Catálogo de Dados

- Execute o comando `add-lf-tags-to-resource`.

O exemplo a seguir atribui a LF-Tag `module=orders` à tabela `orders` no banco de dados `erp`. Ele usa a sintaxe de atalho para o argumento `--lf-tags`. A propriedade `CatalogID` para `--lf-tags` é opcional. Se não for fornecido, o ID do catálogo do recurso (nesse caso, a tabela) será assumido.

```
aws lakeformation add-lf-tags-to-resource --resource '{ "Table":
{"DatabaseName":"erp", "Name":"orders"}}' --lf-tags
CatalogId=111122223333,TagKey=module,TagValues=orders
```

A seguir está o resultado se o comando for bem-sucedido.

```
{
  "Failures": []
}
```

O próximo exemplo atribui duas LF-Tags à tabela `sales` e usa a sintaxe JSON para o argumento `--lf-tags`.

```
aws lakeformation add-lf-tags-to-resource --resource '{ "Table":
{"DatabaseName":"erp", "Name":"sales"}}' --lf-tags '[{"TagKey":
"module","TagValues": ["sales"]}, {"TagKey": "environment","TagValues":
["development"]}]'
```

O próximo exemplo atribui a LF-Tag `level=director` à coluna `total` da tabela `sales`.

```
aws lakeformation add-lf-tags-to-resource --resource '{ "TableWithColumns":
{"DatabaseName":"erp", "Name":"sales", "ColumnNames":["total"]}}' --lf-tags
TagKey=level,TagValues=director
```



## Atualização de LF-Tags para um recurso

Para atualizar uma LF-Tag para um recurso do Catálogo de Dados (AWS CLI).

- Use o comando `add-lf-tags-to-resource`, conforme descrito no procedimento anterior.

Adicionar uma tag do LF com a mesma chave de uma tag do LF existente, mas com um valor diferente, atualiza o valor existente.

## Remoção de LF-Tag de um recurso

Para remover uma LF-Tag de um recurso do Catálogo de Dados (AWS CLI)

- Execute o comando `remove-lf-tags-from-resource`.

Se uma tabela tiver um valor de tag do LF que substitua o valor herdado do banco de dados pai, remover essa tag do LF da tabela vai restaurar o valor herdado. Esse comportamento também se aplica a uma coluna que substitui os valores-chave herdados da tabela.

O exemplo a seguir remove a LF-Tag `level=director` da coluna `total` da tabela `sales`. A propriedade `CatalogID` para `--lf-tags` é opcional. Se não for fornecido, o ID do catálogo do recurso (nesse caso, a tabela) será assumido.

```
aws lakeformation remove-lf-tags-from-resource
--resource ' { "TableWithColumns":
{ "DatabaseName": "erp", "Name": "sales", "ColumnNames": [ "total" ] } } '
--lf-tags CatalogId=111122223333,TagKey=level,TagValues=director
```

## Visualização de tags LF-Tags atribuídas a um recurso

Você pode visualizar as LF-Tags atribuídas a um recurso do Catálogo de Dados. Você deve ter a permissão `DESCRIBE` ou `ASSOCIATE` em uma LF-Tag para visualizá-la.

### Console

Para visualizar as LF-Tags atribuídas a um recurso (console)

1. Abra o console do Lake Formation em <https://console.aws.amazon.com/lakeformation/>.

Faça login como administrador do data lake, proprietário do recurso ou usuário que tenha recebido permissões do Lake Formation no recurso.

2. No painel de navegação, no título Catálogo de dados, faça o seguinte:
  - Para visualizar as LF-Tags atribuídas a um banco de dados, escolha Bancos de dados.
  - Para ver as LF-Tags atribuídas a uma tabela, escolha Tabelas.
3. Na página Tabelas ou Bancos de Dados, escolha o nome do banco de dados ou da tabela. Na seção de detalhes, role para baixo até a seção LF-Tags.

A captura de tela a seguir mostra as LF-Tags atribuídas a uma tabela `customers`, que está contida no banco de dados `retail`. A LF-Tag `module` é herdada do banco de dados. A coluna `credit_limit` tem a LF-Tag `level=vp` atribuída.

### LF-Tags (3) Edit tags

LF-Tags are key-value pairs that you can assign to data catalog resources, such as databases, tables, and columns. You can then grant permissions to principals based on these tags to control access to the resources. Table columns inherit all LF-Tags that are assigned to the table. [Learn More](#)

< 1 >

Resource <span style="float: right;">▲</span>	Key <span style="float: right;">▼</span>	Value <span style="float: right;">▼</span>	Inherited from
customers (table)	module	Customers	retail
customers (table)	environment	Production	-
credit_limit (column)	level	vp	-

## AWS CLI

Para visualizar as LF-Tags atribuídas a um recurso (AWS CLI)

- Digite um comando semelhante ao seguinte:

```
aws lakeformation get-resource-lf-tags --show-assigned-lf-tags --
resource '{ "Table": {"CatalogId":"111122223333", "DatabaseName":"erp",
"Name":"sales"}}'
```

O comando retorna a seguinte saída.

```
{
  "TableTags": [
    {
      "CatalogId": "111122223333",
      "TagKey": "module",
      "TagValues": [
        "sales"
      ]
    },
    {
      "CatalogId": "111122223333",
      "TagKey": "environment",
      "TagValues": [
        "development"
      ]
    }
  ],
  "ColumnTags": [
    {
      "Name": "total",
      "Tags": [
        {
          "CatalogId": "111122223333",
          "TagKey": "level",
          "TagValues": [
            "director"
          ]
        }
      ]
    }
  ]
}
```

Essa saída mostra somente LF-Tags que são atribuídas explicitamente, não herdadas. Se você quiser ver todas as LF-Tags em todas as colunas, incluindo as LF-Tags herdadas, omita a opção `--show-assigned-lf-tags`.

## Visualizando os recursos aos quais uma LF-Tag está atribuída

Você pode visualizar todos os recursos do Catálogo de Dados aos quais uma determinada chave LF-Tag está atribuída. Para fazer isso, você precisa das seguintes permissões do Lake Formation:

- `Describe` ou `Associate` na LF-Tag.
- `Describe` ou qualquer outra permissão do Lake Formation sobre o recurso.

Além disso, você precisa das seguintes permissões do AWS Identity and Access Management (IAM):

- `lakeformation:SearchDatabasesByLFTags`
- `lakeformation:SearchTablesByLFTags`

## Console

Para visualizar os recursos aos quais uma LF-Tag está atribuída (console)

1. Abra o console do Lake Formation em <https://console.aws.amazon.com/lakeformation/>.

Faça login como administrador do data lake ou como usuário que atenda aos requisitos listados anteriormente.

2. No painel de navegação, em Permissões e Perfis e tarefas administrativas, selecione LF-Tags.
3. Escolha uma chave LF-Tag (não o botão de opção ao lado do nome da chave).

A página de detalhes da LF-Tag exibe uma lista dos recursos aos quais a LF-Tag foi atribuída.

# module

**LF-Tag**

Delete

Edit

Key  
module

Values  
Orders, Sales, Customers

## Associated data catalog resources (12)

Q Find resource

Key	Values ▾	Resource type ▾	Resource ▾
module	Customers	DATABASE	<a href="#">retail</a>
module	Customers	TABLE	<a href="#">customers</a>
module	Orders	TABLE	<a href="#">inventory</a>
module	Customers	COLUMN	<a href="#">customers.cust_first_name</a>
module	Customers	COLUMN	<a href="#">customers.work_phone_number</a>
module	Customers	COLUMN	<a href="#">customers.company_name</a>
module	Customers	COLUMN	<a href="#">customers.credit_limit</a>

## AWS CLI

Para visualizar os recursos aos quais uma LF-Tag está atribuída

- Execute um comando `search-tables-by-lf-tags` or `search-databases-by-lf-tags`.

## Example

O exemplo a seguir lista tabelas e colunas que têm a LF-Tag `level=vp` atribuída. Para cada tabela e coluna listada, todas as LF-Tags atribuídas à tabela ou coluna são geradas, não apenas a expressão de pesquisa.

```
aws lakeformation search-tables-by-lf-tags --expression
TagKey=level,TagValues=vp
```

Para mais informações sobre as permissões necessárias, consulte [Referência de personas e permissões do IAM do Lake Formation](#).

### Ciclo de vida de uma LF-tag

1. O criador da LF-Tag, Michael, cria uma LF-Tag `module=Customers`.
2. Michael concede `Associate` na LF-Tag ao engenheiro de dados Eduardo. Conceder `Associate` concede implicitamente `Describe`.
3. Michael concede `Super` na tabela `Custs` a Eduardo com a opção de concessão, para que Eduardo possa atribuir LF-Tags à tabela. Para obter mais informações, consulte [Atribuição de LF-Tags aos recursos do catálogo de dados](#).
4. Eduardo atribui a LF-Tag `module=customers` à tabela `Custs`.
5. Michael faz a seguinte concessão à engenheira de dados Sandra (em pseudocódigo).

```
GRANT (SELECT, INSERT ON TABLES) ON TAGS module=customers TO Sandra WITH GRANT OPTION
```

6. Sandra faz a seguinte concessão à analista de dados Maria.

```
GRANT (SELECT ON TABLES) ON TAGS module=customers TO Maria
```

Agora, Maria pode executar consultas na tabela `Custs`.

### Consulte também

- [Controle de acesso a metadados](#)

## Comparação do controle de acesso baseado em tags do Lake Formation com o controle de acesso baseado em atributos do IAM

O controle de acesso por atributo (ABAC) é uma estratégia de autorização que define permissões com base em atributos. Na AWS, esses atributos são chamados de tags. Você pode anexar tags a recursos do IAM, incluindo entidades (usuários ou funções) do IAM, e a recursos da AWS. É possível criar uma única política de ABAC ou um pequeno conjunto de políticas para suas entidades do IAM. Essas políticas de ABAC podem ser criadas para permitir operações quando a tag da entidade principal corresponder à tag de recurso. O ABAC é útil em ambientes que estão crescendo rapidamente e ajuda em situações onde o gerenciamento de políticas se torna um problema.

As equipes de segurança e governança da nuvem usam o IAM para definir políticas de acesso e permissões de segurança para todos os recursos, incluindo buckets do Amazon S3, instâncias do Amazon EC2 e quaisquer recursos que você possa referenciar com um ARN. As políticas do IAM definem permissões amplas (grosseiras) para seus recursos de data lake, por exemplo, para permitir ou negar acesso no nível de bucket, prefixo ou banco de dados do Amazon S3. Para obter mais informações sobre o ABAC, consulte [Para que é ABAC AWS?](#) no Guia do usuário do IAM.

Por exemplo, é possível criar três funções com a chave de tag `project-access`. Defina o valor da tag da primeira função como `Dev`, a segunda como `Marketing` e a terceira como `Support`. Atribua tags com o valor apropriado aos recursos. Depois disso, é possível usar uma única política que permitirá o acesso quando a função e o recurso estiverem marcados com o mesmo valor para `project-access`.

As equipes de governança de dados usam o Lake Formation para definir permissões refinadas para recursos específicos do data lake. LF-tags são atribuídas a recursos do Catálogo de Dados (bancos de dados, tabelas e colunas) e são concedidas a entidades principais. Uma entidade principal com LF-Tags que correspondem às LF-Tags de um recurso pode acessar esse recurso. As permissões do Lake Formation são secundárias às permissões do IAM. Por exemplo, se as permissões do IAM não permitirem que um usuário acesse um data lake, o Lake Formation não concederá acesso a nenhum recurso dentro desse data lake para esse usuário, mesmo que a entidade principal e o recurso tenham LF-Tags correspondentes.

O controle de acesso baseado em tags do Lake Formation (LF-TBAC) funciona com o IAM ABAC para fornecer níveis adicionais de permissões para dados e recursos do Lake Formation.

- As permissões de TBAC do Lake Formation são dimensionadas com inovação. Não é mais necessário que um administrador atualize as políticas existentes para permitir o acesso a novos recursos. Por exemplo, suponha que você use uma estratégia ABAC do IAM com a tag `project-`

access para fornecer acesso a bancos de dados específicos dentro do Lake Formation. Usando o LF-TBAC, a LF-Tag `Project=SuperApp` é atribuída a tabelas ou colunas específicas, e a mesma LF-Tag é concedida a um desenvolvedor para esse projeto. Pelo IAM, o desenvolvedor pode acessar o banco de dados, e as permissões do LF-TBAC concedem ao desenvolvedor acesso adicional a tabelas ou colunas específicas dentro das tabelas. Se uma nova tabela for adicionada ao projeto, o administrador do Lake Formation só precisará atribuir a tag à nova tabela para que o desenvolvedor tenha acesso a ela.

- O Lake Formation TBAC exige menos políticas de IAM. Como você usa as políticas do IAM para conceder acesso de alto nível aos recursos do Lake Formation e o TBAC do Lake Formation para gerenciar um acesso mais preciso aos dados, você cria menos políticas do IAM.
- Usando o Lake Formation TBAC, as equipes podem mudar e crescer rapidamente. Isso ocorre porque as permissões para novos recursos são concedidas automaticamente com base em atributos. Por exemplo, se um novo desenvolvedor ingressar no projeto, é fácil conceder acesso a esse desenvolvedor associando ao perfil do IAM ao usuário e, em seguida, atribuindo as LF-Tags necessárias ao usuário. Você não precisa alterar a política do IAM para dar suporte a um novo projeto ou criar novas LF-Tags.
- Permissões mais refinadas são possíveis usando o Lake Formation TBAC. As políticas do IAM concedem acesso aos recursos de nível superior, como bancos de dados ou tabelas do catálogo de dados. Usando o Lake Formation TBAC, você pode conceder acesso a tabelas ou colunas específicas que contêm valores de dados específicos.

#### Note

As tags IAM não são iguais às LF-Tags. Essas tags não são intercambiáveis. As LF-Tags são usadas para conceder permissões do Lake Formation e as tags IAM são usadas para definir políticas do IAM.

## Concedendo, revogando e listando permissões de valor da LF-Tag

Você pode conceder as permissões `Drop` e `Alter` nas LF-Tags às entidades principais para gerenciar as expressões de valor da LF-Tags. Você também pode conceder as permissões `Describe`, `Associate`, e `Grant with LF-Tag expressions` de LF-Tags a entidades principais para visualizar as LF-Tags e atribuí-las a recursos do Catálogo de dados (bancos de dados, tabelas e colunas). Quando as LF-Tags são atribuídas aos recursos do Catálogo de dados, você pode usar o método de controle de acesso baseado em tags do Lake Formation (LF-TBAC) para proteger



esses recursos. Para ter mais informações, consulte [Controle de acesso baseado em tags do Lake Formation](#).

Você pode conceder essas permissões com a opção de concessão para que outras entidades principais possam concedê-las. As permissões Grant with LF-Tag expressions, Describe e Associate são explicadas em [Adicionar criadores de LF-Tags](#).

Você pode conceder as Associate permissões Describe e em uma etiqueta LF para uma conta externa AWS . Um administrador de data lake nessa conta pode então conceder essas permissões a outras entidades principais na conta. As entidades principais às quais o administrador do data lake na conta externa concede a permissão Associate podem então atribuir LF-Tags aos recursos do Catálogo de dados que você compartilhou com a conta deles.

Ao conceder para uma conta externa, você deve incluir a opção de concessão.

Você pode conceder permissões em LF-Tags usando o console do Lake Formation, a API ou o AWS Command Line Interface (AWS CLI).

## Tópicos

- [Listando as permissões do LF-Tag usando o console](#)
- [Concedendo permissões de LF-Tag usando o console](#)
- [Conceder, revogar e listar permissões de tag LF usando o AWS CLI](#)

Para obter mais informações, consulte [Gerenciar LF-Tags para controle de acesso a metadados e Controle de acesso baseado em tags do Lake Formation](#).

## Listando as permissões do LF-Tag usando o console

Você pode usar o console do Lake Formation para visualizar as permissões concedidas nas LF-Tags. Você deve ser um criador de LF-Tags, administrador de um data lake ou ter a permissão Describe ou Associate em uma LF-Tag para vê-la.

Para listar as permissões do LF-Tag (console)

1. Abra o console do Lake Formation em <https://console.aws.amazon.com/lakeformation/>.

Faça login como criador da LF-Tags, administrador do data lake ou como usuário a quem as permissões Drop, Alter, Associate ou Describe nas LF-Tags foram concedidas.

- No painel de navegação, em Permissões, escolha LF-Tags e permissões e, então, a seção Permissões de LF-Tag.

A seção de Permissões de LF-Tags mostra uma tabela que contém chaves de entidades principais, tags, valores e permissões.

The screenshot shows the 'LF-Tag permissions' section in the AWS Lake Formation console. It features a search bar with the placeholder text 'Find permissions by LF-Tag key and value'. Below the search bar is a table with the following columns: Principal, Principal type, Keys, Values, LF-Tag permissions, LF-Tag value permissions, and Grantable. The table contains six rows of data, each representing an IAM role with specific permissions.

Principal	Principal type	Keys	Values	LF-Tag permissions	LF-Tag value permissions	Grantable
arn:aws:iam::[redacted]:role/Admin	IAM role	module	All values	Alter, Drop	-	Alter, Drop
arn:aws:iam::[redacted]:role/Admin	IAM role	module	All values	-	Describe	Describe
arn:aws:iam::[redacted]:role/Admin	IAM role	module	All values	-	Associate	Associate
arn:aws:iam::[redacted]:role/Admin	IAM role	module	All values	-	Grant with LF-Tag expression	Grant with LF-Tag expression
arn:aws:iam::[redacted]:role/Admin	IAM role	LF-Test	All values	-	Describe	Describe
arn:aws:iam::[redacted]:role/Admin	IAM role	LF-Test	All values	-	Associate	Associate

## Concedendo permissões de LF-Tag usando o console

As etapas a seguir explicam como conceder permissões a LF-Tags usando a página Conceder permissões a LF-Tags no console do Lake Formation. A página está dividida nas seguintes seções:

- Tipos de permissão: o tipo de permissão a ser concedida.
- Diretores — Os usuários, funções ou AWS contas aos quais conceder permissões.
- LF-Tags – As LF-Tags para as quais conceder permissões.
- Permissões – As permissões a serem concedidas.

## Abra a página Conceder permissões a LF-Tags

- Abra o console do Lake Formation em <https://console.aws.amazon.com/lakeformation/>.

Faça login como criador de LF-Tags, administrador do data lake ou como usuário. As permissões de LF-Tags ou do par de valores-chave de LF-Tags nas LF-Tags foram concedidas com a opção Grant.

- No painel de navegação, escolha LF-Tags e permissões, escolha a seção Permissões de LF-Tags.
- Escolha Grant permissions (Conceder permissões).

## Especifique o tipo de permissão

Na seção Tipo de permissões, escolha um tipo de permissão.

### Permissões de LF-tags

Escolha as Permissões do LF-Tag para permitir que as entidades principais atualizem os valores da LF-Tag ou excluam a LF-Tag.

### Permissões de pares de chave-valor de LF-Tag

Escolha as Permissões do par chave-valor da LF-Tag para permitir que as entidades principais atribuam LF-Tags aos recursos do Catálogo de dados, visualizem LF-Tags e valores e concedam às entidades principais permissões baseadas nos recursos do Catálogo de dados.

As opções disponíveis nas seções a seguir dependem do Tipo de permissões.

## Especifique as entidades principais

### Note

Você não pode conceder permissões de LF-Tags (Alter e Drop) a contas externas ou entidades principais em outra conta.

Na seção Entidades principais, escolha um tipo de entidade principal e especifique as entidades principais às quais conceder permissões.

### Principals

**IAM users and roles**  
Users or roles from this AWS account.

**SAML users and groups**  
SAML users and group or QuickSight ARNs.

**External accounts**  
AWS account, AWS organization or IAM principal outside of this account

**IAM users and roles**  
Add one or more IAM users or roles.

Choose IAM principals to add ▼

## Usuários e perfis do IAM

Escolha um ou mais usuários ou perfis na lista de usuários e perfis do IAM.

## Usuários e grupos SAML

Para QuickSight usuários e grupos do SAML e da Amazon, insira um ou mais nomes de recursos da Amazon (ARNs) para usuários ou grupos federados por meio do SAML ou ARNs para usuários ou grupos da Amazon. QuickSight Pressione Enter após cada ARN.

Para obter informações sobre como construir os ARNs, consulte [Lake Formation concede e revoga comandos AWS CLI](#).

### Note

A integração do Lake Formation com a Amazon QuickSight é compatível somente com o Amazon QuickSight Enterprise Edition.

## Contas externas

Em AWS conta, insira uma ou mais IDs de AWS conta válidas. Pressione Enter após cada ID.

O ID da organização consiste em "o-" seguido por 10 a 32 letras minúsculas ou dígitos.

Um ID de unidade organizacional começa com "ou-" seguido de 4 a 32 letras minúsculas ou dígitos (o ID da raiz que contém a UO). Essa sequência é seguida por um segundo travessão "-" e 8 a 32 letras minúsculas ou dígitos adicionais.

Para a entidade principal do IAM, insira o ARN para o usuário ou o perfil do IAM.

## Especifique as LF-Tags

Para conceder permissões a LF-Tags, na seção Permissões de LF-Tags, especifique as LF-Tags para as quais conceder permissões.

## LF-Tag permissions

**LF-Tags**  
Choose the LF-Tags you want to grant permissions to.

*Choose one or more LF-Tags* ▼

Department ✕

**Permissions**  
Choose the specific LF-Tag permissions to grant.

- Alter**  
Update or delete key values.
- Drop**  
Delete tag(s).

**Grantable permissions**  
Choose the permissions that the grant recipient(s) can grant to other principals.

- Alter**  
Update or delete key values.
- Drop**  
Delete tag(s).

Cancel **Grant**

- Escolha uma ou mais LF-Tags usando o menu suspenso.

### Especificar os pares de chave-valor

1. Para conceder permissões em pares de chave-valor da LF-Tag, (você precisa primeiro escolher as Permissões do par chave-valor da LF-Tag como o Tipo de permissão), escolha Adicionar par chave-valor da LF-Tag para revelar a primeira linha de campos para especificar a chave e os valores da LF-Tag.

## LF-Tag key-value pair permissions

Key  Values

You can add 50 more LF-Tags.

**Permissions**  
Choose the specific key-value pair permissions to grant.

- Describe**  
See keys and values.
- Associate**  
Assign LF-Tags to databases, tables, and columns.
- Grant with LF-Tag expression**  
Allow the principal(s) to grant access permissions using the LF-Tag(s).

**Grantable permissions**  
Choose the permissions that the grant recipient(s) can grant to other principals.

- Describe**  
See keys and values.
- Associate**  
Assign LF-Tags to databases, tables, and columns.
- Grant with LF-Tag expression**  
Allow the principal(s) to grant access permissions using the LF-Tag(s).

2. Posicione o cursor no campo Chave, opcionalmente comece a digitar para restringir a lista de seleção e selecione uma tecla LF-Tag.
3. Na lista Valores, selecione um ou mais valores e pressione Guia ou clique ou toque fora do campo para salvar os valores selecionados.

### Note

Se uma das linhas na lista Valores estiver em foco, pressionar Enter marcará ou desmarcará a caixa de seleção.

Os valores selecionados aparecem como blocos abaixo da lista de Valores. Escolha o ✕ para remover um valor. Escolha Remover para remover toda a LF-Tags.

4. Para adicionar outra LF-Tag, escolha Adicionar LF-Tag novamente e repita as duas etapas anteriores.

### Especifique as permissões

Esta seção mostra as Permissões da LF-Tag ou as Permissões do valor da LF-Tag com base no Tipo de permissão que você escolheu na etapa anterior.

Dependendo do Tipo de permissão que você escolheu conceder, selecione as Permissões de LF-Tag ou as Permissões do par chave-valor da LF-Tag e as permissões concedidas.

1. Em Permissões de LF-Tag, selecione as permissões a serem concedidas.

Conceder Eliminar e Alterar concede implicitamente Descrever

Você precisa conceder as permissões Alterar e Eliminar em todos os valores da tag.

2. Em Permissões de valor-chave de LT\_Tag, selecione as permissões a serem concedidas.

Conceder Associar implicitamente concede Descrever. Escolha Conceder com expressão LF-Tag para permitir que o destinatário da concessão conceda ou revogue permissões de acesso aos recursos do Catálogo de dados usando o método LF-TBAC.

3. (Opcional) Em Permissões concedidas, selecione as permissões que o beneficiário do subsídio pode conceder a outros diretores em sua conta. AWS
4. Selecione Conceder.

### Conceder, revogar e listar permissões de tag LF usando o AWS CLI

Você pode conceder, revogar e listar permissões em LF-Tags usando o AWS Command Line Interface (AWS CLI).

#### Para listar as permissões do LF-Tag (AWS CLI)

- Insira um comando `list-permissions`. Você deve ser o criador da LF-Tag, administrador do data lake ou ter as permissões `Drop`, `Alter`, `Describe`, `Associate`, `Grant with LF-Tag permissions` em uma LF-Tag para vê-la.

O comando a seguir solicita todas as LF-Tags nas quais você tem permissões.

```
aws lakeformation list-permissions --resource-type LF_TAG
```

Veja a seguir um exemplo de saída para um administrador de data lake, que vê todas as LF-Tags concedidas a todas as entidades principais. Usuários não administrativos veem somente as LF-Tags concedidas a eles. As permissões de LF-Tags concedidas por uma conta externa aparecem em uma página de resultados separada. Para vê-los, repita o comando e forneça ao argumento `--next-token` o token retornado da execução anterior do comando.

```
{
  "PrincipalResourcePermissions": [
    {
      "Principal": {
        "DataLakePrincipalIdentifier": "arn:aws:iam::111122223333:user/
datalake_admin"
      },
      "Resource": {
        "LFTag": {
          "CatalogId": "111122223333",
          "TagKey": "environment",
          "TagValues": [
            "*"
          ]
        }
      },
      "Permissions": [
        "ASSOCIATE"
      ],
      "PermissionsWithGrantOption": [
        "ASSOCIATE"
      ]
    },
    {
      "Principal": {
        "DataLakePrincipalIdentifier": "arn:aws:iam::111122223333:user/
datalake_user1"
      },
      "Resource": {
        "LFTag": {
          "CatalogId": "111122223333",
          "TagKey": "module",
          "TagValues": [
            "Orders",
            "Sales"
          ]
        }
      }
    }
  ]
}
```



```

        }
      },
      "Permissions": [
        "DESCRIBE"
      ],
      "PermissionsWithGrantOption": []
    },
    ...
  ],
  "NextToken": "eyJzaG91bGRRdWVy...Wlzc2lvbnMiOnRydWV9"
}

```

Você pode listar todas as concessões para uma chave LF-tag específica. O comando a seguir retorna todas as permissões concedidas na LF-Tag `module`.

```
aws lakeformation list-permissions --resource-type LF_TAG --resource '{ "LFTag": {"CatalogId":"111122223333","TagKey":"module","TagValues":["*"]}}'
```

Você também pode listar os valores da LF-Tag concedidos a uma entidade principal específica para uma LF-Tag específica. Ao fornecer o argumento `--principal`, você deve fornecer o argumento `--resource`. Portanto, o comando só pode solicitar efetivamente os valores concedidos a uma entidade principal específica para uma chave LF-Tag específica. O comando a seguir mostra como fazer isso com a chave de entidade principal `datalake_user1` e a chave de LF-tag `module`.

```
aws lakeformation list-permissions --principal
DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/
datalake_user1 --resource-type LF_TAG --resource '{ "LFTag": {"CatalogId":"111122223333","TagKey":"module","TagValues":["*"]}}'
```

Esta é uma saída de exemplo.

```

{
  "PrincipalResourcePermissions": [
    {
      "Principal": {
        "DataLakePrincipalIdentifier": "arn:aws:iam::111122223333:user/
datalake_user1"
      },

```

```

    "Resource": {
      "LFTag": {
        "CatalogId": "111122223333",
        "TagKey": "module",
        "TagValues": [
          "Orders",
          "Sales"
        ]
      }
    },
    "Permissions": [
      "ASSOCIATE"
    ],
    "PermissionsWithGrantOption": []
  }
]
}

```

### Para conceder permissões em LF-Tags (AWS CLI)

1. Digite um comando semelhante ao seguinte: Este exemplo concede ao usuário `datalake_user1` a permissão `Associate` na LF-Tag com a chave `module`. Ele concede permissões para visualizar e atribuir todos os valores dessa chave, conforme indicado pelo asterisco (\*).

```

aws lakeformation grant-permissions --principal
  DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/
  datalake_user1 --permissions "ASSOCIATE" --resource '{ "LFTag":
  {"CatalogId":"111122223333","TagKey":"module","TagValues":["*"]}}'

```

Conceder a permissão `Associate` implicitamente concede a permissão `Describe`.

O próximo exemplo concede `Associate` à AWS conta externa `1234-5678-9012` na etiqueta LF com a chave, com a opção de concessão. `module` Ele concede permissões para visualizar e atribuir somente os valores `sales` e `orders`.

```

aws lakeformation grant-permissions --principal
  DataLakePrincipalIdentifier=123456789012 --permissions "ASSOCIATE"
  --permissions-with-grant-option "ASSOCIATE" --resource '{ "LFTag":
  {"CatalogId":"111122223333","TagKey":"module","TagValues":["sales", "orders"]}}'

```

2. Conceder a permissão `GrantWithLFTagExpression` implicitamente concede a permissão `Describe`.

O próximo exemplo concede `GrantWithLFTagExpression` a um usuário na LF-Tag com a chave `module`, com a opção de concessão. Ele concede permissões para visualizar e conceder permissões nos recursos do Catálogo de dados usando somente os valores `sales` e `orders`.

```
aws lakeformation grant-permissions --principal
DataLakePrincipalIdentifier=111122223333 --permissions "GrantWithLFTagExpression"
--permissions-with-grant-option "GrantWithLFTagExpression" --resource '{ "LFTag":
{"CatalogId":"111122223333","TagKey":"module","TagValues":["sales", "orders"]}'
```

3. O próximo exemplo concede permissões `Drop` a um usuário na LF-Tag com a chave `module`, com a opção de concessão. Ele concede permissões para excluir a LF-Tag. Para excluir uma LF-Tag, você precisa de permissões em todos os valores dessa chave.

```
aws lakeformation grant-permissions --principal
DataLakePrincipalIdentifier=111122223333 --permissions "DROP"
--permissions-with-grant-option "DROP" --resource '{ "LFTag":
{"CatalogId":"111122223333","TagKey":"module","TagValues":["*"]}'
```

4. O próximo exemplo concede permissões `Alter` ao usuário na LF-Tag com a chave `module`, com a opção de concessão. Ele concede permissões para excluir a LF-Tag. Para atualizar uma LF-Tag, você precisa de permissões em todos os valores dessa chave.

```
aws lakeformation grant-permissions --principal
DataLakePrincipalIdentifier=111122223333 --permissions "ALTER"
--permissions-with-grant-option "ALTER" --resource '{ "LFTag":
{"CatalogId":"111122223333","TagKey":"module","TagValues":["*"]}'
```

#### Para revogar permissões em LF-tags (AWS CLI)

- Digite um comando semelhante ao seguinte: Este exemplo revoga a permissão `Associate` na LF-Tag com a chave `module` do usuário `datalake_user1`.

```
aws lakeformation revoke-permissions --principal
DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/
datalake_user1 --permissions "ASSOCIATE" --resource '{ "LFTag":
{"CatalogId":"111122223333","TagKey":"module","TagValues":["*"]}'
```

## Conceder permissões de data lake usando o método LF-TBAC

É possível conceder as permissões DESCRIBE e ASSOCIATE do Lake Formation em tags do LF a entidades principais para que elas possam visualizar as tags do LF e atribuí-las aos recursos do catálogo de dados (bancos de dados, tabelas, visualizações e colunas). Quando as LF-Tags são atribuídas aos recursos do Catálogo de dados, você pode usar o método de controle de acesso baseado em tags do Lake Formation (LF-TBAC) para proteger esses recursos. Para obter mais informações, consulte [Controle de acesso baseado em tags do Lake Formation](#).

No início, somente o administrador do data lake pode conceder essas permissões. Se o administrador do data lake conceder essas permissões com a opção de concessão, outras entidades principais poderão concedê-las. As permissões DESCRIBE e ASSOCIATE são explicadas em [Considerações e práticas recomendadas de controle de acesso com base em tags do Lake Formation](#).

Você pode conceder as permissões DESCRIBE e ASSOCIATE em uma LF-Tag para uma conta externa AWS. Um administrador de data lake nessa conta pode então conceder essas permissões a outras entidades principais na conta. As entidades principais às quais o administrador do data lake na conta externa concede a permissão ASSOCIATE podem então atribuir LF-Tags aos recursos do Catálogo de dados que você compartilhou com a conta deles.

Ao conceder para uma conta externa, você deve incluir a opção de concessão.

Você pode conceder permissões em LF-Tags usando o console AWS Lake Formation, a API ou o AWS Command Line Interface (AWS CLI).

### Tópicos

- [Conceder permissões do catálogo de dados](#)

#### Consulte também

- [Concedendo, revogando e listando permissões de valor da LF-Tag](#)
- [Gerenciar LF-Tags para controle de acesso a metadados](#)
- [Controle de acesso baseado em tags do Lake Formation](#)

## Conceder permissões do catálogo de dados

Use o console do Lake Formation ou a AWS CLI para conceder permissões do Lake Formation em bancos de dados, tabelas, visualizações e colunas usando o método de controle de acesso baseado em tags do Lake Formation (LF-TBAC).

### Console

As etapas a seguir explicam como conceder permissões usando o método de controle de acesso baseado em tags do Lake Formation (LF-TBAC) e a página Conceder permissões de data lake no console do Lake Formation. A página está dividida nas seguintes seções:

- Entidades principais – Os usuários, perfis e Contas da AWS para os quais conceder permissões.
- LF-Tags ou recursos do catálogo – Os bancos de dados, tabelas ou links de recursos nos quais conceder permissões.
- Permissões – As permissões do Lake Formation devem ser concedidas.

1. Abra a página Conceder permissões de data lake.

Abra o console AWS Lake Formation em <https://console.aws.amazon.com/lakeformation/> e faça login como administrador do data lake ou como usuário que recebeu permissões do Lake Formation nos recursos do catálogo de dados por meio do LF-TBAC com a opção de concessão.

No painel de navegação, em Permissões, escolha Permissões do data lake. Em seguida, escolha Conceder.

2. Especifique as entidades principais.

Na seção Entidades principais, escolha um tipo de principal e, em seguida, especifique às entidades principais aos quais conceder permissões.

[AWS Lake Formation](#) > Grant permissions

## Grant data lake permissions

### Principals

Choose the principals to grant permissions.

**IAM users and roles**  
Users or roles from this AWS account.

**IAM Identity Center - new**  
Users and groups configured in IAM Identity Center.

**SAML users and groups**  
SAML users and group or QuickSight ARNs.

**External accounts**  
AWS account, AWS organization or IAM principal outside of this account

### Users and groups (3)

Choose users and groups to grant permissions.

Remove

Add

< 1 > ⚙

<input type="checkbox"/>	Name <a href="#">↗</a>	Type
<input type="checkbox"/>	<a href="#">DataStewards</a>	Group
<input type="checkbox"/>	<a href="#">user1</a>	User
<input type="checkbox"/>	<a href="#">user2</a>	User

### Usuários e perfis do IAM

Escolha um ou mais usuários ou perfis na lista de usuários e perfis do IAM.


### IAM Identity Center

Selecione um ou mais usuários na lista Usuários e grupos.

### Usuários e grupos SAML

Para usuários e grupos do SAML e do Amazon QuickSight, insira um ou mais nomes dos recursos da Amazon (ARNs) para usuários ou grupos federados por meio do SAML ou ARNs para usuários ou grupos do Amazon QuickSight. Pressione Enter após cada ARN.

Para obter informações sobre como construir os ARNs, consulte [Lake Formation concede e revoga comandos AWS CLI](#).

 Note

A integração do Lake Formation com o Amazon QuickSight é compatível somente com o Amazon QuickSight Enterprise Edition.

## Contas externas

Para Contas da AWS, organização AWS ou Entidade principal do IAM, insira um ou mais IDs válidos Conta da AWS, IDs da organização, IDs da unidade organizacional ou ARN para o usuário ou o perfil do IAM. Pressione Enter após cada ID.

O ID da organização consiste em "o-" seguido por 10 a 32 letras minúsculas ou dígitos.

Um ID de unidade organizacional começa com "ou-" seguido de 4 a 32 letras minúsculas ou dígitos (o ID da raiz que contém a UO). Essa sequência é seguida por um segundo travessão "-" e 8 a 32 letras minúsculas ou dígitos adicionais.

### 3. Especifique as tags do LF.

Certifique-se de que a opção Recursos correspondidos por LF-Tags seja escolhida. Escolha Adicionar LF-Tag.

#### 1. Escolha uma chave e valores da LF-Tag.

Se você escolher mais de um valor, estará criando uma expressão de LF-Tags com um operador OR. Isso significa que, se algum dos valores da LF-Tag corresponder a uma LF-Tag atribuída a um recurso do Catálogo de dados, você receberá permissões sobre o recurso.

### LF-Tags or catalog resources

**Resources matched by LF-Tags (recommended)**  
Manage permissions indirectly for resources or data matched by a specific set of LF-Tags.

**Named data catalog resources**  
Manager permissions for specific databases or tables, in addition to fine-grained data access.

Key:

Values:

- Orders
- Sales
- Customers

2. (Opcional) Escolha Adicionar LF-Tag novamente para especificar outra LF-Tag.

Se você especificar mais de uma LF-Tag, estará criando uma expressão de tag LF com um operador AND. A entidade principal recebe permissões em um recurso do Catálogo de dados somente se o recurso tiver sido atribuído a uma LF-Tag correspondente para cada LF-Tag na expressão da LF-Tag.

4. Especifique as permissões.

Especificar as permissões a serem concedidas à entidade principal em recursos correspondentes do catálogo de dados. Recursos correspondentes são aqueles recursos atribuídos a LF-Tags que correspondem a uma das expressões de LF-Tags concedidas à entidade principal.

É possível especificar as permissões a serem concedidas em bancos de dados, tabelas e visualizações correspondentes.



**▼ Database permissions**

**Database permissions**  
Choose specific access permissions to grant.

Create table    Alter    Drop  
 Describe

**Grantable permissions**  
Choose the permission that may be granted to others.

Create table    Alter    Drop  
 Describe

Super  
 This permission is the union of all the individual permissions to the left, and supersedes them.

**▼ Table permissions**

**Table permissions**  
Choose specific access permissions to grant.

Alter    Insert    Drop  
 Delete    Select    Describe

**Grantable permissions**  
Choose the permission that may be granted to others.

Alter    Insert    Drop  
 Delete    Select    Describe

Super  
 This permission is the union of all the individual permissions to the left, and supersedes them.

Em Permissões do banco de dados, selecione as permissões do banco de dados a serem concedidas à entidade principal nos bancos de dados correspondentes.

Em Permissões de tabela, selecione as permissões de tabela ou visualização a serem concedidas à entidade principal nas tabelas e nas visualizações correspondentes.

Também é possível selecionar as permissões `Select`, `Describe` e `Drop` nas Permissões de tabela a serem aplicadas às visualizações.

## 5. Selecione Conceder.

### AWS CLI

Você pode usar o AWS Command Line Interface (AWS CLI) e o método de controle de acesso baseado em tags do Lake Formation (LF-TBAC) para conceder permissões do Lake Formation em bancos de dados, tabelas e colunas do Catálogo de dados.

Conceder permissões de data lake usando a AWS CLI e o método LF-TBAC

- Use o comando `grant-permissions`.

## Example

O exemplo a seguir concede a expressão LF-Tag "module=\*" (todos os valores da chave LF-Tag module ) ao usuário `datalake_user1`. Esse usuário terá a permissão `CREATE_TABLE` em todos os bancos de dados correspondentes – bancos de dados aos quais foi atribuída a LF-Tag com a chave `module`, com qualquer valor.

```
aws lakeformation grant-permissions --principal
DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/
datalake_user1 --permissions "CREATE_TABLE" --resource '{ "LFTagPolicy":
{"CatalogId":"111122223333","ResourceType":"DATABASE","Expression":
[{"TagKey":"module","TagValues":["*"]}]}'
```

## Example

O próximo exemplo concede a expressão da LF-Tag "(level=director) AND (region=west OR region=south)" ao usuário `datalake_user1`. Esse usuário terá as permissões `SELECT`, `ALTER` e `DROP` com a opção de concessão em tabelas correspondentes – tabelas que foram atribuídas tanto a `level=director` quanto a `region=west` ou `region=south`.

```
aws lakeformation grant-permissions --principal
DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/
datalake_user1 --permissions "SELECT" "ALTER" "DROP" --permissions-
with-grant-option "SELECT" "ALTER" "DROP" --resource '{ "LFTagPolicy":
{"CatalogId":"111122223333","ResourceType":"TABLE","Expression": [{"TagKey":
"level","TagValues": ["director"]}, {"TagKey": "region","TagValues": ["west",
"south"]}]}'
```

## Example

O próximo exemplo concede a expressão LF-tag "module=orders" à conta AWS 1234-5678-9012. O administrador do data lake nessa conta pode então conceder a expressão "module=orders" às entidades principais em sua conta. Essas entidades principais terão então a permissão `CREATE_TABLE` para combinar bancos de dados pertencentes à conta 1111-2222-3333 e compartilhados com a conta 1234-5678-9012 usando o método de recurso nomeado ou o método LF-TBAC.

```
aws lakeformation grant-permissions --principal
DataLakePrincipalIdentifier=123456789012 --permissions "CREATE_TABLE" --
permissions-with-grant-option "CREATE_TABLE" --resource '{ "LFTagPolicy":
{"CatalogId":"111122223333","ResourceType":"DATABASE","Expression":
[{"TagKey":"module","TagValues":["orders"]}]}'
```

## Exemplo de cenário de permissões

O cenário a seguir ajuda a demonstrar como você pode configurar permissões para proteger o acesso aos dados em AWS Lake Formation.

Shirley é administradora de dados. Ela quer criar um data lake para sua empresa, a AnyCompany. Atualmente, todos os dados são armazenados no Amazon S3. John é gerente de marketing e precisa ter acesso por escrito às informações de compra do cliente (contidas em `s3://customerPurchases`). Um analista de marketing, Diego, se junta a John neste verão. John precisa da capacidade de conceder acesso a Diego para realizar consultas nos dados sem envolver Shirley.

Mateo, do setor financeiro, precisa acessar dados contábeis de consulta (por exemplo, `s3://transactions`). Ele quer consultar os dados das transações em tabelas em um banco de dados (`Finance_DB`) que a equipe financeira usa. Seu gerente, Arnav, pode dar a ele acesso ao `Finance_DB`. Embora ele não deva ser capaz de modificar os dados contábeis, ele precisa converter os dados em um formato (esquema) adequado para previsões. Esses dados serão armazenados em um bucket separado (`s3://financeForecasts`) que ele pode modificar.

Para resumir:

- Shirley é a administradora do data lake.
- John exige permissão `CREATE_DATABASE` e `CREATE_TABLE` para criar novos bancos de dados e tabelas no Catálogo de dados.
- John também exige `SELECT`, `INSERT` e `DELETE` permissões nas tabelas que ele cria.
- Diego exige permissão `SELECT` na tabela para executar consultas.

Os funcionários da AnyCompany realizam as seguintes ações para configurar as permissões. As operações de API mostradas nesse cenário mostram uma sintaxe simplificada para maior clareza.

1. Shirley registra o caminho do Amazon S3 contendo informações de compra do cliente no Lake Formation.

```
RegisterResource(ResourcePath("s3://customerPurchases"), false, Role_ARN )
```

2. Shirley concede a John o acesso ao caminho do Amazon S3 contendo informações de compra do cliente.

```
GrantPermissions(John, S3Location("s3://customerPurchases"),  
[DATA_LOCATION_ACCESS]) )
```

3. Shirley concede permissão a John para criar bancos de dados.

```
GrantPermissions(John, catalog, [CREATE_DATABASE])
```

4. John cria o banco de dados John\_DB. John tem permissão CREATE\_TABLE automática nesse banco de dados porque ele o criou.

```
CreateDatabase(John_DB)
```

5. John cria a tabela John\_Table apontando para s3://customerPurchases. Como ele criou a tabela, ele tem todas as permissões nela e pode conceder permissões sobre ela.

```
CreateTable(John_DB, John_Table)
```

6. John permite que seu analista, Diego, tenha acesso à tabela John\_Table.

```
GrantPermissions(Diego, John_Table, [SELECT])
```

7. John permite que seu analista, Diego, acesse o s3://customerPurchases/London/. Como Shirley já está registrada s3://customerPurchases, suas subpastas são registradas no Lake Formation.

```
GrantDataLakePrivileges( 123456789012/datalake, Diego, [DATA_LOCATION_ACCESS], [],  
S3Location("s3://customerPurchases/London/") )
```

8. John permite que seu analista, Diego, crie tabelas no banco de dados John\_DB.

```
GrantDataLakePrivileges( 123456789012/datalake, Diego, John_DB, [CREATE_TABLE],  
[] )
```

9. Diego cria uma tabela John\_DB em s3://customerPurchases/London/ e obtém automaticamente as permissões ALTER, DROP, SELECT, INSERT e DELETE.

```
CreateTable( 123456789012/datalake, John_DB, Diego_Table )
```

## Filtragem de dados e segurança em nível de célula no Lake Formation

Ao conceder permissões do Lake Formation em uma tabela do Catálogo de Dados, você pode incluir especificações de filtragem de dados para restringir o acesso a determinados dados nos resultados da consulta e nos mecanismos integrados ao Lake Formation. O Lake Formation usa a filtragem de dados para obter segurança em nível de coluna, segurança em nível de linha e segurança de nível de célula. Será possível definir e aplicar filtros de dados em colunas aninhadas se os dados de origem contiverem estruturas aninhadas.

### Tópicos

- [Visão geral da filtragem de dados](#)
- [Filtros de dados no Lake Formation](#)
- [Suporte partiQL em expressões de filtro de linha](#)
- [Notas e restrições para filtragem em nível de coluna](#)
- [Permissões necessárias para consultar tabelas com filtragem em nível de célula](#)
- [Como gerenciar filtros de dados](#)

## Visão geral da filtragem de dados

Com os recursos de filtragem de dados do Lake Formation, você pode implementar os seguintes níveis de segurança de dados.

### Segurança em nível de coluna

Conceder permissões em uma tabela do catálogo de dados com segurança em nível de coluna (filtragem de colunas) permite que os usuários visualizem somente colunas específicas e colunas aninhadas às quais eles têm acesso na tabela. Considere uma tabela do `persons` usada em vários aplicativos para uma grande empresa de comunicações multirregional. A concessão de permissões em tabelas do Catálogo de Dados com filtragem de colunas pode impedir que usuários que não trabalham no departamento de RH vejam informações de identificação pessoal (PII), como número

de previdência social ou data de nascimento. Também é possível definir políticas de segurança e conceder acesso somente a subestruturas parciais de colunas aninhadas.

### Segurança no nível da linha

A concessão de permissões em uma tabela do Catálogo de Dados com segurança em nível de linha (filtragem de linha) permite que os usuários visualizem somente linhas específicas de dados às quais eles têm acesso na tabela. A filtragem é baseada nos valores de uma ou mais colunas. É possível incluir estruturas de colunas aninhadas ao definir expressões de filtro de linha. Por exemplo, se diferentes escritórios regionais da empresa de comunicações tiverem seus próprios departamentos de RH, você poderá limitar os registros pessoais que os funcionários de RH podem ver a apenas registros de funcionários em sua região.

### Segurança em nível de célula

A segurança em nível de célula combina filtragem de linhas e filtragem de colunas para um modelo de permissões altamente flexível. Se você exibir as linhas e colunas de uma tabela como uma grade, usando a segurança em nível de célula, poderá restringir o acesso a elementos individuais (células) da grade em qualquer lugar nas duas dimensões. Ou seja, você pode restringir o acesso a diferentes colunas, dependendo da linha. Veja o diagrama a seguir, no qual as colunas restritas são sombreadas.

	Col1	Col2	Col3	Col4	Col5	Col6
Row1						
Row2						
Row3						
Row4						
Row5						

Continuando com o exemplo da tabela de pessoas, você pode criar um filtro de dados no nível da célula que restringe o acesso à coluna de endereço da rua se a linha tiver a coluna do país definida como “Reino Unido”, mas permite o acesso à coluna do endereço da rua se a linha tiver a coluna do país definida como “EUA”.

Os filtros se aplicam somente a operações de leitura. Portanto, você pode conceder somente a permissão SELECT do Lake Formation com filtros.

### Segurança em nível de célula em colunas aninhadas

O Lake Formation permite definir e aplicar filtros de dados com segurança em nível de célula em colunas aninhadas. No entanto, os mecanismos analíticos integrados, como Amazon Athena, Amazon EMR e Amazon Redshift Spectrum, aceitam a execução de consultas em tabelas aninhadas gerenciadas pelo Lake Formation com segurança em nível de linha e coluna.

Para obter limitações, consulte [Limitações de filtragem de dados](#).

## Filtros de dados no Lake Formation

É possível implementar segurança em nível de coluna, de linha e de célula ao criar filtros de dados. Você seleciona um filtro de dados ao conceder permissão SELECT ao Lake Formation nas tabelas. Se a tabela contiver estruturas de colunas aninhadas, será possível definir um filtro de dados incluindo ou excluindo as colunas secundárias e definir expressões de filtro em nível de linha em atributos aninhados.

Cada filtro de dados pertence a uma tabela específica em seu catálogo de dados. Um filtro de dados inclui as seguintes informações:

- Nome do filtro
- Os IDs de catálogo da tabela associada ao filtro
- Nome da tabela
- Nome do banco de dados que contém a tabela
- Especificação de coluna: uma lista de colunas e colunas aninhadas (com tipos de dados `struct`) a serem incluídas ou excluídas dos resultados da consulta.
- Expressão de filtro de linha — uma expressão que especifica as linhas a serem incluídas nos resultados da consulta. Com algumas restrições, a expressão tem a sintaxe de uma cláusula `WHERE` na linguagem `partiQL`. Para especificar todas as linhas, selecione Acesso a todas as linhas em Acesso em nível de linha no console ou use `AllRowsWildcard` em chamadas de API.

Para obter mais informações sobre o que é suportado em expressões de filtro de linha, consulte [Suporte partiQL em expressões de filtro de linha](#).

O nível de filtragem obtido depende de como você preenche o filtro de dados.

- Ao especificar o curinga “todas as colunas” e fornecer uma expressão de filtro de linha, você está estabelecendo apenas a segurança no nível da linha (filtragem de linhas).
- Ao incluir ou excluir colunas específicas e colunas aninhadas e especificar “todas as linhas” utilizando o curinga de todas as linhas, você está estabelecendo somente a segurança em nível de coluna (filtragem de colunas).
- Quando você inclui ou exclui colunas específicas e também fornece uma expressão de filtro de linha, está estabelecendo a segurança no nível da célula (filtragem de células).

A captura de tela a seguir do console do Lake Formation mostra um filtro de dados que realiza a filtragem em nível de célula. Para consultas na tabela `orders`, ela restringe o acesso à coluna `customer_name` e os resultados da consulta retornam somente as linhas em que a coluna `product_type` contém “pharma”.



## Create data filter



### Data filter name

Enter a name that describes this data access filter.

Name may contain letters (A-Z), numbers (0-9), hyphens (-), or under-scores (\_), and be less than 256 characters.

### Target database

Select the database that contains the target table.



### Target table

Select the table for which the data filter will be created.



### Column-level access

Choose whether this filter should have column-level restrictions.

- Access to all columns  
Filter won't have any column restrictions.
- Include columns  
Filter will only allow access to specific columns.
- Exclude columns  
Filter will allow access to all but specific columns.

### Select columns



Observe o uso de aspas simples para incluir o literal da string de caracteres, 'pharma'.

Você pode usar o console do Lake Formation para criar esse filtro de dados ou fornecer o seguinte objeto de solicitação para a operação `CreateDataCellsFilter` da API.

```
{
  "Name": "restrict-pharma",
  "DatabaseName": "sales",
  "TableName": "orders",
  "TableCatalogId": "111122223333",
  "RowFilter": {"FilterExpression": "product_type='pharma'"},
  "ColumnWildcard": {
    "ExcludedColumnNames": ["customer_name"]
  }
}
```

É possível criar tantos filtros de dados quantos necessários para uma tabela. Para fazer isso, você precisa de permissão `SELECT` com a opção de concessão em uma tabela. Por padrão, os administradores do Data Lake têm permissão para criar filtros de dados em todas as tabelas dessa conta. Normalmente, você usa apenas um subconjunto dos filtros de dados possíveis ao conceder permissões na tabela a uma entidade principal. Por exemplo, você pode criar um segundo filtro de dados para a tabela `orders`, que é um filtro de dados somente de segurança de linha. Referindo-se à captura de tela anterior, você pode escolher a opção `Acesso a todas as colunas` e incluir uma expressão de filtro de linha de `product_type<>pharma`. O nome desse filtro de dados pode ser `no-pharma`. Ele restringe o acesso a todas as linhas que têm a coluna `product_type` definida como "pharma".

O objeto de solicitação para a operação `CreateDataCellsFilter` da API desse filtro de dados é o seguinte.

```
{
  "Name": "no-pharma",
  "DatabaseName": "sales",
  "TableName": "orders",
  "TableCatalogId": "111122223333",
  "RowFilter": {"FilterExpression": "product_type<>'pharma'"},
  "ColumnNames": ["customer_id", "customer_name", "order_num",
    "product_id", "purchase_date", "product_type",
    "product_manufacturer", "quantity", "price"]
}
```

Em seguida, você poderia conceder `SELECT` na tabela `orders` com o filtro de dados `restrict-pharma` a um usuário administrativo, e `SELECT` na tabela `orders` com o filtro de dados `no-pharma` a usuários não administrativos. Para usuários do setor de saúde, você concederia `SELECT` na tabela `orders` com acesso total a todas as linhas e colunas (sem filtro de dados), ou talvez com outro filtro de dados que restringe o acesso às informações de preços.

É possível incluir ou excluir colunas aninhadas ao especificar a segurança em nível de coluna e em nível de linha em um filtro de dados. No exemplo a seguir, o acesso ao campo `product.offer` é especificado usando nomes de colunas qualificados (entre aspas duplas). Isso é importante para campos aninhados, a fim de evitar a ocorrência de erros quando os nomes das colunas contêm caracteres especiais e para manter a compatibilidade com versões anteriores das definições de segurança em nível de coluna de nível superior.

```
{
  "Name": "example_dcf",
  "DatabaseName": "example_db",
  "TableName": "example_table",
  "TableCatalogId": "111122223333",
  "RowFilter": { "FilterExpression": "customer.customerName <> 'John'" },
  "ColumnNames": ["customer", "\"product\".\"offer\""]
}
```

#### Consulte também

- [Como gerenciar filtros de dados](#)

## Suporte partiQL em expressões de filtro de linha

Você pode estruturar expressões de filtro de linha usando um subconjunto de tipos de dados, operadores e agregações do PartiQL. O Lake Formation não permite nenhum perfil PartiQL padrão ou definida pelo usuário na expressão de filtro. Você pode usar operadores de comparação para comparar colunas com constantes (por exemplo, `views >= 10000`), mas não pode comparar colunas com outras colunas.

Uma expressão de filtro de linha pode ser uma expressão simples ou uma expressão composta. O tamanho total da expressão deve ter menos de 2048 caracteres.

## Expressões simples

Uma expressão simples terá o formato: `<column name > <comparison operator ><value >`

- Nome da coluna

Deve ser uma coluna de dados de nível superior, uma coluna de partição ou uma coluna aninhada presente no esquema da tabela e deve pertencer aos [Tipos de dados compatíveis](#) listados abaixo.

- Operador de comparação

Os seguintes operadores são compatíveis: `=`, `>`, `<`, `>=`, `<=`, `<>`, `!=`, `BETWEEN`, `IN`, `LIKE`, `NOT`, `IS [NOT] NULL`

- Todas as comparações de strings e correspondências de padrão LIKE diferenciam maiúsculas e minúsculas. Não é possível usar o operador `IS [NOT] NULL` em colunas de partição.

- Valor da coluna

O valor da coluna deve corresponder ao tipo de dados do nome da coluna.

## Expressão composta

Uma expressão composta terá o formato: `( <simple expression > ) <AND/OR > ( <simple expression > )`. Expressões compostas podem ser combinadas ainda mais usando operadores lógicos AND/OR.

## Tipos de dados compatíveis

Filtros de linha que se referem a uma tabela AWS Glue Data Catalog que contém tipos de dados incompatíveis resultarão em um erro. A seguir estão os tipos de dados compatíveis com colunas e constantes da tabela, que são mapeados para tipos de dados: Amazon Redshift

- `STRING`, `CHAR`, `VARCHAR`
- `INT`, `LONG`, `BIGINT`, `FLOAT`, `DECIMAL`, `DOUBLE`
- `BOOLEAN`
- `STRUCT`

Para obter mais informações sobre tipos de dados no Amazon Redshift, consulte [Tipos de dados](#) no Guia do desenvolvedor de banco de dados do Amazon Redshift.

## Expressões de filtro em linha

### Example

Veja a seguir exemplos de expressões de filtro de linha válidas para uma tabela com colunas: `country` (String), `id` (Long), `year` (partition column of type Integer), `month` (partition column of type Integer)

- `year > 2010 and country != 'US'`
- `(year > 2010 and country = 'US') or (month < 8 and id > 23)`
- `(country between 'Z' and 'U') and (year = 2018)`
- `(country like '%ited%') and (year > 2000)`

### Example

Veja a seguir exemplos de expressões de filtro de linha válidas para uma tabela com colunas aninhadas: `year > 2010 and customer.customerId <> 1`

Os campos aninhados em colunas de partição não devem ser referenciados ao definir expressões aninhadas em nível de linha.

As constantes de string devem estar entre aspas simples.

## Palavras-chave reservadas

Se sua expressão de filtro de linha contiver palavras-chave PartiQL, você receberá um erro de análise, pois os nomes das colunas podem entrar em conflito com as palavras-chave. Quando isso acontecer, substitua os nomes das colunas por aspas duplas. Alguns exemplos de palavras-chave reservadas são “primeiro”, “último”, “asc”, “ausente”. Consulte a especificação do PartiQL para obter uma lista de palavras-chave reservadas.

## Referência do PartiQL

Para obter mais informações sobre o PartiQL, consulte <https://partiql.org/>.

## Notas e restrições para filtragem em nível de coluna

Há três maneiras de especificar a filtragem de colunas:

- Ao usar filtros de dados, como descrito anteriormente.
- Usando filtragem de colunas simples ou filtragem de colunas aninhada.
- Usando tags.

A filtragem simples de colunas apenas especifica uma lista de colunas a serem incluídas ou excluídas. Tanto o console do Lake Formation quanto a API oferecem suporte do AWS CLI à filtragem simples de colunas. Para ver um exemplo, consulte [Grant with Simple Column Filtering](#).

As seguintes notas e restrições se aplicam à filtragem de colunas:

- As tarefas de ETL do AWS Glue oferecem suporte à filtragem de colunas somente usando filtros de dados (segurança em nível de célula). A tarefa falhará se a filtragem simples de colunas for aplicada a qualquer tabela referenciada pela tarefa. Se você quiser apenas filtragem de colunas, conceda acesso às tabelas usando filtros de dados e digite a expressão `true` do filtro de linha no console ou use `AllRowsWildcard` em suas chamadas de API.
- Para conceder com a opção `SELECT` e a filtragem de colunas, você deve usar uma lista de inclusão, não uma lista de exclusão. Sem a opção de concessão, você pode usar listas de inclusão ou exclusão.
- Para conceder `SELECT` em uma tabela com filtragem de colunas, você deve ter recebido a opção de concessão `SELECT` na tabela e sem nenhuma restrição de linha. Você deve ter acesso a todas as linhas.
- Se você conceder com a opção `SELECT` e a filtragem de colunas a uma entidade principal em sua conta, essa entidade principal deverá especificar a filtragem de colunas para as mesmas colunas ou um subconjunto das colunas concedidas ao conceder a outra entidade principal. Se você conceder com a opção `SELECT` e a filtragem de colunas a uma conta externa, o administrador do data lake na conta externa poderá conceder `SELECT` a todas as colunas a outra entidade principal em sua conta. No entanto, mesmo com todas as colunas com `SELECT`, esse entidade principal terá visibilidade somente nas colunas concedidas à conta externa.
- Você não pode aplicar a filtragem de colunas nas chaves de partição.
- Uma entidade principal com a permissão `SELECT` em um subconjunto de colunas em uma tabela não pode receber a permissão `ALTER`, `DROP`, `DELETE`, ou `INSERT` nessa tabela. Para uma entidade principal com a permissão `ALTER`, `DROP`, `DELETE`, ou `INSERT` em uma tabela, se você conceder a permissão `SELECT` com a filtragem de colunas, ela não terá efeito.

As seguintes notas e restrições se aplicam à filtragem de colunas aninhadas:

- É possível incluir ou excluir cinco níveis de campos aninhados em um filtro de dados.

#### Example

```
Col1.Col1_1.Col1_1_1.Col1_1_1_1.Col1_1_1_1_1
```

- Não é possível aplicar a filtragem de colunas em campos aninhados em colunas de partição.
- Se o esquema da tabela contiver um nome de coluna de nível superior ("customer"."address") que tenha o mesmo padrão de representação de campo aninhado em um filtro de dados (uma coluna aninhada com um nome de coluna de nível superior `customer` e um nome de campo aninhado `address` especificado como `"customer"."address"` em um filtro de dados), não será possível especificar explicitamente o acesso à coluna de nível superior nem ao campo aninhado porque ambos são representados usando o mesmo padrão nas listas de inclusão/exclusão. Isso é ambíguo, e o Lake Formation não poderá resolver se você estiver especificando a coluna de nível superior ou o campo aninhado.
- Se uma coluna de nível superior ou um campo aninhado contiver aspas duplas no nome, será necessário incluir uma segunda aspa dupla ao especificar o acesso a um campo aninhado na lista de inclusão e exclusão de um filtro de células de dados.

#### Example

Exemplo de nome de coluna aninhada com aspas duplas: `a.b.double"quote`

#### Example

Exemplo de representação de coluna aninhada em um filtro de dados:

```
"a"."b"."double""quote"
```

## Permissões necessárias para consultar tabelas com filtragem em nível de célula

As permissões AWS Identity and Access Management a seguir (IAM) são necessárias para executar consultas em tabelas com filtragem em nível de célula.

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

```
    "Action": [  
      "lakeformation:StartQueryPlanning",  
      "lakeformation:GetQueryState",  
      "lakeformation:GetWorkUnits",  
      "lakeformation:GetWorkUnitResults"  
    ],  
    "Resource": "*"    
  }  
]  
}
```

Para obter mais informações sobre as permissões de Lake Formation, consulte [Referência de personas e permissões do IAM do Lake Formation](#).

## Como gerenciar filtros de dados

Para implementar a segurança em nível de coluna, de linha e de célula, é possível criar e manter filtros de dados. Cada filtro de dados pertence a uma tabela do Catálogo de dados. Você pode criar vários filtros de dados para uma tabela e, em seguida, usar um ou mais deles ao conceder permissões na tabela. Também é possível definir e aplicar filtros de dados em colunas aninhadas que têm tipos de dados `struct` que permitem aos usuários acessar somente subestruturas de colunas aninhadas.

Você precisa da permissão `SELECT` com a opção de concessão para criar ou visualizar um filtro de dados. Para permitir que as entidades principais da sua conta visualizem e usem um filtro de dados, você pode conceder a permissão `DESCRIBE` sobre ele.

### Note

O Lake Formation não dá suporte à concessão da permissão `Describe` em um filtro de dados, que é compartilhado de outra conta.

Você pode gerenciar filtros de dados usando o console AWS Lake Formation, a API ou o AWS Command Line Interface (AWS CLI).

Para obter informações sobre filtros de dados, consulte [Filtros de dados no Lake Formation](#)

## Criando um filtro de dados

Você pode criar um ou mais filtros de dados para cada tabela do Catálogo de dados.



## Para criar um filtro de dados para uma tabela do Catálogo de dados (console)

1. Abra o console do Lake Formation em <https://console.aws.amazon.com/lakeformation/>.

Assine como administrador do data lake, proprietário da tabela de destino ou entidade principal que tenha uma permissão do Lake Formation na tabela de destino.

2. No painel de navegação, em Catálogo de dados, escolha Filtros de dados.
3. Na página Filtros de dados, escolha Criar novo filtro.
4. Na caixa de diálogo Criar filtro de dados, insira as seguintes informações:
  - Nome do filtro de dados
  - Banco de dados de destino – especifique o banco de dados que contém a tabela.
  - Tabela de destino
  - Acesso em nível de coluna – deixe essa opção definida como Acesso a todas as colunas para especificar somente a filtragem de linhas. Escolha Incluir colunas ou Excluir colunas para especificar a filtragem de colunas ou células e, em seguida, especifique as colunas a serem incluídas ou excluídas.

Colunas aninhadas: se você estiver aplicando o filtro em uma tabela que contenha colunas aninhadas, poderá especificar explicitamente as subestruturas das colunas da estrutura aninhada em um filtro de dados.

Ao conceder a permissão SELECT a uma entidade principal nesse arquivador, a entidade principal que executa a consulta a seguir verá apenas os dados de `customer.customerName` e não de `customer.customerId`.

```
SELECT "customer" FROM "example_db"."example_table";
```

### Column-level access

Choose whether this filter should have column-level restrictions.

#### Column-level access

Choose whether this filter should have column-level restrictions.

- Access to all columns  
Filter won't have any column restrictions.
- Include columns  
Filter will only allow access to specific columns.
- Exclude columns  
Filter will allow access to all but specific columns.

### Included columns (4/11)

Choose the columns for column-level access

< 1 >

	Name	Type
<input type="checkbox"/>	customer	struct
<input type="checkbox"/>	customerId	string
<input checked="" type="checkbox"/>	customerName	string
<input checked="" type="checkbox"/>	customerapplication	struct
<input type="checkbox"/>	appld	string
<input checked="" type="checkbox"/>	product	struct
<input type="checkbox"/>	offer	struct
<input type="checkbox"/>	listingId	string
<input type="checkbox"/>	prodId	string
<input type="checkbox"/>	type	string
<input checked="" type="checkbox"/>	purchaseid	string

### Row-level access

Choose whether this filter should have row-level restrictions.

- Access to all rows
- Filter rows

### Row filter expression

Enter the rest of the following query statement `SELECT * FROM nested-table WHERE...`  
Please see the documentation for examples of filter expressions.

`customer.customerName <> 'John'`

Ao conceder permissões à coluna `customer`, a entidade principal recebe o acesso à coluna e aos campos aninhados abaixo da coluna (`customerName` e `customerID`).

- Expressão de filtro de linha – insira uma expressão de filtro para especificar a filtragem de linha ou célula. Para obter os tipos de dados e operadores compatíveis, consulte [Suporte partiQL em expressões de filtro de linha](#). Selecione Acesso a todas as linhas para conceder acesso a todos.

É possível incluir estruturas de coluna parciais de colunas aninhadas em uma expressão de filtro de linha para filtrar linhas que contenham valores específicos.

Quando uma entidade principal recebe permissões para uma tabela com uma expressão `Select * from example_nesttable where customer.customerName <> 'John'` de filtro de linha e o acesso em nível de coluna é definido como Acesso a todas as colunas, os resultados da consulta mostram somente as linhas em que `customerName <> 'John'` é avaliado como verdadeiro.

A captura de tela a seguir mostra um filtro de dados que implementa a filtragem de células. Nas consultas com base na tabela `orders`, ela nega o acesso à coluna `customer_name` e mostra somente as linhas que têm “pharma” na coluna `product_type`.

## Create data filter



### Data filter name

Enter a name that describes this data access filter.

restrict-pharma

Name may contain letters (A-Z), numbers (0-9), hyphens (-), or under-scores (\_), and be less than 256 characters.

### Target database

Select the database that contains the target table.

Choose databases



Load more

sales



054881201579

### Target table

Select the table for which the data filter will be created.

Choose tables



Load more

orders



054881201579

### Column-level access

Choose whether this filter should have column-level restrictions.

- Access to all columns  
Filter won't have any column restrictions.
- Include columns  
Filter will only allow access to specific columns.
- Exclude columns  
Filter will allow access to all but specific columns.

### Select columns

Choose one or more columns



customer\_name  
string



## 5. Selecione Criar filtro.

Como criar um filtro de dados com políticas de filtro de células em um campo aninhado

Esta seção usa o seguinte exemplo de esquema para mostrar como criar um filtro de células de dados:

```
[
  { name: "customer", type: "struct<customerId:string,customerName:string>" },
  { name: "customerApplication", type: "struct<appId:string>" },
  { name: "product", type:
"struct<offer:struct<prodId:string,listingId:string>,type:string>" },
  { name: "purchaseId", type: "string" },
]
```

1. Na página Criar um filtro de dados, insira um nome para o filtro de dados.
2. Depois, use o menu suspenso para escolher o nome do banco de dados e o nome da tabela.
3. Na seção Acesso em nível de coluna, selecione Colunas incluídas e uma coluna aninhada (`customer.customerName`).
4. Na seção Acesso em nível de linha, selecione a opção Acesso a todas as linhas.
5. Selecione Criar filtro.

Ao conceder a permissão SELECT nesse filtro, a entidade principal obtém acesso a todas as linhas na coluna `customerName`.

6. Depois, defina outro filtro de dados para o mesmo banco de dados/tabela.
7. Na seção Acesso em nível de coluna, selecione Colunas incluídas e outra coluna aninhada (`customer.customerid`).
8. Na seção Acesso em nível de linha, escolha Filtrar linhas e insira uma Expressão de filtro de linha (`customer.customerid <> 5`).
9. Selecione Criar filtro.

Ao conceder a permissão SELECT nesse filtro, a entidade principal recebe acesso a todas as linhas no `customerName` e aos campos `customerid`, exceto à célula em que o valor é 5 na coluna `customerid`.

## Concedendo permissões de filtro de dados

Você pode conceder as permissões SELECT, DESCRIBE e DROP do Lake Formation sobre filtros de dados às entidades principais.

No início, somente você pode visualizar os filtros de dados que você cria para uma tabela. Para permitir que outra entidade principal visualize um filtro de dados e conceda permissões do Catálogo de dados com o filtro de dados, você deve:

- Conceder SELECT uma tabela à entidade principal com a opção de concessão e aplique o filtro de dados à concessão.
- Conceder a permissão DESCRIBE ou DROP no filtro de dados da entidade principal.

Você pode conceder a permissão SELECT a uma conta externa AWS. Um administrador do data lake nessa conta pode então conceder essa permissão a outras entidades principais da conta. Ao conceder a uma conta externa, você deve incluir a opção de concessão para que o administrador da conta externa possa transmitir ainda mais a permissão para outros usuários em sua conta. Ao conceder a uma entidade principal em sua conta, a concessão com a opção de concessão é opcional.

Você pode conceder e revogar permissões em filtros de dados usando o console AWS Lake Formation, a API ou o AWS Command Line Interface (AWS CLI).

### Console

1. Faça login no AWS Management Console e abra o console do Lake Formation em <https://console.aws.amazon.com/lakeformation/>.
2. No painel de navegação, em Permissões, escolha Permissões do data lake.
3. Na página Permissões, na seção Permissões de dados, escolha Conceder.
4. Na página Conceder permissões de dados, escolha as entidades principais aos quais conceder as permissões.
5. Na seção LF-Tags ou recursos do catálogo, escolha Recursos do Catálogo de dados nomeados. Em seguida, escolha o banco de dados, a tabela e o filtro de dados para os quais você deseja conceder permissões.

### LF-Tags or catalog resources

Resources matched by LF-Tags (recommended)  
Manage permissions indirectly for resources or data matched by a specific set of LF-Tags.

Named data catalog resources  
Manager permissions for specific databases or tables, in addition to fine-grained data access.

**Databases**  
Select one or more databases.

Choose databases ▼ Load more

cloudtrail X  
106567286946

**Tables - optional**  
Select one or more tables.

Choose tables ▼ Load more

cloudtrail\_logs\_awslogs X  
106567286946

**Data filters - optional**  
Select one or more data filters.

Choose data filters ▼ Load more Create new

cloudtrail\_lakeformation\_filter X  
106567286946

[Manage data filters](#) ↗

6. Na seção Permissões do filtro de dados, escolha as permissões que você deseja conceder às entidades principais selecionadas.

### Data filter permissions

**Data filter permissions**  
Choose specific access permissions to grant.

Select  Describe  Drop

**Grantable permissions**  
Choose the permission that may be granted to others.

Select  Describe  Drop

## AWS CLI

- Insira um comando `grant-permissions`. Especifique `DataCellsFilter` para o argumento `resource` e especifique `DESCRIBE` ou `DROP` para o argumento `Permissions` e, opcionalmente, para o argumento `PermissionsWithGrantOption`.

O exemplo a seguir concede a `DESCRIBE` a opção de concessão ao usuário `datalake_user1` no filtro de dados `restrict-pharma`, que pertence à tabela `orders` no banco de dados `sales` na conta AWS `1111-2222-3333`.

```
aws lakeformation grant-permissions --cli-input-json file://grant-params.json
```

A seguir está o conteúdo do arquivo `grant-params.json`.

```
{
  "Principal": {"DataLakePrincipalIdentifier":
    "arn:aws:iam::111122223333:user/datalake_user1"},
  "Resource": {
    "DataCellsFilter": {
      "TableCatalogId": "111122223333",
      "DatabaseName": "sales",
      "TableName": "orders",
      "Name": "restrict-pharma"
    }
  },
  "Permissions": ["DESCRIBE"],
  "PermissionsWithGrantOption": ["DESCRIBE"]
}
```

## Conceder permissões de dados fornecidas por filtros de dados

Os filtros de dados representam um subconjunto de dados em uma tabela. Para fornecer acesso aos dados às entidades principais, as permissões `SELECT` precisam ser concedidas a essas entidades principais. Com essa permissão, as entidades principais podem:

- Veja o nome real da tabela na lista de tabelas compartilhadas com suas contas.
- Crie filtros de dados na tabela compartilhada e conceda permissões a seus usuários nesses filtros de dados.



## Console

Para conceder permissões SELECIONAR

1. Acesse a página Permissões no console do Lake Formation e escolha Conceder.

AWS Lake Formation > Permissions

Too many permissions? Filter by database or table. In the navigation page, choose **Databases** or **Tables**. Then choose a database or table, and on the **Actions** menu, choose **View Permissions**.

**Data permissions** ↻ Revoke Grant

🔍 *Filter permissions by property or value* < 1 ... > ⚙️

Principal ▲ Principal type ▼ Resource type ▼ Database ▼ Table ▼ Resource ▼ Catalog ▼

2. Selecione as entidades principais às quais você deseja fornecer acesso e selecione Recursos do catálogo de dados nomeados.

### LF-Tags or catalog resources

Resources matched by LF-Tags (recommended)  
Manage permissions indirectly for resources or data matched by a specific set of LF-Tags.

Named data catalog resources  
Manager permissions for specific databases or tables, in addition to fine-grained data access.

**Databases**  
Select one or more databases.

Choose databases ▼

Load more

cloudtrail X  
106567286946

**Tables - optional**  
Select one or more tables.

Choose tables ▼

Load more

cloudtrail\_logs\_awslogs X  
106567286946

**Data filters - optional**  
Select one or more data filters.

Choose data filters ▼

Load more

Create new

cloudtrail\_lakeformation\_filter X  
106567286946

[Manage data filters](#) ↗

3. Para fornecer acesso aos dados que o filtro representa, escolha Selecionar em Permissões do filtro de dados.


### Data filter permissions

**Data filter permissions**  
Choose specific access permissions to grant.

Select     Describe     Drop

**Grantable permissions**  
Choose the permission that may be granted to others.

Select     Describe     Drop

 Select permissions on data filters will grant access to the table 'cloudtrail\_logs\_awslogs'.

## CLI

Insira um comando `grant-permissions`. Especifique `DataCellsFilter` para o argumento do recurso e especifique `SELECT` para o argumento `Permissões`.

O exemplo a seguir concede `SELECT` a opção de concessão ao usuário `datalake_user1` no filtro de dados `restrict-pharma`, que pertence à tabela `orders` no banco de dados `sales` em Conta da AWS `1111-2222-3333`.

```
aws lakeformation grant-permissions --cli-input-json file://grant-params.json
```

A seguir está o conteúdo do arquivo `grant-params.json`.

```
{
  "Principal": {
    "DataLakePrincipalIdentifier": "arn:aws:iam::111122223333:user/
datalake_user1"
  },
  "Resource": {
    "DataCellsFilter": {
      "TableCatalogId": "111122223333",
      "DatabaseName": "sales",
      "TableName": "orders",
      "Name": "restrict-pharma"
    }
  },
}
```

```
"Permissions": ["SELECT"]
}
```

## Visualizando filtros de dados

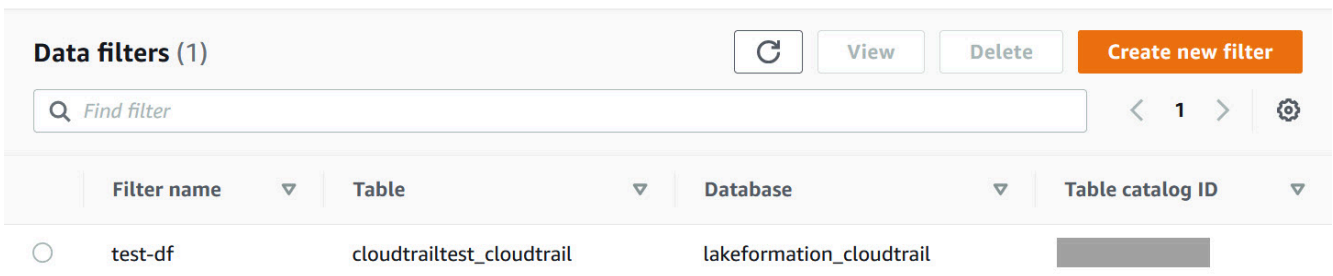
Você pode usar o console do Lake Formation, AWS CLI, ou a API do Lake Formation para visualizar os filtros de dados.

Para visualizar os filtros de dados, você deve ser administrador do data lake ou ter as permissões necessárias nos filtros de dados.

### Console

1. Faça login no AWS Management Console e abra o console do Lake Formation em <https://console.aws.amazon.com/lakeformation/>.
2. No painel de navegação, em Catálogo de dados, escolha Filtros de dados.

A página exibe os filtros de dados aos quais você tem acesso.



Filter name	Table	Database	Table catalog ID
test-df	cloudtrailtest_cloudtrail	lakeformation_cloudtrail	

3. Para visualizar os detalhes do filtro de dados, escolha o filtro de dados e, em seguida, escolha Exibir. Uma nova janela aparece com as informações detalhadas do filtro de dados.

**View data filter** [X]

Name  
test-df

---

Database  
lakeformation\_cloudtrail

Table  
cloudtrailtest\_cloudtrail

---

Column-level access  
Include

Row filter expression  
true

Columns  
eventversion, useridentity, eventtime,  
eventsource, eventname

Close

## AWS CLI

Insira um comando `list-data-cells-filter` e especifique um recurso de tabela.

O exemplo a seguir lista os filtros de dados para a tabela `cloudtrailtest_cloudtrail`.

```
aws lakeformation list-data-cells-filter --table '{"CatalogId":"123456789012",
"DatabaseName":"lakeformation_cloudtrail", "Name":"cloudtrailtest_cloudtrail"}
```

## API/SDK

Use a API `ListDataCellsFilter` e especifique um recurso de tabela.

O exemplo a seguir usa Python para listar os 20 primeiros filtros de dados da tabela `myTable`.

```
response = client.list_data_cells_filter(
    Table = {
        'CatalogId': '111122223333',
        'DatabaseName': 'mydb',
        'Name': 'myTable'
    },
```

```
MaxResults=20
```

```
)
```

## Listando permissões de filtro de dados

Você pode usar o console do Lake Formation para visualizar as permissões concedidas nos filtros de dados.

Para ver as permissões em um filtro de dados, você deve ser administrador do data lake ou ter as permissões necessárias no filtro de dados.

### Console

1. Faça login no AWS Management Console e abra o console do Lake Formation em <https://console.aws.amazon.com/lakeformation/>.
2. No painel de navegação, em Permissões, escolha Permissões de dados.
3. Na página Permissões de dados, clique ou toque no campo de pesquisa e, no menu Propriedades, escolha Tipo de recurso.
4. No menu Tipo de recurso, escolha Tipo de recurso: filtro de células de dados.

Os filtros de dados nos quais você tem permissões estão listados. Talvez seja necessário rolar horizontalmente para ver as colunas de Permissões e de Concessão.

Data Permissions (58)							
Principal	Resource type	Database	Table	Resource	Catalog	Permissions	
<input type="radio"/> datalake_admin	Data cell filter	sales	orders	no-pharma	111122223333	Describe, Drop, Select	
<input type="radio"/> datalake_admin	Data cell filter	sales	orders	restrict-pharma	111122223333	Describe, Drop, Select	
<input type="radio"/> datalake_user1	Data cell filter	sales	orders	restrict-pharma	111122223333	Describe	
<input type="radio"/> datalake_user2	Data cell filter	sales	orders	restrict-pharma	111122223333	Select	

### AWS CLI

- Insira um comando `list-permissions`. Especifique `DataCellsFilter` para o argumento `resource` e especifique `DESCRIBE` ou `DROP` para o argumento `Permissions` e, opcionalmente, para o argumento `PermissionsWithGrantOption`.

O exemplo a seguir lista as permissões DESCRIBE com a opção de concessão no filtro de dados `restrict-pharma`. Os resultados são limitados às permissões concedidas para a entidade principal `datalake_user1` e a tabela `orders` no banco de dados `sales` na conta AWS 1111-2222-3333.

```
aws lakeformation list-permissions --cli-input-json file://list-params.json
```


A seguir está o conteúdo do arquivo `grant-params.json`.

```
{
  "Principal": {"DataLakePrincipalIdentifier":
    "arn:aws:iam::111122223333:user/datalake_user1"},
  "Resource": {
    "DataCellsFilter": {
      "TableCatalogId": "111122223333",
      "DatabaseName": "sales",
      "TableName": "orders",
      "Name": "restrict-pharma"
    }
  },
  "Permissions": ["DESCRIBE"],
  "PermissionsWithGrantOption": ["DESCRIBE"]
}
```

## Visualizar permissões de banco de dados e tabelas no Lake Formation

Você pode visualizar as permissões do Lake Formation concedidas em um banco de dados ou uma tabela do Catálogo de Dados. Você pode gerenciar LF-Tags usando o console do Lake Formation, a API ou a AWS Command Line Interface (AWS CLI).

Usando o console, você pode visualizar as permissões nas páginas Bancos de dados, Tabelas ou Permissões de dados.

 Note

Se você não for administrador de banco de dados ou proprietário de recursos, poderá ver as permissões que outras entidades principais têm sobre o recurso somente se tiver uma permissão do Lake Formation sobre o recurso com a opção de concessão.


Além das permissões necessárias do Lake Formation, você precisa das permissões do AWS Identity and Access Management (IAM) `glue:GetDatabases`, `glue:GetDatabase`, `glue:GetTables`, `glue:GetTable` e `glue:ListPermissions`.

Para visualizar as permissões em um banco de dados (console, a partir da página Bancos de dados)

1. Abra o console do Lake Formation em <https://console.aws.amazon.com/lakeformation/>.

Faça login como administrador do data lake, criador do banco de dados ou como usuário que tenha qualquer permissão do Lake Formation no banco de dados com a opção de concessão.

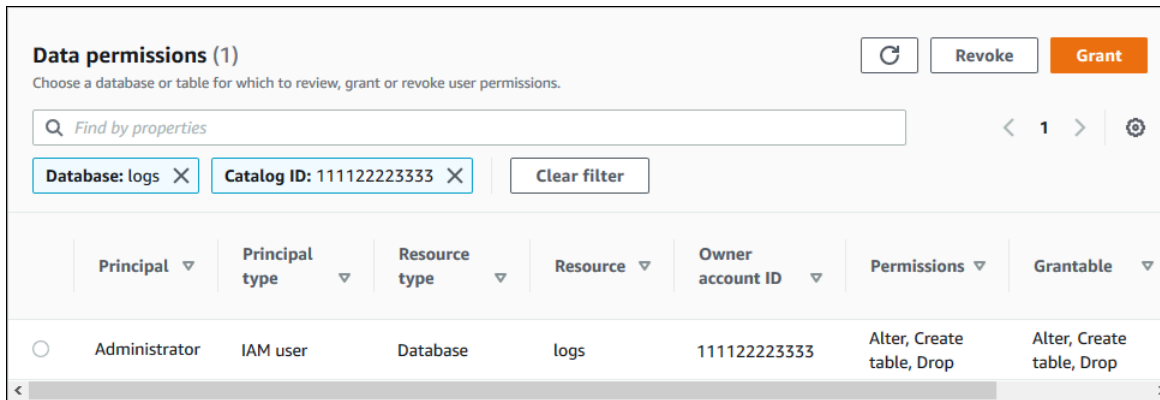
2. No painel de navegação, escolha Bancos de dados.
3. Escolha um banco de dados e, no menu Ações, escolha Exibir permissões.

 Note

Se você escolher um link de recurso de banco de dados, o Lake Formation exibirá as permissões no link do recurso, não no banco de dados de destino do link do recurso.

A página Permissões de dados lista todas as permissões do Lake Formation para o banco de dados. O nome do banco de dados e o ID do catálogo (ID da conta AWS) do proprietário do banco de dados aparecem como rótulos na caixa de pesquisa. Os mosaicos indicam que um filtro foi aplicado para listar permissões somente para esse banco de dados. Você pode ajustar o filtro fechando um quadro ou escolhendo Limpar filtro.





Para visualizar as permissões em um banco de dados (console, a partir da página Permissões de dados)

1. Abra o console do Lake Formation em <https://console.aws.amazon.com/lakeformation/>.

Faça login como administrador do data lake, criador do banco de dados ou como usuário que tenha qualquer permissão do Lake Formation no banco de dados com a opção de concessão.

2. No painel de navegação, selecione Permissões de dados.
3. Posicione o cursor na caixa de pesquisa na parte superior da página e, no menu Propriedades exibido, escolha Banco de dados.
4. No menu de bancos de dados exibido, escolha um banco de dados.

#### Note

Se você escolher um link de recurso de banco de dados, o Lake Formation exibirá as permissões no link do recurso, não no banco de dados de destino do link do recurso.

A página Permissões de dados lista todas as permissões do Lake Formation para o banco de dados. O nome do banco de dados aparece como um quadro abaixo da caixa de pesquisa. Os blocos indicam que um filtro foi aplicado para listar permissões somente para esse banco de dados. Você pode remover o filtro fechando o quadro ou escolhendo Limpar filtro.

Para visualizar as permissões em uma tabela (console, a partir da página Tabelas)

1. Abra o console do Lake Formation em <https://console.aws.amazon.com/lakeformation/>.

Faça login como administrador do data lake, criador da tabela ou como usuário que tenha qualquer permissão do Lake Formation na tabela com a opção de concessão.

2. No painel de navegação, selecione Tabelas.
3. Escolha uma tabela e, no menu Ações, selecione Exibir permissões.

### Note

Se você escolher um link de recurso de tabela, o Lake Formation exibirá as permissões no link do recurso, não na tabela de destino do link do recurso.

A página Permissões de dados lista todas as permissões do Lake Formation para a tabela. O nome da tabela, o nome do banco de dados que contém a tabela e o ID do catálogo (ID da conta AWS) do proprietário do banco de dados aparecem como rótulos na caixa de pesquisa. As etiquetas indicam que um filtro foi aplicado para listar permissões apenas para essa tabela. Você pode ajustar o filtro fechando uma etiqueta ou escolhendo Limpar filtro.

Principal	Principal type	Resource type	Resource	Owner account ID	Permissions	Grantable
Administrator	IAM user	Table	alexa-logs	111122223333	Super	Super

Para visualizar as permissões em uma tabela (console, a partir da página Permissões de dados)

1. Abra o console do Lake Formation em <https://console.aws.amazon.com/lakeformation/>.

Faça login como administrador do data lake, criador do banco de dados ou como usuário que tenha qualquer permissão do Lake Formation no banco de dados com a opção de concessão.

2. No painel de navegação, selecione Permissões de dados.
3. Posicione o cursor na caixa de pesquisa na parte superior da página e, no menu Propriedades exibido, escolha Banco de dados.
4. No menu de Bancos de dados exibido, escolha um banco de dados.

**⚠ Important**

Se você quiser ver as permissões em uma tabela que foi compartilhada com sua conta AWS a partir de uma conta externa, você deve escolher o banco de dados na conta externa que contém a tabela, não um link de recurso para o banco de dados.

A página Permissões de dados lista todas as permissões do Lake Formation para o banco de dados.

5. Posicione o cursor na caixa de pesquisa na parte superior da página e, no menu Propriedades exibido, escolha Tabelas.
6. No menu Tabelas exibido, escolha uma tabela.

A página Permissões de dados lista todas as permissões do Lake Formation para o banco de dados. O nome da tabela e o nome do banco de dados que contém a tabela aparecem como blocos sob a caixa de pesquisa. Os mosaicos indicam que um filtro foi aplicado para listar permissões somente para essa tabela. Você pode ajustar o filtro fechando um quadro ou escolhendo Limpar filtro.

Para ver as permissões em uma tabela (AWS CLI)

- Insira um comando `list-permissions`.

O exemplo a seguir lista as permissões em uma tabela compartilhada de uma conta externa. A propriedade `CatalogId` é o ID da conta AWS da conta externa, e o nome do banco de dados se refere ao banco de dados na conta externa que contém a tabela.

```
aws lakeformation list-permissions --resource-type TABLE --resource '{ "Table":  
{"DatabaseName":"logs", "Name":"alex-log", "CatalogId":"123456789012"} }'
```

## Revogando a permissão usando o console Lake Formation

Você pode usar o console para revogar todos os tipos de permissões do Lake Formation – permissões de Catálogo de dados, permissões de etiquetas de políticas, permissões de filtro de dados e permissões de localização.

Para revogar as permissões do Lake Formation em um recurso (console)

1. Abra o console do Lake Formation em <https://console.aws.amazon.com/lakeformation/>.

Faça login como administrador do data lake ou como usuário que recebeu permissões com a opção de concessão no recurso.

2. No painel de navegação, em Permissões, escolha Permissões de data lake, Tags do LF e permissões ou Localizações de dados.
3. Selecione a permissão ou o local e, em seguida, escolha Revogar.
4. Na caixa de diálogo exibida, escolha Revogar.

## Compartilhamento de dados entre contas no Lake Formation

Os recursos de várias contas do Lake Formation permitem que os usuários compartilhem com segurança lagos de dados distribuídos em várias AWS organizações ou diretamente com os diretores do IAM em outra conta Contas da AWS, fornecendo acesso refinado aos metadados do Catálogo de Dados e aos dados subjacentes. As grandes empresas geralmente usam várias Contas da AWS, e muitas dessas contas podem precisar acessar um data lake gerenciado por uma única Conta da AWS. Os usuários e as tarefas de AWS Glue extração, transformação e carregamento (ETL) podem consultar e unir tabelas em várias contas e ainda aproveitar as proteções de dados em nível de tabela e coluna do Lake Formation.

Quando você concede permissões do Lake Formation em um recurso do Catálogo de Dados para uma conta externa ou diretamente para um diretor do IAM em outra conta, o Lake Formation usa o serviço AWS Resource Access Manager (AWS RAM) para compartilhar o recurso. Se a conta do concedido estiver na mesma organização da conta do concedente, o recurso compartilhado estará disponível imediatamente para o concedido. Se a conta do beneficiário não estiver na mesma organização, AWS RAM envia um convite à conta do beneficiário para aceitar ou rejeitar a concessão do recurso. Em seguida, para disponibilizar o recurso compartilhado, o administrador do data lake na conta do beneficiário deve usar o AWS RAM console ou AWS CLI aceitar o convite.

O Lake Formation permite o compartilhamento de recursos do Catálogo de Dados com contas externas no modo de acesso híbrido. O modo de acesso híbrido oferece a flexibilidade de habilitar seletivamente as permissões do Lake Formation para bancos de dados e tabelas em seu AWS Glue Data Catalog.

Com o modo de acesso híbrido, agora você tem um caminho incremental que permite definir permissões do Lake Formation para um conjunto específico de usuários sem interromper as políticas de permissão de outros usuários ou workload existentes.

Para ter mais informações, consulte [Modo de acesso híbrido](#).

## Compartilhamento direto entre contas

As entidades principais autorizadas podem compartilhar recursos explicitamente com uma entidade principal do IAM em uma conta externa. Esse atributo é útil quando o proprietário da conta deseja ter controle sobre quem na conta externa pode acessar os recursos. As permissões que a entidade principal do IAM receberá serão uma união de concessões diretas e concessões em nível de conta que serão transferidas em cascata para as entidades principais. O administrador do data lake da conta do destinatário pode ver as concessões diretas entre contas, mas não pode revogar as permissões. A entidade principal que recebe o compartilhamento de recursos não pode compartilhar o recurso com outras entidades principais.

## Métodos para compartilhar recursos do Catálogo de Dados

Com uma única operação de concessão do Lake Formation, você pode conceder permissões entre contas nos seguintes recursos do Catálogo de Dados:

- Um banco de dados
- Uma tabela individual (com filtragem de coluna opcional)
- Algumas tabelas selecionadas
- Todas as tabelas em um banco de dados (usando o curinga Todas as Tabelas)

Há duas opções para compartilhar seus bancos de dados e tabelas com outra conta Conta da AWS ou com diretores do IAM em outra conta.

- Controle de acesso baseado em tags do Lake Formation (LF-TBAC) (recomendado)

O controle de acesso baseado em tags do Lake Formation é uma estratégia de autorização que define permissões com base em atributos. Você pode usar o controle de acesso baseado em tags para compartilhar recursos do Catálogo de Dados (bancos de dados, tabelas e colunas) com diretores externos do IAM, Contas da AWS Organizations and Organizational Units (OUs). No Lake Formation, esses atributos são chamados de LF-tags. Para obter mais informações, consulte [Gerenciamento de um data lake usando o controle de acesso baseado em tags do Lake Formation](#).

**Note**

O método LF-TBAC de conceder permissões de uso do Catálogo de Dados para concessões entre contas. AWS Resource Access Manager

O Lake Formation agora oferece suporte à concessão de permissões entre contas para organizações e unidades organizacionais usando o método LF-TBAC.

Para ativar esse recurso, você precisa atualizar as Configurações de versão entre contas para a Versão 3.

Para ter mais informações, consulte [Como atualizar as configurações da versão de compartilhamento de dados entre contas](#).

- Recursos nomeados do Lake Formation

O compartilhamento de dados entre contas do Lake Formation usando o método de recurso nomeado permite que você conceda permissões do Lake Formation com uma opção de concessão em tabelas e bancos de dados do Catálogo de Dados para entidades externas Contas da AWS, diretores do IAM, organizações ou unidades organizacionais. A operação de concessão compartilha automaticamente esses recursos.

**Note**

Você também pode permitir que o AWS Glue rastreador acesse um armazenamento de dados em uma conta diferente usando as credenciais do Lake Formation. Para obter mais informações, consulte [Rastreamento entre contas no AWS Glue Guia do desenvolvedor](#).

Serviços integrados, como o Athena e o Amazon Redshift Spectrum, exigem links de recursos para poder incluir recursos compartilhados nas consultas. Para obter mais informações sobre os links de recursos, consulte [Como os links de recursos funcionam no Lake Formation](#).

Para conhecer as limitações e as considerações, consulte [Práticas recomendadas e considerações sobre compartilhamento de dados entre contas](#).

## Tópicos

- [Pré-requisitos](#)
- [Como atualizar as configurações da versão de compartilhamento de dados entre contas](#)

- [Compartilhamento de tabelas e bancos de dados do Catálogo de Dados entre Contas da AWS e entidades principais do IAM a partir de contas externas](#)
- [Conceder permissões em um banco de dados ou tabela compartilhada com sua conta](#)
- [Como conceder permissões de links de recursos](#)
- [Como acessar os dados subjacentes de uma tabela compartilhada](#)
- [Registro em várias contas CloudTrail](#)
- [Gerenciamento de permissões entre contas usando A e Lake Formation](#)
- [Visualizando todas as concessões entre contas usando a operação de GetResourceShares API](#)

#### Tópicos relacionados da

- [Visão geral das permissões do Lake Formation](#)
- [Acessar e visualizar tabelas e bancos de dados compartilhados do Catálogo de Dados](#)
- [Criação de links de recursos](#)
- [Resolução de problemas de acesso entre contas](#)

## Pré-requisitos

Antes que sua AWS conta possa compartilhar recursos do Catálogo de Dados (bancos de dados e tabelas) com outra conta ou diretores em outra conta e antes que você possa acessar os recursos compartilhados com sua conta, os seguintes pré-requisitos devem ser atendidos.

### Requisitos gerais de compartilhamento de dados entre contas

- Para compartilhar bancos de dados e tabelas do Catálogo de Dados no modo de acesso híbrido, você precisa atualizar as Configurações de versão entre contas para a Versão 4.
- Antes de conceder permissões entre contas em um recurso do Catálogo de Dados, você deve revogar todas as permissões do Lake Formation do grupo `IAMAllowedPrincipals` para o recurso. Se a entidade principal solicitante tiver permissões entre contas para acessar um recurso, e a permissão `IAMAllowedPrincipals` existir no recurso, Lake Formation exibirá `AccessDeniedException`.

Esse requisito é válido somente quando você registra a localização dos dados subjacentes no modo Lake Formation. Se você registrar a localização dos dados no modo híbrido, as

permissões do grupo `IAMAllowedPrincipals` poderão existir no banco de dados ou na tabela compartilhada.

- Nos bancos de dados que contêm tabelas que você cogita compartilhar, é necessário evitar que novas tabelas tenham uma concessão padrão de `Super` para `IAMAllowedPrincipals`. No console do Lake Formation, edite o banco de dados e desative Usar somente o controle de acesso do IAM para novas tabelas nesse banco de dados ou insira o AWS CLI comando a seguir, `database` substituindo-o pelo nome do banco de dados. Se a localização dos dados subjacentes estiver registrada no modo de acesso híbrido, você não precisará alterar essa configuração padrão. No modo de acesso híbrido, o Lake Formation permite que você aplique seletivamente as permissões do Lake Formation e as políticas de permissões do IAM para o Amazon S3 AWS Glue e no mesmo recurso.

```
aws glue update-database --name database --database-input  
'{"Name": "database", "CreateTableDefaultPermissions": []}'
```

- Para conceder permissões entre contas, o concedente deve ter as permissões necessárias AWS Identity and Access Management (IAM) sobre AWS Glue o serviço. AWS RAM A política AWS gerenciada `AWSLakeFormationCrossAccountManager` concede as permissões necessárias.

Os administradores de data lake em contas que recebem compartilhamentos de recursos usando AWS RAM devem ter a seguinte política adicional. Ele permite que o administrador aceite convites AWS RAM de compartilhamento de recursos. Permitir também que o administrador habilite o compartilhamento de recursos com organizações.

```
{  
  "Version": "2012-10-17",  
  "Statement": [  
    {  
      "Effect": "Allow",  
      "Action": [  
        "ram:AcceptResourceShareInvitation",  
        "ram:RejectResourceShareInvitation",  
        "ec2:DescribeAvailabilityZones",  
        "ram:EnableSharingWithAwsOrganization"  
      ],  
      "Resource": "*"   
    }  
  ]  
}
```



```
}
```

- Se você quiser compartilhar recursos do Catálogo de Dados com AWS Organizations ou com unidades organizacionais, o compartilhamento com organizações deve estar ativado em AWS RAM.

Para obter informações sobre como habilitar o compartilhamento com organizações, consulte [Habilitar o compartilhamento com AWS organizações](#) no Guia AWS RAM do usuário.

Você deve ter a permissão `ram:EnableSharingWithAwsOrganization` para habilitar o compartilhamento com organizações.

- Para compartilhar recursos diretamente com uma entidade principal do IAM em outra conta, você precisa atualizar as Configurações Versão Entre Contas para a Versão 3. Essa configuração está disponível na página de Configurações do Catálogo de Dados. Se você estiver usando a Versão 1, consulte as instruções para atualizar a configuração [Como atualizar as configurações da versão de compartilhamento de dados entre contas](#).
- Você não pode compartilhar recursos do Catálogo de Dados criptografados com a chave gerenciada do AWS Glue serviço com outra conta. Você pode compartilhar somente recursos do Catálogo de Dados criptografados com a chave de criptografia do cliente, e a conta que recebe o compartilhamento de recursos deve ter permissões na chave de criptografia do Catálogo de Dados para descriptografar os objetos.

### Compartilhamento de dados entre contas usando os requisitos do LF-TBAC

- Para compartilhar recursos do Catálogo de Dados com AWS Organizations unidades organizacionais (OUs), você precisa atualizar as configurações da versão da conta cruzada para a versão 3.
- Para compartilhar os recursos do Catálogo de Dados com a versão 3 das Configurações de versão entre contas, o concedente precisa ter as permissões de IAM definidas na política `AWSLakeFormationCrossAccountManager` gerenciada por AWS em sua conta.
- Se você estiver usando a versão 1 ou a versão 2 das Configurações de versão entre contas, deverá ter uma política de recursos do Catálogo de Dados (`glue:PutResourcePolicy`) que habilite o LF-TBAC. Para ter mais informações, consulte [Gerenciamento de permissões entre contas usando A e Lake Formation](#).
- Se você estiver usando atualmente uma política de recursos do Catálogo de Dados AWS Glue para compartilhar recursos e quiser conceder permissões entre contas usando a versão 3 das Configurações de versão entre contas, deverá adicionar a permissão `glue:ShareResource` nas

Configurações do Catálogo de Dados usando a operação de API `glue:PutResourcePolicy`, conforme mostrado na seção [Gerenciamento de permissões entre contas usando A e Lake Formation](#). Essa política não é necessária se sua conta não fez concessões entre contas usando a política de recursos do Catálogo de Dados AWS Glue (permissão de uso `glue:PutResourcePolicy` da versão 1 e versão 2) para conceder acesso entre contas.

```
{
  "Effect": "Allow",
  "Action": [
    "glue:ShareResource"
  ],
  "Principal": {"Service": [
    "ram.amazonaws.com"
  ]},
  "Resource": [
    "arn:aws:glue:<region>:<account-id>:table/*/*",
    "arn:aws:glue:<region>:<account-id>:database/*",
    "arn:aws:glue:<region>:<account-id>:catalog"
  ]
}
```

- Se sua conta fez compartilhamentos entre contas usando a política de recursos do Catálogo de Dados AWS Glue e você está usando o método de recurso nomeado ou LF-TBAC com a versão 3 das Configurações de versão entre contas para compartilhar recursos, que usam AWS RAM para compartilhar recursos, você deve definir o argumento `EnableHybrid` para `'true'` quando invocar a operação da API `glue:PutResourcePolicy`. Para ter mais informações, consulte [Gerenciamento de permissões entre contas usando A e Lake Formation](#).

Configuração necessária em cada conta que acessa o recurso compartilhado

- Se você estiver compartilhando recursos com Contas da AWS, pelo menos um usuário na conta do consumidor deve ser administrador do data lake para visualizar os recursos compartilhados. Para obter informações sobre como criar um administrador de data lake, consulte [Crie um administrador de data lake](#).

O administrador do data lake pode conceder permissões do Lake Formation sobre os recursos compartilhados com outras entidades principais da conta. Outras entidades principais não podem acessar recursos compartilhados até que o administrador do data lake conceda permissões sobre os recursos.

- Serviços integrados, como o Athena e o Redshift Spectrum, exigem links de recursos para poder incluir recursos compartilhados nas consultas. Entidades principais precisam criar um link de recurso em seu Catálogo de Dados para um recurso compartilhado de outra Conta da AWS. Para obter mais informações sobre os links de recursos, consulte [Como os links de recursos funcionam no Lake Formation](#).
- Quando um recurso é compartilhado diretamente com uma entidade principal do IAM, para consultar a tabela usando o Athena, a entidade principal precisa criar um link de recurso. Para criar um link de recurso, a entidade principal precisa da permissão `CREATE_TABLE` ou `CREATE_DATABASE` do Lake Formation, e da permissão `glue:CreateTable` ou `glue:CreateDatabase` do IAM.

Se a conta do produtor compartilhar uma tabela diferente no mesmo banco de dados com o mesmo ou outra entidade principal, esse principal poderá consultar imediatamente a tabela.

#### Note

Para o administrador do data lake e para as entidades principais às quais o administrador do data lake concedeu permissões, os recursos compartilhados aparecem no Catálogo de Dados como se fossem recursos locais (de propriedade). Tarefas de extração, transformação e carregamento (ETL) podem acessar os dados subjacentes dos recursos compartilhados. Para recursos compartilhados, as páginas Tabelas e bancos de dados no console do Lake Formation exibem o ID da conta do proprietário.

Quando os dados subjacentes de um recurso compartilhado são acessados, os eventos de CloudTrail registro são gerados na conta do destinatário do recurso compartilhado e na conta do proprietário do recurso. Os CloudTrail eventos podem conter o ARN do principal que acessou os dados, mas somente se a conta do destinatário optar por incluir o ARN principal nos registros. Para ter mais informações, consulte [Registro em várias contas CloudTrail](#).

## Como atualizar as configurações da versão de compartilhamento de dados entre contas

De tempos em tempos, AWS Lake Formation atualiza as configurações de compartilhamento de dados entre contas para distinguir as alterações feitas no AWS RAM uso e para oferecer suporte às atualizações feitas no recurso de compartilhamento de dados entre contas. Quando o Lake Formation faz isso, ele cria uma nova versão das Configurações de versão entre contas.

## Principais diferenças entre as versões de Configurações Entre Contas

Para obter mais informações sobre como o compartilhamento de dados entre contas funciona em diferentes versões de Configurações de versão entre contas, consulte as seções a seguir.

### Note

Para compartilhar dados com outra conta, o concedente deve ter permissões de política do IAM gerenciadas por `AWSLakeFormationCrossAccountManager`. Isso é um pré-requisito para todas as versões.

A atualização da versão das Configurações de versão entre contas não afeta as permissões que o destinatário tem nos recursos compartilhados. Isso é aplicável ao atualizar da versão 1 para a versão 2, da versão 2 para a versão 3 e da versão 1 para a versão 3. Veja as considerações listadas abaixo ao atualizar as versões.

### Versão 1

Método de recurso nomeado: mapeia cada concessão de permissão entre contas do Lake Formation para um compartilhamento AWS RAM de recursos. O usuário (no perfil de concedente ou entidade principal) não precisa de permissões adicionais.

Método LF-TBAC: as permissões entre contas do Lake Formation não AWS RAM são usadas para compartilhar dados. O usuário deve ter a permissão `glue:PutResourcePolicy`.

Benefícios da atualização de versões: Versão inicial — não aplicável.

Considerações ao atualizar versões: Versão inicial — não aplicável.

### Versão 2

Método de recurso nomeado: otimiza o número de compartilhamentos de AWS RAM recursos mapeando várias concessões de permissão entre contas com um compartilhamento de AWS RAM recursos. Usuários não precisam de permissões adicionais.

Método LF-TBAC: as permissões entre contas do Lake Formation não AWS RAM são usadas para compartilhar dados. O usuário deve ter a permissão `glue:PutResourcePolicy`.

Benefícios da atualização de versões: configuração escalável entre contas por meio da utilização ideal da capacidade. AWS RAM

Considerações ao atualizar versões: os usuários que desejam conceder permissões entre contas do Lake Formation devem ter as permissões na política `AWSLakeFormationCrossAccountManager` AWS gerenciada. Caso contrário, você precisará ter permissões `ram:AssociateResourceShare` e `ram:DisassociateResourceShare` para compartilhar recursos com sucesso com outra conta.

### Versão 3

Método de recurso nomeado: otimiza o número de compartilhamentos de AWS RAM recursos mapeando várias concessões de permissão entre contas com um compartilhamento de AWS RAM recursos. Usuários não precisam de permissões adicionais.

Método LF-TBAC: o Lake Formation usa AWS RAM para doações entre contas. O usuário deve adicionar `ShareResource` declaração à `glue:PutResourcePolicy` permissão. O destinatário deve aceitar convites de compartilhamento de recursos de AWS RAM

Benefícios da atualização de versões: Suporta os seguintes recursos:

- Permite compartilhar recursos explicitamente com uma entidade principal do IAM em uma conta externa.

Para ter mais informações, consulte [Concedendo e revogando permissões nos recursos do Catálogo de dados](#).

- Permite o compartilhamento entre contas usando o método LF-TBAC para organizações ou unidades organizacionais (UOs).
- Elimina a sobrecarga de manter AWS Glue políticas adicionais para subsídios entre contas.

Considerações ao atualizar as versões: Se o concedente usar uma versão inferior à versão 3 e o destinatário estiver usando a versão 3 ou superior, o concedente receberá a seguinte mensagem de erro: "Solicitação de concessão entre contas inválida. A conta do consumidor optou pela versão entre contas: v3. `CrossAccountVersionAtualize DataLakeSetting` para a versão mínima v3 (Serviço: `AmazonDataCatalog`; Código de status: 400; Código de erro: `InvalidInputException`)". No entanto, se o concedente usar a versão 3 e o destinatário estiver usando a versão 1 ou a versão 2, as concessões entre contas serão aprovadas com êxito.

Para compartilhar recursos diretamente com as entidades principais do IAM em outra conta, somente o concedente precisa usar a versão 3.

As concessões entre contas feitas usando o método LF-TBAC exigem que os usuários tenham uma política de recursos AWS Glue Data Catalog na conta. Quando você atualiza

para a versão 3, o LF-TBAC concede usos de AWS RAM. Para permitir que as concessões AWS RAM baseadas em várias contas sejam bem-sucedidas, você deve adicionar a `glue:ShareResource` declaração às suas políticas de recursos existentes do Catálogo de Dados, conforme mostrado na [Gerenciamento de permissões entre contas usando A e Lake Formation](#) seção.

#### Versão 4

O concedente precisa da versão 4 ou superior para compartilhar os recursos do Catálogo de Dados no modo de acesso híbrido.

## Otimize os compartilhamentos AWS RAM de recursos

As novas versões (versão 2 e superior) de concessões entre contas utilizam de forma otimizada a AWS RAM capacidade para maximizar o uso de várias contas. Quando você compartilha um recurso com um diretor externo Conta da AWS ou do IAM, o Lake Formation pode criar um novo compartilhamento de recursos ou associar o recurso a um compartilhamento existente. Ao se associar aos compartilhamentos existentes, o Lake Formation reduz o número de convites de compartilhamento de recursos que um consumidor precisa aceitar.

## Habilite AWS RAM compartilhamentos via TBAC ou compartilhe recursos diretamente com os diretores

Para compartilhar recursos diretamente com entidades principais do IAM em outra conta, ou para habilitar compartilhamentos entre contas TBAC para organizações ou unidades organizacionais, é necessário atualizar as Configurações de versão entre contas para a versão 3. Para obter mais informações sobre limites AWS RAM de recursos, consulte [Práticas recomendadas e considerações sobre compartilhamento de dados entre contas](#).

### Permissões necessárias para atualizar a versão de Configurações Entre Contas

Se um concedente de permissão entre contas tiver permissões de política do IAM gerenciadas por `AWSLakeFormationCrossAccountManager`, não será necessária nenhuma configuração de permissão extra para o perfil de concedente ou entidade principal para permissão entre contas. No entanto, se o concedente entre contas não estiver usando a política gerenciada, o perfil do concedente ou entidade principal deverá ter as seguintes permissões do IAM concedidas para que a nova versão da concessão entre contas seja bem-sucedida.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "VisualEditor1",
      "Effect": "Allow",
      "Action": [
        "ram:AssociateResourceShare",
        "ram:DisassociateResourceShare",
        "ram:GetResourceShares"
      ],
      "Resource": "*",
      "Condition": {
        "StringLike": {
          "ram:ResourceShareName": "LakeFormation*"
        }
      }
    }
  ]
}
```

## Como habilitar a nova versão

Siga estas etapas para atualizar as configurações da versão da conta cruzada por meio do AWS Lake Formation console ou do AWS CLI.

### Console

1. Selecione **Versão 2**, **Versão 3** ou **Versão 4** em **Configurações de versão entre contas** na página de **Configurações do Catálogo de Dados**. Se você selecionar a **Versão 1**, o Lake Formation usará o modo padrão de compartilhamento de recursos.

AWS Lake Formation &gt; Data catalog settings

## Data catalog settings

### Default permissions for newly created databases and tables

These settings maintain existing AWS Glue Data Catalog behavior. You can still set individual permissions on databases and tables, which will take effect when you revoke the Super permission from IAMAllowedPrincipals. See [Changing Default Settings for Your Data Lake](#).

- Use only IAM access control for new databases
- Use only IAM access control for new tables in new databases

### Default permissions for AWS CloudTrail

These settings specify the information being shown in AWS CloudTrail.

#### Resource owners

Enter resource owners you wish to share your CloudTrail access details with.

Enter one or more AWS account IDs. Press Enter after each ID.

### Cross account version settings

Version 1

Version 2

Version 3

Version 3 ▲

Cross account permissions. See

Cancel

Save

## 2. Escolha Salvar

### AWS Command Line Interface (AWS CLI)

Use o `put-data-lake-settings` AWS CLI comando para definir o `CROSS_ACCOUNT_VERSION` parâmetro. Os valores aceitos são 1, 2, 3 e 4.

```
aws lakeformation put-data-lake-settings --region us-east-1 --data-lake-settings
file://settings
{
```



```
"DataLakeAdmins": [  
  {  
    "DataLakePrincipalIdentifier": "arn:aws:iam::111122223333:user/test"  
  }  
],  
"CreateDatabaseDefaultPermissions": [],  
"CreateTableDefaultPermissions": [],  
"Parameters": {  
  "CROSS_ACCOUNT_VERSION": "3"  
}  
}
```

### Important

Após escolher a Versão 2 ou a Versão 3, todas as novas concessões de recursos nomeados passarão pelo novo modo de concessão entre contas. Para otimizar o uso AWS RAM da capacidade de seus compartilhamentos entre contas existentes, recomendamos que você revogue as concessões que foram feitas com a versão mais antiga e conceda novamente no novo modo.

## Compartilhamento de tabelas e bancos de dados do Catálogo de Dados entre Contas da AWS e entidades principais do IAM a partir de contas externas

Esta seção inclui instruções sobre como habilitar permissões entre contas em tabelas e bancos de dados do Catálogo de Dados para uma AWS conta externa, diretor do IAM, organização ou unidade organizacional. A operação de concessão compartilha automaticamente esses recursos.

### Tópicos

- [Compartilhamento de dados usando controle de acesso baseado em tags](#)
- [Compartilhamento de dados entre contas usando o método de recurso nomeado](#)

## Compartilhamento de dados usando controle de acesso baseado em tags

### Configuração obrigatória na conta do produtor/concedente

1. Defina uma LF-tag. Para obter instruções sobre como criar uma LF-tag, consulte [Criação de LF-Tags](#).
2. Atribua a LF-tag ao recurso de destino. Para ter mais informações, consulte [Atribuição de LF-Tags aos recursos do catálogo de dados](#).
3. Conceda permissão da LF-tag à conta externa. Para ter mais informações, consulte [Concedendo permissões de LF-Tag usando o console](#).

Nesse ponto, o administrador do data lake do consumidor deve ser capaz de encontrar a tag de política que está sendo compartilhada por meio do console Lake Formation da conta do concedido, em Permissões, Perfis e tarefas administrativas, LF-tags.

4. Conceda permissão de dados para a conta externa ou do concedido.
  - a. No painel de navegação, em Permissões, Data Lake, escolha Conceder.
  - b. Para Diretores, escolha Contas externas e insira o Conta da AWS ID de destino ou a função IAM do principal ou o Amazon Resource Name (ARN) do principal (ARN principal).
  - c. Para LF-tags ou recursos de catálogo, escolha a chave e os valores da LF-tag que está sendo compartilhada com a conta do consumidor (chave Confidentiality e valor public).
  - d. Para Permissões, em Recursos combinados com LF-tags (recomendado), escolha Adicionar LF-tag.
  - e. Selecione a chave e o valor da tag que está sendo compartilhada com a conta do concedido (chave Confidentiality e valor public).
  - f. Para permissões de banco de dados, selecione Descrever em Permissões de banco de dados para conceder permissões de acesso no nível do banco de dados.
  - g. O administrador do Data Lake do consumidor deve ser capaz de encontrar a tag de política que está sendo compartilhada por meio da conta do consumidor no console do Lake Formation em <https://console.aws.amazon.com/lakeformation/>, em Permissões, Perfis e tarefas administrativas, LF-tags.
  - h. Selecione Descrever em Permissões concedíveis para que a conta do consumidor possa conceder permissões em nível de banco de dados a seus usuários.

Como o administrador do data lake deve conceder permissões em recursos compartilhados para as entidades principais na conta do concedido, as permissões entre contas devem sempre ser concedidas com a opção de concessão.

**Note**

Entidades principais que receberem concessões diretas entre contas não terão a opção de Permissões concedíveis.

- i. Para Permissões de tabela e coluna, selecione Selecionar e descrever em Permissões de tabela.
- j. Selecione Selecionar e Descrever em Permissões concedíveis.
- k. Escolha Conceder

### Configuração necessária na conta de recebedor/concedido

1. Quando você compartilha um recurso com outra conta, o recurso ainda pertence à conta do produtor e não será visível no console do Athena. Para tornar o recurso visível no console do Athena, você precisa criar um link de recurso direcionando para o recurso compartilhado. Para obter instruções sobre como criar um link de recurso, consulte [Como criar um link de recurso para uma tabela compartilhada do Catálogo de Dados](#) e [Como criar um link de recurso para um banco de dados compartilhado do Catálogo de Dados](#)
2. Você precisa criar um conjunto separado de LF-tags na conta do consumidor para usar o controle de acesso baseado em LF-tags ao compartilhar os links de recursos. Crie e atribua as LF-tags necessárias ao banco de dados/tabelas compartilhados e aos links de recursos.
3. Conceda permissões sobre essas LF-tags às entidades principais do IAM na conta do concedido.

### Compartilhamento de dados entre contas usando o método de recurso nomeado

Você pode conceder permissões diretamente aos diretores em outra AWS conta, ou para uma conta externa Contas da AWS ou AWS Organizations. Conceder permissões do Lake Formation a organizações ou unidades organizacionais é equivalente a conceder a permissão a todos Conta da AWS nessa organização ou unidade organizacional.

Ao conceder permissões a contas ou organizações externas, você deve incluir a opção Permissões concedíveis. Somente o administrador do data lake na conta externa pode acessar os recursos compartilhados até que o administrador conceda permissões sobre os recursos compartilhados a outras entidades principais na conta externa.

**Note**

A opção de Permissões concedíveis não é suportada ao conceder permissões diretamente às entidades principais do IAM a partir de contas externas.

Siga as instruções em [Conceder permissões de banco de dados usando o método de recurso nomeado](#) para conceder permissões entre contas usando o método de recurso nomeado.

## Conceder permissões em um banco de dados ou tabela compartilhada com sua conta

Depois que um recurso do Catálogo de Dados pertencente a outra AWS conta for compartilhado com sua AWS conta, como administrador do data lake, você poderá conceder permissões sobre o recurso compartilhado a outros diretores da sua conta. No entanto, você não pode conceder permissões sobre o recurso a outras contas AWS ou organizações.

Você pode usar o AWS Lake Formation console, a API ou o AWS Command Line Interface (AWS CLI) para conceder as permissões.

Como conceder permissões em um banco de dados compartilhado (método de recurso nomeado, console)

- Siga as instruções em [Conceder permissões de banco de dados usando o método de recurso nomeado](#). Na lista Banco de dados, em LF-tags ou recursos do catálogo, certifique-se de selecionar o banco de dados na conta externa, não um link de recurso para o banco de dados.

Se você não encontrar o banco de dados na lista de bancos de dados, certifique-se de ter aceitado o convite de compartilhamento de recursos AWS Resource Access Manager (AWS RAM) para o banco de dados. Para ter mais informações, consulte [Aceitar um convite de compartilhamento de recursos do AWS RAM](#).

Além disso, para obter as permissões CREATE\_TABLE e ALTER, siga as instruções em [Concessão de permissões de localização de dados \(mesma conta\)](#) e não se esqueça de inserir o ID da conta proprietária no campo Localização da conta registrada.

Como conceder permissões em uma tabela compartilhada (método de recurso nomeado, console)

- Siga as instruções em [Conceder permissões de tabela usando o método de recurso nomeado](#). Na lista Banco de dados, em LF-tags ou recursos do catálogo, certifique-se de selecionar o banco de dados na conta externa, não um link de recurso para o banco de dados.

Se você não encontrar tabela na lista de tabelas, certifique-se de ter aceitado o convite de compartilhamento de recursos AWS RAM para a tabela. Para ter mais informações, consulte [Aceitar um convite de compartilhamento de recursos do AWS RAM](#).

Além disso, para obter as permissões ALTER, siga as instruções em [Concessão de permissões de localização de dados \(mesma conta\)](#) e não se esqueça de inserir o ID da conta proprietária no campo Localização da conta registrada.

Como conceder permissões em recursos compartilhados (método LF-TBAC, console)

- Siga as instruções em [Conceder permissões do catálogo de dados](#). Na seção LF-tags ou recursos do catálogo, conceda a expressão exata da LF-tag que a conta externa concedeu à sua conta ou um subconjunto dessa expressão.

Por exemplo, se uma conta externa concedeu a expressão da LF-tag `module=customers AND environment=production` à sua conta com a opção de concessão, como administrador do data lake, você pode conceder essa mesma expressão, `module=customers` ou `environment=production` a uma entidade principal em sua conta. Você pode conceder somente as mesmas permissões ou um subconjunto das permissões do Lake Formation (por exemplo: SELECT, ALTER e assim por diante) que foram concedidas aos recursos por meio da expressão da LF-tag.

Para conceder permissões em uma tabela compartilhada (chamado método de recurso AWS CLI)

- Digite um comando semelhante ao seguinte: Neste exemplo:
  - O ID AWS da sua conta é 1111-2222-3333.
  - A conta que possui a tabela e que a concedeu à sua conta é 1234-5678-9012.
  - A permissão SELECT está sendo concedida na tabela compartilhada `pageviews` ao usuário `dataLake_user1`. Esse usuário é a entidade principal da sua conta.

- A tabela `pageviews` está no banco de dados `analytics`, que pertence à conta `1234-5678-9012`.

```
aws lakeformation grant-permissions --principal
DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1
--permissions "SELECT" --resource '{ "Table": {"CatalogId":"123456789012",
"DatabaseName":"analytics", "Name":"pageviews"} }'
```

Observe que a conta proprietária deve ser especificada na propriedade `CatalogId`, no argumento `resource`.

## Como conceder permissões de links de recursos

Siga estas etapas para conceder AWS Lake Formation permissões em um ou mais links de recursos a um diretor em sua AWS conta.

Após criar um link de recurso, somente você poderá visualizá-lo e acessá-lo. (Isso pressupõe que Usar somente o controle de acesso do IAM para novas tabelas nesse banco de dados não esteja habilitado para o banco de dados.) Para permitir que outras entidades principais da sua conta acessem o link do recurso, conceda pelo menos a permissão `DESCRIBE`.

### Important

Conceder permissões em um link de recurso não concede permissões no banco de dados ou tabela de destino (vinculado). Você deve conceder permissões no destino separadamente.

Você pode conceder permissões usando o console do Lake Formation, a API ou o AWS Command Line Interface (AWS CLI).

console

Como conceder permissões de links de recursos usando o console do Lake Formation

1. Execute um destes procedimentos:

- Para obter links de recursos de banco de dados, siga as etapas em [Conceder permissões de banco de dados usando o método de recurso nomeado](#) para fazer o seguinte:

1. Abra a página Conceder permissões de data lake.
  2. Especifique os bancos de dados. Especifique um ou mais links de recursos do banco de dados.
  3. Especifique as entidades principais.
- Para obter links de recursos de tabela, siga as etapas em [Conceder permissões de tabela usando o método de recurso nomeado](#) para fazer o seguinte:
    1. Abra a página Conceder permissões de data lake.
    2. Especifique as tabelas. Especifique um ou mais links de recursos de tabela.
    3. Especifique as entidades principais.
2. Em Permissões, selecione as permissões a serem concedidas. Como opção, selecione permissões concedíveis.

### Permissions

Select the permissions to grant.

**Resource link permissions**  
Grant resource-wide permissions.

**Column-based permissions**  
Grant data access to specific columns.

---

**Resource link permissions**  
Choose specific access permissions to grant.

Drop  Describe

---

Super  
This permission is the union of the individual permissions above and supercedes them. [Learn More](#)

**Grantable permissions**  
Choose the permission that may be granted to others.

Drop  Describe

---

Super  
This permission is the union of the individual permissions above and supercedes them. [Learn More](#)

3. Selecione Conceder.

## AWS CLI

Para conceder permissões de links de recursos usando AWS CLI

- Execute o comando `grant-permissions`, especificando um link de recurso como recurso.

### Example

Este exemplo concede `DESCRIBE` ao usuário `datalake_user1` na tabela o link do recurso no banco de dados `incidents-link issues` na AWS conta `1111-2222-3333`.

```
aws lakeformation grant-permissions --principal
DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1
--permissions "DESCRIBE" --resource '{ "Table": {"DatabaseName":"issues",
"Name":"incidents-link"}}'
```

### Consulte também:

- [Criação de links de recursos](#)
- [Referência de permissões do Lake Formation](#)

## Como acessar os dados subjacentes de uma tabela compartilhada

Suponha que a AWS conta A compartilhe uma tabela do Catálogo de Dados com a conta B — por exemplo, concedendo a opção de concessão `SELECT` na tabela à conta B. Para que um principal na conta B possa ler os dados subjacentes da tabela compartilhada, as seguintes condições devem ser atendidas:

- O administrador do data lake na conta B deve aceitar o compartilhamento. (Isso não é necessário se as contas A e B estiverem na mesma organização ou se a concessão tiver sido feita com o método de controle de acesso baseado em tags do Lake Formation.)
- O administrador do data lake deve conceder novamente à entidade principal a permissão `SELECT` do Lake Formation que a conta A concedeu na tabela compartilhada.
- A entidade principal deve ter as seguintes permissões de IAM na tabela, no banco de dados que a contém e no Catálogo de Dados da conta A.



**Note**

Na seguinte política do IAM:

- <account-id-A>Substitua pelo AWS ID da conta A.
- Substitua <region> por uma região válida.
- Substitua <database> pelo nome do banco de dados na conta A que contém a tabela compartilhada.
- Substitua <table> pelo nome da tabela compartilhada.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "glue:GetTable",
        "glue:GetTables",
        "glue:GetPartition",
        "glue:GetPartitions",
        "glue:BatchGetPartition",
        "glue:GetDatabase",
        "glue:GetDatabases"
      ],
      "Resource": [
        "arn:aws:glue:<region>:<account-id-A>:table/<database>/<table>",
        "arn:aws:glue:<region>:<account-id-A>:database/<database>",
        "arn:aws:glue:<region>:<account-id-A>:catalog"
      ]
    },
    {
      "Effect": "Allow",
      "Action": [
        "lakeformation:GetDataAccess"
      ],
      "Resource": [
        "*"
      ],
      "Condition": {
        "StringEquals": {
```

```
        "lakeformation:GlueARN": "arn:aws:glue:<region>:<account-id-  
A>:table/<database>/<table>"  
    }  
} ]  
}
```

### Consulte também:

- [Aceitar um convite de compartilhamento de recursos do AWS RAM](#)

## Registro em várias contas CloudTrail

O Lake Formation fornece uma trilha de auditoria centralizada de todo o acesso entre contas aos dados em seu data lake. Quando uma AWS conta de destinatário acessa dados em uma tabela compartilhada, o Lake Formation copia o CloudTrail evento para os registros da CloudTrail conta proprietária. Os eventos copiados incluem consultas aos dados por serviços integrados, como o Amazon Redshift Spectrum, Amazon Athena e acessos a dados por trabalhos. AWS Glue

CloudTrail os eventos para operações entre contas nos recursos do Catálogo de Dados são copiados de forma semelhante.

Como proprietário do recurso, se você habilitar o registro em nível de objeto no Amazon S3, poderá executar consultas que unem eventos do S3 aos CloudTrail eventos do CloudTrail Lake Formation para determinar as contas que acessaram seus buckets do S3.

### Tópicos

- [Incluindo identidades principais em registros de várias contas CloudTrail](#)
- [Consultando CloudTrail registros para acesso entre contas do Amazon S3](#)

## Incluindo identidades principais em registros de várias contas CloudTrail

Por padrão, CloudTrail os eventos entre contas adicionados aos registros do destinatário do recurso compartilhado e copiados para os registros do proprietário do recurso contêm somente o ID AWS principal do responsável externo da conta, não o nome de recurso da Amazon (ARN) legível por

humanos do principal (ARN principal). Ao compartilhar recursos dentro de limites confiáveis, como dentro da mesma organização ou equipe, você pode optar por incluir o ARN principal nos CloudTrail eventos. As contas do proprietário do recurso podem então rastrear as entidades principais nas contas de destinatários que acessam seus recursos próprios.

### Important

Como destinatário do recurso compartilhado, para ver o ARN principal em eventos em seus próprios CloudTrail registros, você deve optar por compartilhar o ARN principal com a conta do proprietário.

Se o acesso aos dados ocorrer por meio de um link de recurso, dois eventos serão registrados na conta do destinatário do recurso compartilhado: um para o acesso ao link do recurso e outro para o acesso ao recurso de destino. O evento para o acesso ao link de recurso inclui o ARN da entidade principal. O evento para o acesso ao recurso de destino não inclui o ARN da entidade principal sem selecionar a opção de inclusão. O evento de acesso ao link do recurso não é copiado para a conta do proprietário.

A seguir está um trecho de um CloudTrail evento padrão entre contas (sem aceitação). A conta que executa o acesso aos dados é 1111-2222-3333. Este é o log mostrado na conta de origem e na conta do proprietário do recurso. Lake Formation preenche logs em ambas as contas no caso entre contas.

```
{
  "eventVersion": "1.05",
  "userIdentity": {
    "type": "AWSAccount",
    "principalId": "AROAQGFTBBBG0BWV2EMZA:GlueJobRunnerSession",
    "accountId": "111122223333"
  },
  "eventSource": "lakeformation.amazonaws.com",
  "eventName": "GetDataAccess",
  ...
  ...
  "additionalEventData": {
    "requesterService": "GLUE_JOB",
    "lakeFormationRoleSessionName": "AWSLF-00-GL-111122223333-G13T0Rmng2"
  },
  ...
}
```

Como consumidor de recursos compartilhados, quando você opta por incluir o ARN da entidade principal, o trecho se torna o seguinte. O campo `lakeFormationPrincipal` representa o perfil final ou o usuário que executa a consulta por meio do Amazon Athena, Amazon Redshift Spectrum ou tarefas AWS Glue.

```
{
  "eventVersion": "1.05",
  "userIdentity": {
    "type": "AWSAccount",
    "principalId": "AROAQGFTBBBG0BWV2EMZA:GlueJobRunnerSession",
    "accountId": "111122223333"
  },
  "eventSource": "lakeformation.amazonaws.com",
  "eventName": "GetDataAccess",
  ...
  ...
  "additionalEventData": {
    "requesterService": "GLUE_JOB",
    "lakeFormationPrincipal": "arn:aws:iam::111122223333:role/ETL-Glue-Role",
    "lakeFormationRoleSessionName": "AWSLF-00-GL-111122223333-G13T0Rmng2"
  },
  ...
}
```

Para optar por incluir ARNs principais nos registros de várias contas CloudTrail

1. Abra o console do Lake Formation em <https://console.aws.amazon.com/lakeformation/>.

Faça login como usuário Administrator ou como usuário com a política do IAM Administrator Access.

2. No painel de navegação, selecione Configurações.
3. Na página Configurações do catálogo de dados, na AWS CloudTrail seção Permissões padrão para, para proprietários de recursos, insira uma ou mais IDs de conta do proprietário do AWS recurso.

Pressione Enter após cada ID da conta.

4. Escolha Salvar.

Agora, CloudTrail os eventos entre contas armazenados nos registros do destinatário do recurso compartilhado e do proprietário do recurso contêm o ARN principal.

## Consultando CloudTrail registros para acesso entre contas do Amazon S3

Como proprietário de um recurso compartilhado, você pode consultar CloudTrail os registros do S3 para determinar as contas que acessaram seus buckets do Amazon S3 (desde que você tenha habilitado o registro em nível de objeto no Amazon S3). Isso se aplica somente aos locais do S3 que você registrou no Lake Formation. Se os consumidores de recursos compartilhados optarem por incluir os Rans principais nos CloudTrail registros do Lake Formation, você poderá determinar as funções ou os usuários que acessaram os buckets.

Ao executar consultas com Amazon Athena, você pode unir eventos do Lake Formation e CloudTrail eventos do S3 na CloudTrail propriedade do nome da sessão. As consultas também podem filtrar eventos do Lake Formation em `eventName="GetDataAccess"` e eventos do S3 em `eventName="Get Object"` ou `eventName="Put Object"`.

A seguir está um trecho de um CloudTrail evento entre contas do Lake Formation em que dados em um local registrado do S3 foram acessados.

```
{
  "eventSource": "lakeformation.amazonaws.com",
  "eventName": "GetDataAccess",
  .....
  .....
  "additionalEventData": {
    "requesterService": "GLUE_JOB",
    "lakeFormationPrincipal": "arn:aws:iam::111122223333:role/ETL-Glue-Role",
    "lakeFormationRoleSessionName": "AWSLF-00-GL-111122223333-B8JSAjo5QA"
  }
}
```

O valor da `lakeFormationRoleSessionName` chave `AWSLF-00-GL-111122223333-B8JSAjo5QA`, pode ser associado ao nome da sessão na `principalId` chave do CloudTrail evento do S3. A seguir está um trecho do evento CloudTrail S3. Veja a localização do nome da sessão.

```
{
  "eventSource": "s3.amazonaws.com",
  "eventName": "Get Object"
  .....
  .....
  "principalId": "AROAQSOX5XXUR7D6RMYLR:AWSLF-00-GL-111122223333-B8JSAjo5QA",
```

```

"arn": "arn:aws:sets::111122223333:assumed-role/Deformationally/AWSLF-00-
GL-111122223333-B8JSAjo5QA",
"session Context": {
  "session Issuer": {
    "type": "Role",
    "principalId": "AROAQS0X5XXUR7D6RMYLR",
    "arn": "arn:aws:iam::111122223333:role/aws-service-role/
lakeformation.amazonaws.com/Deformationally",
    "accountId": "111122223333",
    "user Name": "Deformationally"
  },
  .....
  .....
}

```

O nome da sessão é formatado da seguinte forma:

```
AWSLF-<version-number>-<query-engine-code>-<account-id>-<suffix>
```

### version-number

A versão desse formato, atualmente 00. Se o formato do nome da sessão mudar, a próxima versão será 01.

### query-engine-code

Indica a entidade que acessou os dados. Os valores atuais são:

GL	Tarefa de ETL AWS Glue
AT	Athena
RE	Amazon Redshift Spectrum

### account-id

O ID da AWS conta que solicitou as credenciais do Lake Formation.

### suffix

Uma string gerada aleatoriamente.

## Gerenciamento de permissões entre contas usando A e Lake Formation

É possível conceder acesso entre contas a recursos do Catálogo de Dados e dados subjacentes usando um AWS Glue ou AWS Lake Formation.

Em AWS Glue, você concede permissões entre contas criando ou atualizando uma política de recursos do Catálogo de Dados. No Lake Formation, você concede permissões entre contas usando o modelo de permissões GRANT/REVOKE do Lake Formation e a operação da API `Grant Permissions`.

### Tip

Recomendamos que você confie somente nas permissões do Lake Formation para proteger seu data lake.

Você pode ver os subsídios entre contas do Lake Formation usando o console do Lake Formation ou o console AWS Resource Access Manager (AWS RAM). No entanto, essas páginas do console não mostram permissões entre contas concedidas pela política de recursos do Catálogo de Dados AWS Glue. Da mesma forma, você pode visualizar as concessões entre contas na política de recursos do Catálogo de Dados usando a página Configurações do console AWS Glue, mas essa página não mostra as permissões entre contas concedidas usando o Lake Formation.

Para garantir que você não perca nenhuma concessão ao visualizar e gerenciar permissões entre contas, o Lake Formation e AWS Glue exigem que você execute as seguintes ações para indicar que está ciente e permite concessões entre contas pelo Lake Formation e AWS Glue.

Ao conceder permissões entre contas usando a política de recursos do Catálogo de Dados AWS Glue

Se sua conta (conta do concedente ou conta do produtor) não tiver concedido doações entre contas que sejam usadas AWS RAM para compartilhar os recursos, você poderá salvar uma política de recursos do Catálogo de Dados normalmente em AWS Glue. No entanto, se concessões que envolvem compartilhamentos de AWS RAM recursos já tiverem sido feitas, você deverá fazer o seguinte para garantir que o salvamento da política de recursos seja bem-sucedido:

- Quando você salva a política de recursos na página Configurações do console AWS Glue, o console emite um alerta informando que as permissões na política serão adicionais às permissões

concedidas usando o console do Lake Formation. Você deve escolher Continuar para salvar a política.

- Ao salvar a política de recursos usando a operação da API `glue:PutResourcePolicy`, você deve definir o campo `EnableHybrid` como `'TRUE'` (type = string). O exemplo de código a seguir mostra como fazer isso em Python.

```
import boto3
import json

REGION = 'us-east-2'
PRODUCER_ACCOUNT_ID = '123456789012'
CONSUMER_ACCOUNT_IDS = ['111122223333']

glue = glue_client = boto3.client('glue')

policy = {
    "Version": "2012-10-17",
    "Statement": [
        {
            "Sid": "Cataloguers",
            "Effect": "Allow",
            "Action": [
                "glue:*"
            ],
            "Principal": {
                "AWS": CONSUMER_ACCOUNT_IDS
            },
            "Resource": [
                f"arn:aws:glue:{REGION}:{PRODUCER_ACCOUNT_ID}:catalog",
                f"arn:aws:glue:{REGION}:{PRODUCER_ACCOUNT_ID}:database/*",
                f"arn:aws:glue:{REGION}:{PRODUCER_ACCOUNT_ID}:table/*/*"
            ]
        }
    ]
}

policy = json.dumps(policy)
glue.put_resource_policy(PolicyInJson=policy, EnableHybrid='TRUE')
```

Para obter mais informações, consulte [PutResourcePolicy Action \(Python: put\\_resource\\_policy\)](#) no Guia do desenvolvedor.AWS Glue



## Ao conceder permissões entre contas usando o método de recursos nomeados do Lake Formation

Se não houver uma política de recursos do Catálogo de Dados em sua conta, as concessões entre contas do Lake Formation que você fizer procederão normalmente. No entanto, se existir uma política de recursos do Catálogo de Dados, você deverá adicionar a seguinte declaração a ela para permitir que suas concessões entre contas sejam bem-sucedidas se forem feitas com o método de recurso nomeado. <region>Substitua por um nome de região válido e pelo <account-id>ID AWS da sua conta.

```
{
  "Effect": "Allow",
  "Action": [
    "glue:ShareResource"
  ],
  "Principal": {"Service": [
    "ram.amazonaws.com"
  ]},
  "Resource": [
    "arn:aws:glue:<region>:<account-id>:table/*/*",
    "arn:aws:glue:<region>:<account-id>:database/*",
    "arn:aws:glue:<region>:<account-id>:catalog"
  ]
}
```

Sem essa declaração adicional, o subsídio do Lake Formation é bem-sucedido, mas fica bloqueado AWS RAM e a conta do destinatário não pode acessar o recurso concedido.

### Important

Ao usar o método de controle de acesso baseado em tags do Lake Formation (LF-TBAC) para fazer concessões entre contas, você deve ter uma política de recursos do Catálogo de Dados com pelo menos as permissões especificadas em [Pré-requisitos](#).

### Consulte também:

- [Controle de acesso a metadados](#) (para uma análise do método de recurso nomeado em comparação com o método de controle de acesso baseado em tags do Lake Formation (LF-TBAC)).

- [Visualizando tabelas e bancos de dados compartilhados do Catálogo de Dados](#)
- [Como trabalhar com as configurações do Catálogo de Dados no console AWS Glue](#) no Guia do Desenvolvedor AWS Glue
- [Como conceder acesso entre contas](#) no Guia do Desenvolvedor AWS Glue (para exemplos de políticas de recursos do Catálogo de Dados)

## Visualizando todas as concessões entre contas usando a operação de GetResourceShares API

Se sua empresa concede permissões entre contas usando uma política de AWS Glue Data Catalog recursos e subsídios do Lake Formation, a única maneira de visualizar todas as concessões entre contas em um só lugar é usando a operação de `glue:GetResourceShares` API.

Quando você concede permissões do Lake Formation em todas as contas usando o método de recurso nomeado, AWS Resource Access Manager (AWS RAM) cria uma política de recursos AWS Identity and Access Management (IAM) e a armazena em sua AWS conta. A política concede as permissões necessárias para acessar o recurso. AWS RAM cria uma política de recursos separada para cada concessão entre contas. Você pode ver todas essas políticas usando a operação da API `glue:GetResourceShares`.

### Note

Essa operação também restaura a política de recursos do Catálogo de Dados. No entanto, se você habilitou a criptografia de metadados nas configurações do Catálogo de Dados e não tem permissão na AWS KMS chave, a operação não retornará a política de recursos do Catálogo de Dados.

Como visualizar todas as concessões entre contas

- Digite o AWS CLI comando a seguir.

```
aws glue get-resource-policies
```

A seguir está um exemplo de política de recursos que AWS RAM cria e armazena quando você concede permissões na tabela t no banco de dados db1 para a AWS conta 1111-2222-3333.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "glue:GetTable",
        "glue:GetTables",
        "glue:GetTableVersion",
        "glue:GetTableVersions",
        "glue:GetPartition",
        "glue:GetPartitions",
        "glue:BatchGetPartition",
        "glue:SearchTables"
      ],
      "Principal": {"AWS": [
        "111122223333"
      ]},
      "Resource": [
        "arn:aws:glue:<region>:111122223333:table/db1/t"
      ]
    }
  ]
}
```

 Consulte também:

- [GetResourceShares Ação \(Python: get\\_resource\\_policies\)](#) no Guia do desenvolvedor AWS Glue

## Acessar e visualizar tabelas e bancos de dados compartilhados do Catálogo de Dados

Para o administrador do data lake e para as entidades principais que receberam permissões, os recursos que são compartilhados com sua conta AWS aparecem no Catálogo de Dados como se fossem recursos em sua conta. O console exibe a conta que tem o recurso.

Você pode visualizar os recursos que são compartilhados com sua conta usando o console do Lake Formation. Você também pode usar o console do AWS Resource Access Manager (AWS RAM) para visualizar os recursos que são compartilhados com sua conta e os recursos que você compartilhou com outras contas AWS usando o método de recurso nomeado.

### Important

Quando alguém usa o método de recurso nomeado para conceder permissões entre contas em um recurso do Catálogo de Dados para sua conta ou organização AWS, o Lake Formation usa o serviço do AWS Resource Access Manager (AWS RAM) para compartilhar o recurso. Se a conta do concedido estiver na mesma organização AWS da conta do concedente, o recurso compartilhado estará disponível imediatamente para o concedido. No entanto, se a sua conta não estiver na mesma organização, o AWS RAM envia um convite à sua conta para aceitar ou rejeitar o compartilhamento do recurso. Em seguida, para disponibilizar o recurso compartilhado, o administrador do data lake deve usar o console do AWS RAM ou a CLI para aceitar o convite na conta do concedido.

O console do Lake Formation exibe um alerta se houver um convite de compartilhamento de recursos do AWS RAM aguardando para ser aceito. Somente usuários autorizados a ver os convites do AWS RAM recebem o alerta.

### Consulte também:

- [Compartilhamento de tabelas e bancos de dados do Catálogo de dados entre contas AWS](#)
- [Compartilhamento de dados entre contas no Lake Formation](#)
- [Como acessar os dados subjacentes de uma tabela compartilhada](#)
- [Controle de acesso a metadados](#) (para obter informações sobre o método de recurso nomeado versus o método LF-TBAC para compartilhar recursos.)

## Tópicos

- [Aceitar um convite de compartilhamento de recursos do AWS RAM](#)
- [Visualizando tabelas e bancos de dados compartilhados do Catálogo de Dados](#)

## Aceitar um convite de compartilhamento de recursos do AWS RAM

Se um recurso do Catálogo de Dados for compartilhado com sua conta AWS e sua conta não estiver na mesma organização AWS da conta de compartilhamento, você não terá acesso ao recurso compartilhado até aceitar um convite de compartilhamento de recursos do AWS Resource Access Manager (AWS RAM). Como administrador do data lake, você deve primeiro consultar o AWS RAM em busca de convites pendentes e depois aceitá-los.

Você pode usar o console do AWS RAM, a API ou a AWS Command Line Interface (AWS CLI) para ver e aceitar convites.

Para ver e aceitar um convite de compartilhamento de recursos do AWS RAM (console)

1. Verifique se você tem as permissões do AWS Identity and Access Management (IAM) necessárias para visualizar e aceitar convites de compartilhamento de recursos.

Para obter informações sobre as políticas de IAM sugeridas para administradores de data lake, consulte [the section called “Permissões de administrador do data lake”](#).

2. Siga as instruções em [Aceitar e rejeitar convites](#) no Guia do usuário do AWS RAM.

Para ver e aceitar um convite de compartilhamento de recursos do AWS RAM (AWS CLI)

1. Verifique se você tem as permissões do AWS Identity and Access Management (IAM) necessárias para visualizar e aceitar convites de compartilhamento de recursos.

Para obter informações sobre as políticas de IAM sugeridas para administradores de data lake, consulte [the section called “Permissões de administrador do data lake”](#).

2. Insira o comando a seguir para visualizar os convites de compartilhamento de recursos pendentes.

```
aws ram get-resource-share-invitations
```

A saída deve ser semelhante a esta.

```
{
  "resourceShareInvitations": [
    {
```

```

        "resourceShareInvitationArn": "arn:aws:ram:us-
east-1:111122223333:resource-share-invitation/a93aa60a-1bd9-46e8-96db-
a4e72eec1d9f",
        "resourceShareName": "111122223333-123456789012-uswuU",
        "resourceShareArn": "arn:aws:ram:us-east-1:111122223333:resource-
share/2a4ab5fb-d859-4751-84f7-8760b35fc1fe",
        "senderAccountId": "111122223333",
        "receiverAccountId": "123456789012",
        "invitationTimestamp": 1589576601.79,
        "status": "PENDING"
    }
]
}

```

Observe o status de PENDING.

3. Copie o valor da chave `resourceShareInvitationArn` para a área de transferência.
4. Cole o valor no comando a seguir `<invitation-arn>`, substitua-o e insira o comando.

```
aws ram accept-resource-share-invitation --resource-share-invitation-
arn <invitation-arn>
```

A saída deve ser semelhante a esta.

```

{
  "resourceShareInvitations": [
    {
      "resourceShareInvitationArn": "arn:aws:ram:us-
east-1:111122223333:resource-share-invitation/a93aa60a-1bd9-46e8-96db-
a4e72eec1d9f",
      "resourceShareName": "111122223333-123456789012-uswuU",
      "resourceShareArn": "arn:aws:ram:us-east-1:111122223333:resource-
share/2a4ab5fb-d859-4751-84f7-8760b35fc1fe",
      "senderAccountId": "111122223333",
      "receiverAccountId": "123456789012",
      "invitationTimestamp": 1589576601.79,
      "status": "ACCEPTED"
    }
  ]
}

```

Observe o status de ACCEPTED.

## Visualizando tabelas e bancos de dados compartilhados do Catálogo de Dados

Você pode visualizar os recursos que são compartilhados com sua conta usando o console do Lake Formation ou a AWS CLI. Você também pode usar o console do AWS Resource Access Manager (AWS RAM) ou a CLI para visualizar os recursos que são compartilhados com sua conta e os recursos que você compartilhou com outras contas AWS.

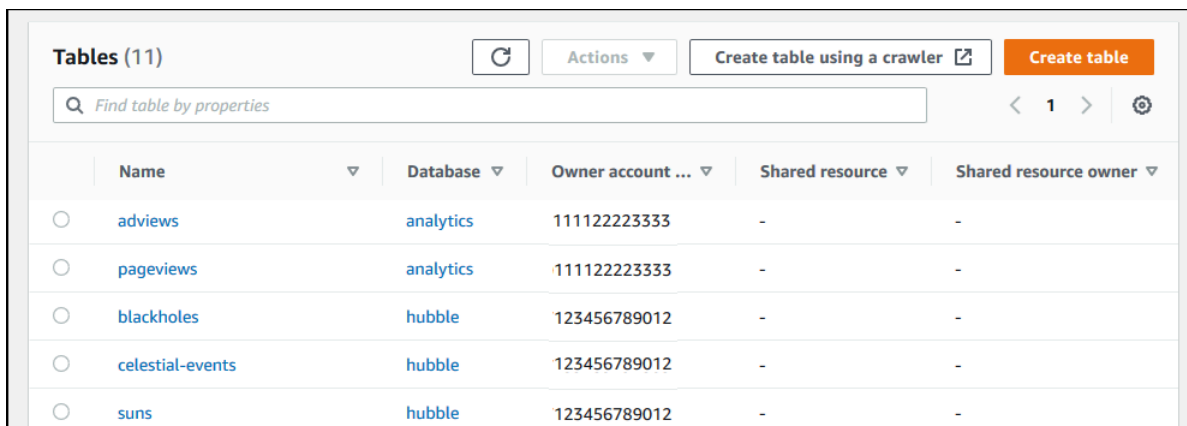
Para visualizar recursos compartilhados usando o console do Lake Formation

1. Abra o console do Lake Formation em <https://console.aws.amazon.com/lakeformation/>.

Faça login como administrador do data lake ou como usuário que recebeu permissões em uma tabela compartilhada.

2. Para visualizar recursos que são compartilhados com sua conta AWS, siga um destes procedimentos:
  - Para visualizar tabelas que são compartilhadas com sua conta, escolha Tabelas no painel de navegação.
  - Para visualizar tabelas que são compartilhadas com sua conta, escolha Banco de dados no painel de navegação.

O console exibe uma lista de bancos de dados ou tabelas em sua conta e compartilhados com sua conta. Para recursos compartilhados com sua conta, o console exibe o ID da conta AWS do proprietário na coluna ID da conta do proprietário (a terceira coluna na captura de tela a seguir).



	Name	Database	Owner account ...	Shared resource	Shared resource owner
<input type="radio"/>	adviews	analytics	111122223333	-	-
<input type="radio"/>	pageviews	analytics	111122223333	-	-
<input type="radio"/>	blackholes	hubble	123456789012	-	-
<input type="radio"/>	celestial-events	hubble	123456789012	-	-
<input type="radio"/>	suns	hubble	123456789012	-	-

3. Para visualizar os recursos que você compartilhou com outras contas AWS ou organizações, no painel de navegação, escolha Permissões de dados.

Os recursos que você compartilhou estão listados na página Permissões de dados com o número da conta externa mostrado na coluna Entidade principal, conforme mostrado na imagem a seguir.

Principal	Principal type	Resource type	Resource	Owner account ID	Permissions
<input type="radio"/> datalake_admin	IAM user	Table	clickthroughs	123456789012	Super, Alter, Delete, Drop, Insert
<input type="radio"/> datalake_admin	IAM user	Column	analytics.clickthroughs.*	123456789012	Select
<input type="radio"/> 111122223333	AWS account	Table	clickthroughs	123456789012	Insert
<input type="radio"/> 111122223333	AWS account	Column	analytics.clickthroughs.*	123456789012	Select

Para visualizar os recursos compartilhados usando o console do AWS RAM

1. Verifique se você tem as permissões do AWS Identity and Access Management (IAM) necessárias para visualizar recursos compartilhados usando o AWS RAM.

No mínimo, você deve ter a permissão `ram:ListResources`. Essa permissão está incluída na política gerenciada pela AWS `AWSLakeFormationCrossAccountManager`.

2. Faça login no AWS Management Console e abra o console do AWS RAM em <https://console.aws.amazon.com/ram>.
3. Faça um dos seguintes procedimentos:
  - Para ver os recursos que você compartilhou, no painel de navegação, em Compartilhado por mim, escolha Recursos compartilhados.
  - Para ver os recursos compartilhados com você, no painel de navegação, em Compartilhado comigo, escolha Recursos compartilhados.



# Criação de links de recursos

Links de recursos são objetos do Catálogo de Dados que são links para bancos de dados e tabelas de metadados — normalmente para bancos de dados e tabelas compartilhados de outras contas da AWS. Eles ajudam a permitir o acesso entre contas aos dados no data lake em todas as regiões da AWS.

## Note

O Lake Formation suporta a consulta de tabelas do Catálogo de Dados em todas as regiões da AWS. Você pode acessar os bancos de dados e tabelas do Catálogo de Dados de qualquer região da AWS, criando links de recursos nessas regiões que apontam para bancos de dados e tabelas compartilhados em diferentes regiões.

## Tópicos

- [Como os links de recursos funcionam no Lake Formation](#)
- [Como criar um link de recurso para uma tabela compartilhada do Catálogo de Dados](#)
- [Como criar um link de recurso para um banco de dados compartilhado do Catálogo de Dados](#)
- [Gestão de links de recursos em APIs do AWS Glue.](#)

## Como os links de recursos funcionam no Lake Formation

Um link de recurso é um objeto de Catálogo de Dados que é um link para um banco de dados, ou tabela local, ou compartilhada. Depois de criar um link de recurso para um banco de dados ou tabela, você pode usar o nome do link de recurso onde quer que use o nome do banco de dados ou da tabela. Juntamente com as tabelas que você possui ou que são compartilhadas com você, os links de recursos de tabela são retornados pelo `glue:GetTables()` e aparecem como entradas na página Tabelas do console do Lake Formation. Os links de recursos para bancos de dados agem de maneira semelhante.

A criação de um link de recurso para um banco de dados ou tabela permite que você:

- Atribua um nome diferente a um banco de dados ou tabela em seu Catálogo de Dados. Isso é especialmente útil se contas da AWS diferentes compartilharem bancos de dados ou tabelas com o mesmo nome, ou se vários bancos de dados em sua conta tiverem tabelas com o mesmo nome.

- Acesse os bancos de dados e tabelas do Catálogo de Dados de qualquer região da AWS, criando links de recursos nessas regiões apontando para o banco de dados e tabelas em outra região. Você pode executar consultas em qualquer região com esses links de recursos usando o Athena, o Amazon EMR e executar trabalhos do AWS Glue no ETL Spark, sem copiar os dados de origem nem os metadados no Catálogo de Dados Glue.
- Use os serviços integrados da AWS, como o Amazon Athena e o Amazon Redshift Spectrum, para executar consultas que acessam bancos de dados ou tabelas compartilhados. Alguns serviços integrados não podem acessar diretamente bancos de dados ou tabelas entre contas. No entanto, eles podem acessar links de recursos em sua conta para bancos de dados e tabelas em outras contas.

### Note


Você não precisa criar um link de recurso para referenciar um banco de dados ou tabela compartilhados em scripts de extração, transformação e carregamento (ETL) do AWS Glue. No entanto, para evitar ambiguidades quando várias contas da AWS compartilham um banco de dados ou tabela com o mesmo nome, você pode criar e usar um link de recurso ou especificar a ID do catálogo ao invocar operações de ETL.

O exemplo a seguir mostra a página Tabelas do console Lake Formation, que lista dois links de recursos. Os nomes dos links de recursos são sempre exibidos em itálico. Cada link de recurso é exibido junto com o nome e o proprietário do recurso compartilhado vinculado. Neste exemplo, um administrador de data lake na conta da AWS 1111-2222-3333 compartilhou as tabelas `inventory` e `incidents` com a conta 1234-5678-9012. Em seguida, um usuário dessa conta criou links de recursos para essas tabelas compartilhadas.

Tables (30)					
Name	Database	Owner account ...	Shared resource	Shared resource owner	
<i>inventory-link</i>	retail	123456789012	inventory	111122223333	<input type="radio"/>
<i>incidents-link</i>	issues-local	123456789012	incidents	111122223333	<input type="radio"/>
site-logs	logs	123456789012	-	-	<input type="radio"/>
alexa-logs	logs	123456789012	-	-	<input type="radio"/>

A seguir estão notas e restrições sobre links de recursos:

- Os links de recursos são necessários para permitir que serviços integrados, como Athena e Redshift Spectrum, consultem os dados subjacentes das tabelas compartilhadas. As consultas nesses serviços integrados são construídas com base nos nomes dos links de recursos.
- Supondo que a configuração Usar somente o controle de acesso do IAM para novas tabelas nesse banco de dados esteja desativada para o banco de dados que o contém, somente a entidade principal que criou um link de recurso pode visualizá-lo e acessá-lo. Para permitir que outras entidades principais da sua conta acessem um link de recurso, conceda a permissão DESCRIBE nela. Para permitir que outras pessoas descartem um link de recurso, conceda a permissão DROP para ele. Os administradores do data lake podem acessar todos os links de recursos na conta. Para descartar um link de recurso criado por outra entidade principal, o administrador do data lake deve primeiro conceder a si a permissão DROP no link do recurso. Para obter mais informações, consulte [Referência de permissões do Lake Formation](#).

 Important

Conceder permissões em um link de recurso não concede permissões no banco de dados ou tabela de destino (vinculado). Você deve conceder permissões no destino separadamente.

- Para criar um link de recurso, você precisa da permissão CREATE\_TABLE ou CREATE\_DATABASE do Lake Formation, e da permissão glue:CreateTable ou glue:CreateDatabase AWS Identity and Access Management do IAM.
- Você pode criar links de recursos para recursos locais (próprios) do Catálogo de Dados, bem como para recursos compartilhados com sua conta da AWS.
- Quando você cria um link de recurso, nenhuma verificação é executada para ver se o recurso compartilhado de destino existe ou se você tem permissões entre contas no recurso. Isso permite que você crie o link do recurso e o recurso compartilhado em qualquer ordem.
- Se você excluir um link de recurso, o recurso compartilhado vinculado não será descartado. Se você descartar um recurso compartilhado, os links de recursos para esse recurso não serão excluídos.
- É possível criar cadeias de links de recursos. No entanto, não há vantagem em fazer isso, pois as APIs seguem apenas o primeiro link de recurso.

 Consulte também:

- [Concedendo e revogando permissões nos recursos do Catálogo de dados](#)

## Como criar um link de recurso para uma tabela compartilhada do Catálogo de Dados

Você pode criar um link de recurso para uma tabela compartilhada em qualquer região da AWS usando o console do AWS Lake Formation, a API, ou o AWS Command Line Interface (AWS CLI).

Como criar um link de recurso para a tabela compartilhada

1. Abra o console do AWS Lake Formation em <https://console.aws.amazon.com/lakeformation/>. Faça login como entidade principal que tem a permissão `CREATE_TABLE` do Lake Formation no banco de dados para conter o link do recurso.
2. No painel de navegação, selecione Tables (Tabelas) e Create table (Criar tabela).
3. Na página Criar tabela, selecione o quadro Link do recurso e, em seguida, forneça as seguintes informações:

Nome do link de recurso

Digite um nome que siga as mesmas regras de um nome de tabela. O nome pode ser o mesmo da tabela compartilhada de destino.

Banco de dados

O banco de dados no Catálogo de Dados local para conter o link do recurso.

Região do proprietário de tabela compartilhada

Se você estiver criando o link do recurso em uma região diferente, selecione a região da tabela compartilhada de destino.

Tabela compartilhada

Selecione uma tabela compartilhada na lista ou digite um nome de tabela local (de propriedade) ou compartilhada.

A lista contém todas as tabelas compartilhadas com sua conta. Anote o banco de dados e o ID da conta do proprietário listado em cada tabela. Caso você não veja uma tabela que saiba que foi compartilhada com sua conta, verifique o seguinte:

- Se você não for administrador do data lake, verifique se o administrador do data lake concedeu a você as permissões do Lake Formation na tabela.
- Se você for administrador de um data lake e sua conta não estiver na mesma organização da AWS da conta concedente, certifique-se de ter aceitado o convite de compartilhamento de recursos AWS Resource Access Manager (AWS RAM) para o banco de dados. Para obter mais informações, consulte [Aceitar um convite de compartilhamento de recursos do AWS RAM](#).

#### Banco de dados da tabela compartilhada

Se você selecionou uma tabela compartilhada na lista, esse campo será preenchido com o banco de dados da tabela compartilhada na conta externa. Caso contrário, digite um banco de dados local (para um link de recurso para uma tabela local) ou o banco de dados da tabela compartilhada na conta externa.

#### Proprietário de tabela compartilhada

Se você selecionou uma tabela compartilhada na lista, esse campo será preenchido com o ID da conta do proprietário da tabela compartilhada. Caso contrário, digite o ID da sua conta da AWS (para um link de recurso para uma tabela local) ou o ID da conta da AWS que compartilhou a tabela.

AWS Lake Formation &gt; Tables &gt; Create table

## Create table

### Table details

Create a table in the AWS Glue Data Catalog.

Table  
Create a table in my account.

Resource link  
Create a resource link to a shared table.

#### Resource link name

Name may contain letters (A-Z), numbers (0-9), hyphens (-), or underscores (\_), and must be less than 256 characters long.

#### Database

Resource link will be contained in this database.

#### Shared table owner region

Select the region where the table is shared

#### Shared table

Enter or choose a shared table.

#### Shared table's database

Enter the database containing the shared table.

#### Shared table's owner ID

Enter the AWS account ID of the shared table owner.

Cancel

Create

4. Selecione Criar para criar o link do recurso.

Em seguida, você pode exibir o nome do link do recurso na coluna Nome na página Tabelas.

5. (Opcional) Conceda a permissão DESCRIBE do Lake Formation no link do recurso às entidades principais que devem ser capazes de exibir o link e acessar a tabela de destino.

Como criar um link de recurso para uma tabela compartilhada na mesma região (AWS CLI)

1. Digite um comando semelhante ao seguinte:

```
aws glue create-table --database-name myissues --table-input
'{"Name":"my_customers","TargetTable":
{"CatalogId":"111122223333","DatabaseName":"issues","Name":"customers"}}'
```

Esse comando cria um link de recurso chamado `my_customers` para a tabela compartilhada `customers`, que está no banco de dados `issues` na conta AWS 1111-2222-3333. O link do recurso é armazenado no banco de dados local `myissues`.

2. (Opcional) Conceda a permissão DESCRIBE do Lake Formation no link do recurso às entidades principais que devem ser capazes de exibir o link e acessar a tabela de destino.

Como criar um link de recurso para uma tabela compartilhada em uma região diferente (AWS CLI)

1. Digite um comando semelhante ao seguinte:

```
aws glue create-table --region eu-west-1 --cli-input-json '{
  "CatalogId": "111122223333",
  "DatabaseName": "ireland_db",
  "TableInput": {
    "Name": "rl_useast1salestb_ireland",
    "TargetTable": {
      "CatalogId": "444455556666",
      "DatabaseName": "useast1_salesdb",
      "Region": "us-east-1",
      "Name": "useast1_salestb"
    }
  }
}'
```

Esse comando cria um link de recurso nomeado `rl_useast1salestb_ireland` na região da Europa (Irlanda) para a tabela compartilhada `useast1_salestb`, que está no banco de dados `useast1_salesdb` na conta AWS 444455556666 na região Leste dos EUA (Norte da Virgínia). O link do recurso é armazenado no banco de dados local `ireland_db`.

2. Conceda permissão DESCRIBE ao Lake Formation às entidades principais que devem ser capazes de exibir o link e acessar o destino do link por meio do link.

 Consulte também:

- [Como os links de recursos funcionam no Lake Formation](#)
- [DESCRIBE](#)

## Como criar um link de recurso para um banco de dados compartilhado do Catálogo de Dados

Você pode criar um link de recurso para um banco de dados compartilhado usando o console do AWS Lake Formation, a API, ou AWS Command Line Interface (AWS CLI).

Como criar um link de recurso para um banco de dados compartilhado (console)

1. Abra o console do AWS Lake Formation em <https://console.aws.amazon.com/lakeformation/>. Faça login como administrador de data lake ou como criador de banco de dados.

Um criador de banco de dados é um entidade principal que recebeu a permissão CREATE\_DATABASE do Lake Formation.

2. No painel de navegação, selecione Bancos de dados e, em seguida, selecione Criar banco de dados.
3. Na página Criar banco de dados, selecione o quadro Link do recurso e, em seguida, forneça as seguintes informações:

Nome do link de recurso

Digite um nome que siga as mesmas regras de um nome de banco de dados. O nome pode ser o mesmo do banco de dados compartilhado de destino.

Região do proprietário do banco de dados compartilhado

Se você estiver criando o link do recurso em uma região diferente, selecione a região do banco de dados compartilhado de destino.



## Banco de dados compartilhado

Selecione um banco de dados na lista ou digite um nome de banco de dados local (de propriedade) ou compartilhado.

A lista contém todos os bancos de dados compartilhados com sua conta. Anote o ID da conta do proprietário que está listado em cada banco de dados. Caso você não veja um banco de dados que você sabe que foi compartilhado com sua conta, verifique o seguinte:

- Se você não for administrador do data lake, verifique se o administrador do data lake concedeu a você as permissões do Lake Formation no banco de dados.
- Se você for administrador de um data lake e sua conta não estiver na mesma organização da AWS da conta concedente, certifique-se de ter aceitado o convite de compartilhamento de recursos AWS Resource Access Manager (AWS RAM) para o banco de dados. Para obter mais informações, consulte [Aceitar um convite de compartilhamento de recursos do AWS RAM](#).

## Proprietário do banco de dados compartilhado

Se você selecionou um banco de dados compartilhado na lista, esse campo será preenchido com o ID da conta do proprietário do banco de dados compartilhado. Caso contrário, digite o ID da sua conta da AWS (para um link de recurso para um banco de dados local) ou o ID da conta da AWS que compartilhou o banco de dados.

[AWS Lake Formation](#) > [Databases](#) > [Create database](#)

## Create database

**Database details**  
Create a database in the AWS Glue Data Catalog.

Database  
Create a database in my account.

Resource link  
Create a resource link to a shared database.

**Resource link name**  
rl\_useast1shared\_irelanddb

Name may contain letters (A-Z), numbers (0-9), hyphens (-), or underscores (\_), and must be less than 256 characters long.

**Shared database owner region**  
Select the region where the database is shared

US East (N. Virginia) ▼

**Shared database**  
Enter or choose a shared database.

Q useast1shared\_db X

**Shared database's owner ID**  
Enter the AWS account ID of the shared database owner.

444455556666

[Cancel](#) [Create](#)

4. Selecione Criar para criar o link do recurso.

Em seguida, você pode exibir o nome do link do recurso na coluna Nome na página Bancos de dados.

5. (Opcional) Conceda a permissão DESCRIBE do Lake Formation no link do recurso às entidades principais da região da Europa (Irlanda), que devem ser capazes de exibir o link e acessar o banco de dados de destino.

## Como criar um link de recurso para um banco de dados compartilhado na mesma região (AWS CLI)

1. Digite um comando semelhante ao seguinte:

```
aws glue create-database --database-input '{"Name":"myissues","TargetDatabase":
{"CatalogId":"111122223333","DatabaseName":"issues"}}'
```

Esse comando cria um link de recurso chamado `myissues` para o banco de dados compartilhado `issues`, que está na conta da AWS 1111-2222-3333.

2. (Opcional) Conceda a permissão `DESCRIBE` do Lake Formation às entidades principais no link do recurso que devem ser capazes de exibir o link e acessar o banco de dados de destino.

## Como criar um link de recurso para um banco de dados compartilhado em uma região diferente (AWS CLI)

1. Digite um comando semelhante ao seguinte:

```
aws glue create-database --region eu-west-1 --cli-input-json '{
  "CatalogId": "111122223333",
  "DatabaseInput": {
    "Name": "rl_useast1shared_irelanddb",
    "TargetDatabase": {
      "CatalogId": "444455556666",
      "DatabaseName": "useast1shared_db",
      "Region": "us-east-1"
    }
  }
}'
```

Esse comando cria um link de recurso chamado `rl_useast1shared_irelanddb` na conta da AWS 111122223333 na região Europa (Irlanda) para o banco de dados compartilhado `useast1shared_db`, que está na conta da AWS 444455556666 na região Leste dos EUA (Norte da Virgínia).

2. Conceda a permissão `DESCRIBE` do Lake Formation às entidades principais da região da Europa (Irlanda), que devem ser capazes de exibir o link e acessar o destino do link por meio do link.

 Consulte também:

- [Como os links de recursos funcionam no Lake Formation](#)
- [DESCRIBE](#)

## Gestão de links de recursos em APIs do AWS Glue.

As tabelas a seguir explicam como as APIs do Catálogo de Dados do AWS Glue lidam com links de recursos de banco de dados e tabelas. Para todas as operações de API do Get\*, somente bancos de dados e tabelas nos quais o chamador tem permissões são retornados. Além disso, ao acessar um banco de dados ou tabela de destino por meio de um link de recurso, você deve ter as permissões do AWS Identity and Access Management (IAM) e Lake Formation no link de destino e no link do recurso. A permissão do Lake Formation exigida nos links de recursos é DESCRIBE. Para obter mais informações, consulte [DESCRIBE](#).

### Operações de API de banco de dados

Operação de API	Gestão de links de recursos
CreateDatabase	Se o banco de dados for um link de recurso, ele cria o link de recurso para o banco de dados de destino designado.
UpdateDatabase	Se o banco de dados designado for um link de recurso, ele segue o link e atualiza o banco de dados de destino. Se o link de recurso precisar ser modificado para vincular a um banco de dados diferente, você deverá excluí-lo e criar um novo.
DeleteDatabase	Exclua o link do recurso. Isto não exclui o banco de dados vinculado (de destino).
GetDatabase	Se o chamador tiver permissões no alvo, ele segue o link para retornar as propriedades do alvo. Caso contrário, ele retornará as propriedades do link.
GetDatabases	Retorna uma lista de bancos de dados, incluindo links de recursos. Para cada link de recurso no conjunto de resultados, a operação segue o link para obter as propriedades do link de destino. Você

Operação de API	Gestão de links de recursos
	deve especificar <code>ResourceShareType = ALL</code> para ver os bancos de dados compartilhados com sua conta.

## Operações de API de tabela

Operação de API	Gestão de links de recursos
<code>CreateTable</code>	Se o banco de dados for um link de recurso, ele segue o link do banco de dados e cria uma tabela no banco de dados de destino. Se a tabela for um link de recurso, a operação criará o link de recurso no banco de dados designado. Não há suporte para criar um link de recurso de tabela por meio de um link de recurso de banco de dados.
<code>UpdateTable</code>	Se a tabela ou o banco de dados designado for um link de recurso, isso atualiza a tabela de destino. Se a tabela e o banco de dados forem links de recursos, a operação falhará.
<code>DeleteTable</code>	Se o banco de dados designado for um link de recurso, ele segue o link e exclui a tabela ou o link do recurso da tabela no banco de dados de destino. Se a tabela for um link de recurso, a operação excluirá o link de recurso da tabela no banco de dados designado. A exclusão de um link de recurso de tabela não exclui a tabela de destino.
<code>BatchDeleteTable</code>	Igual a <code>DeleteTable</code> .
<code>GetTable</code>	Se o banco de dados designado for um link de recurso, ele segue o link do banco de dados e retorna a tabela ou o link do recurso da tabela do banco de dados de destino. Caso contrário, se a tabela for um link de recurso, a operação seguirá o link e retornará as propriedades da tabela de destino.
<code>GetTables</code>	Se o banco de dados designado for um link de recurso, ele segue o link do banco de dados e retorna as tabelas e os links de recursos da tabela do banco de dados de destino. Se o banco de dados de

Operação de API	Gestão de links de recursos
	destino for um banco de dados compartilhado de outra conta da AWS, a operação retornará somente as tabelas compartilhadas nesse banco de dados. Ele não segue os links de recursos da tabela no banco de dados de destino. Caso contrário, se o banco de dados designado for um banco de dados local (de propriedade), a operação retornará todas as tabelas no banco de dados local e seguirá cada link de recurso da tabela para retornar as propriedades da tabela de destino.
SearchTables	Retorna tabelas e links de recursos de tabelas. Ele não segue links para retornar as propriedades da tabela de destino. Você deve especificar <code>ResourceShareType = ALL</code> para ver as tabelas compartilhadas com sua conta.
GetTableVersion	Igual a GetTable.
GetTableVersions	Igual a GetTable.
DeleteTableVersion	Igual a DeleteTable .
BatchDeleteTableVersion	Igual a DeleteTable .

### Operações de API de partição

Operação de API	Gestão de links de recursos
CreatePartition	Se o banco de dados designado for um link de recurso, ele segue o link do banco de dados e cria uma partição na tabela designada no banco de dados de destino. Se a tabela for um link de recurso, a operação segue o link do recurso e cria a partição na tabela de destino. A criação de uma partição por meio de um link de recurso de tabela e de um link de recurso de banco de dados não é suportada.

Operação de API	Gestão de links de recursos
BatchCreatePartiti on	Igual a CreatePartition .
UpdatePartition	Se o banco de dados designado for um link de recurso, ele segue o link do banco de dados e atualiza a partição na tabela designada no banco de dados de destino. Se a tabela for um link de recurso, a operação segue o link do recurso e atualiza a partição na tabela de destino. A atualização de uma partição por meio de um link de recurso de tabela e de um link de recurso de banco de dados não é suportada.
DeletePartition	Se o banco de dados designado for um link de recurso, ele segue o link do banco de dados e exclui a partição na tabela designada no banco de dados de destino. Se a tabela for um link de recurso, a operação segue o link do recurso e exclui a partição na tabela de destino. A exclusão de uma partição por meio de um link de recurso de tabela e de um link de recurso de banco de dados não é suportada.
BatchDeletePartiti on	Igual a DeletePartition .
GetPartition	Se o banco de dados designado for um link de recurso, ele segue o link do banco de dados e retorna as informações da partição da tabela designada. Caso contrário, se a tabela for um link de recurso, a operação seguirá o link e retornará as informações da partição. Se a tabela e o banco de dados forem links de recursos, ele retornará um conjunto de resultados vazio.
GetPartitions	Se o banco de dados designado for um link de recurso, ele segue o link do banco de dados e retorna as informações de partição de todas as partições na tabela designada. Caso contrário, se a tabela for um link de recurso, a operação seguirá o link e retornará as informações da partição. Se a tabela e o banco de dados forem links de recursos, ele retornará um conjunto de resultados vazio.

Operação de API	Gestão de links de recursos
BatchGetPartition	Igual a GetPartition .

Operações de API de perfis definidas pelo usuário

Operação de API	Gestão de links de recursos
(Todas as operações de API)	Se o banco de dados for um link de recurso, ele segue o link do recurso e executa a operação no banco de dados de destino.

 Consulte também:

- [Como os links de recursos funcionam no Lake Formation](#)

## Acessar tabelas entre regiões

O Lake Formation suporta a consulta de tabelas do Catálogo de Dados em todas as regiões da AWS. Você pode acessar dados em uma região de outras regiões usando o Amazon Athena, o Amazon EMR e o ETL do AWS Glue [criando links de recursos](#) em outras regiões apontando para os bancos de dados e tabelas de origem. Com o acesso à tabela entre regiões, você pode acessar dados entre regiões sem copiar os dados subjacentes ou os metadados no Catálogo de Dados.

Por exemplo, você pode compartilhar um banco de dados ou uma tabela em uma conta de produtor com uma conta de consumidor na Região A. Após aceitar o convite de compartilhamento de recursos na Região A, o administrador do data lake da conta do consumidor pode criar links de recursos para o recurso compartilhado na Região B. O administrador da conta do consumidor pode conceder permissões sobre o recurso compartilhado com as entidades principais do IAM nessa conta na Região A, e conceder permissões de link de recurso na Região B. Usando o link do recurso, as entidades principais da conta do consumidor podem consultar os dados compartilhados da Região B.

Você também ser host da fonte de dados do Amazon S3 na Região A em uma conta de produtor, e registrar o local dos dados em uma conta central na Região B. Você pode criar recursos do Catálogo de Dados na conta central, configurar permissões do Lake Formation, e compartilhar dados com consumidores em sua conta ou com contas externas na Região B. O atributo entre regiões



permite que os usuários acessem essas tabelas do Catálogo de Dados da Região C usando links de recursos.

Usando esse atributo, você pode consultar bancos de dados federados em Apache Hive repositórios entre regiões e também unir tabelas na região local com tabelas em outra região ao executar consultas.

O Lake Formation oferece suporte aos seguintes recursos com acesso a tabelas entre regiões:

- Controle de acesso baseado em LF-tags
- Permissões de acesso refinado
- Operações de gravação no banco de dados compartilhado ou na tabela com as permissões apropriadas
- Compartilhamento de dados entre contas no nível da conta e direto com as entidades principais do IAM

Usuários não administrativos com permissões `Create_Database` e permissões `Create_Table` podem criar links de recursos entre regiões.

#### Note

Você pode criar links de recursos entre regiões em qualquer região e acessar dados sem aplicar as permissões do Lake Formation. Para dados de origem no Amazon S3 que não estão registrados no Lake Formation, o acesso é determinado pelas políticas de permissões do IAM para ações do Amazon S3 e do AWS Glue.

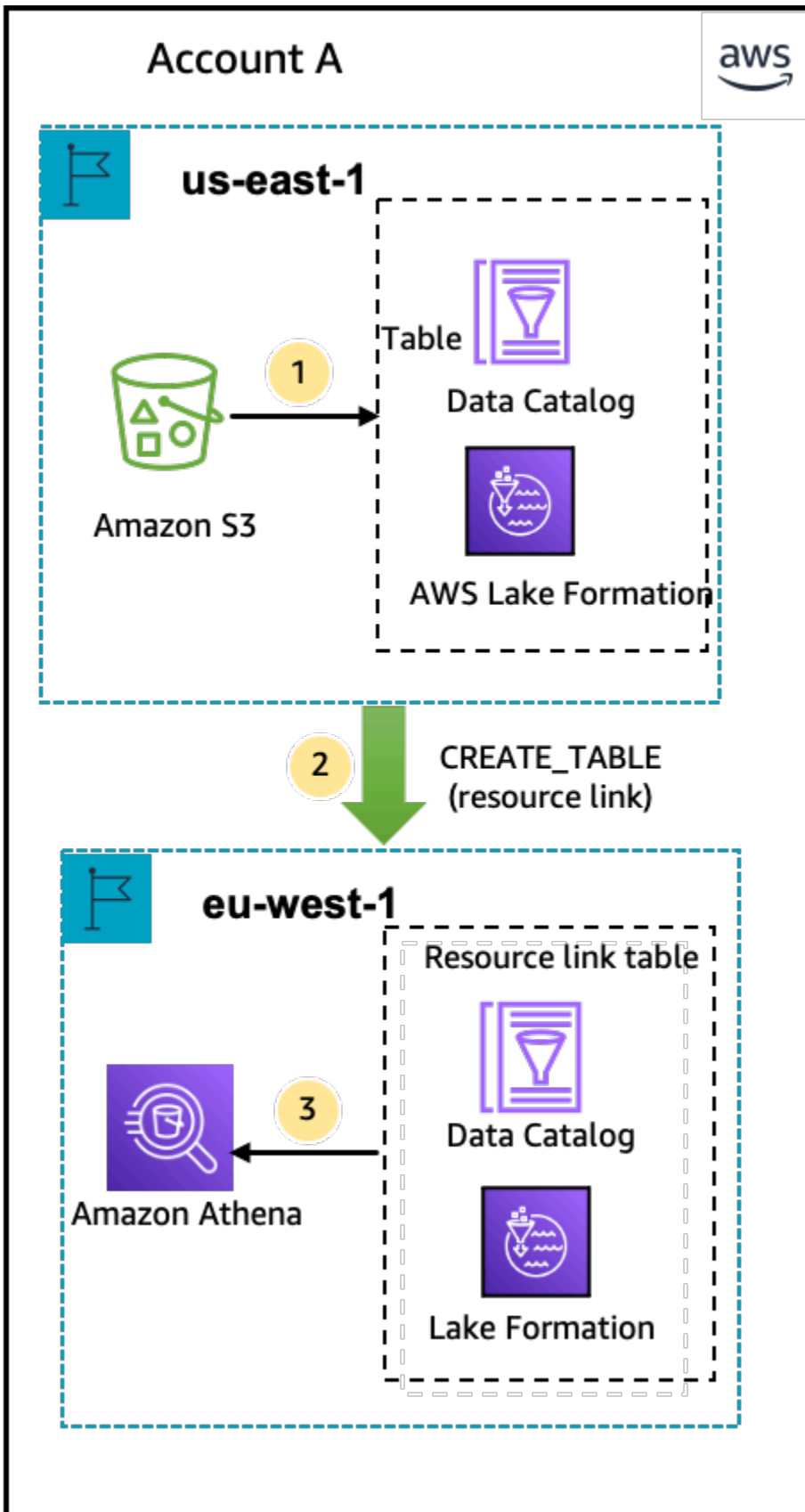
Para obter limitações, consulte [Limitações de acesso aos dados entre regiões](#).

## Fluxos de trabalho

Os diagramas a seguir mostram os fluxos de trabalho para acessar dados entre regiões da AWS da mesma conta da AWS e de uma conta externa.

### Fluxo de trabalho para acessar tabelas compartilhadas na mesma conta AWS

No diagrama abaixo, os dados são compartilhados com um usuário da mesma conta da AWS na região Leste dos EUA (Norte da Virgínia), e o usuário consulta os dados compartilhados na região Europa (Irlanda).



O administrador do data lake executa as seguintes atividades (etapas 1 e 2):

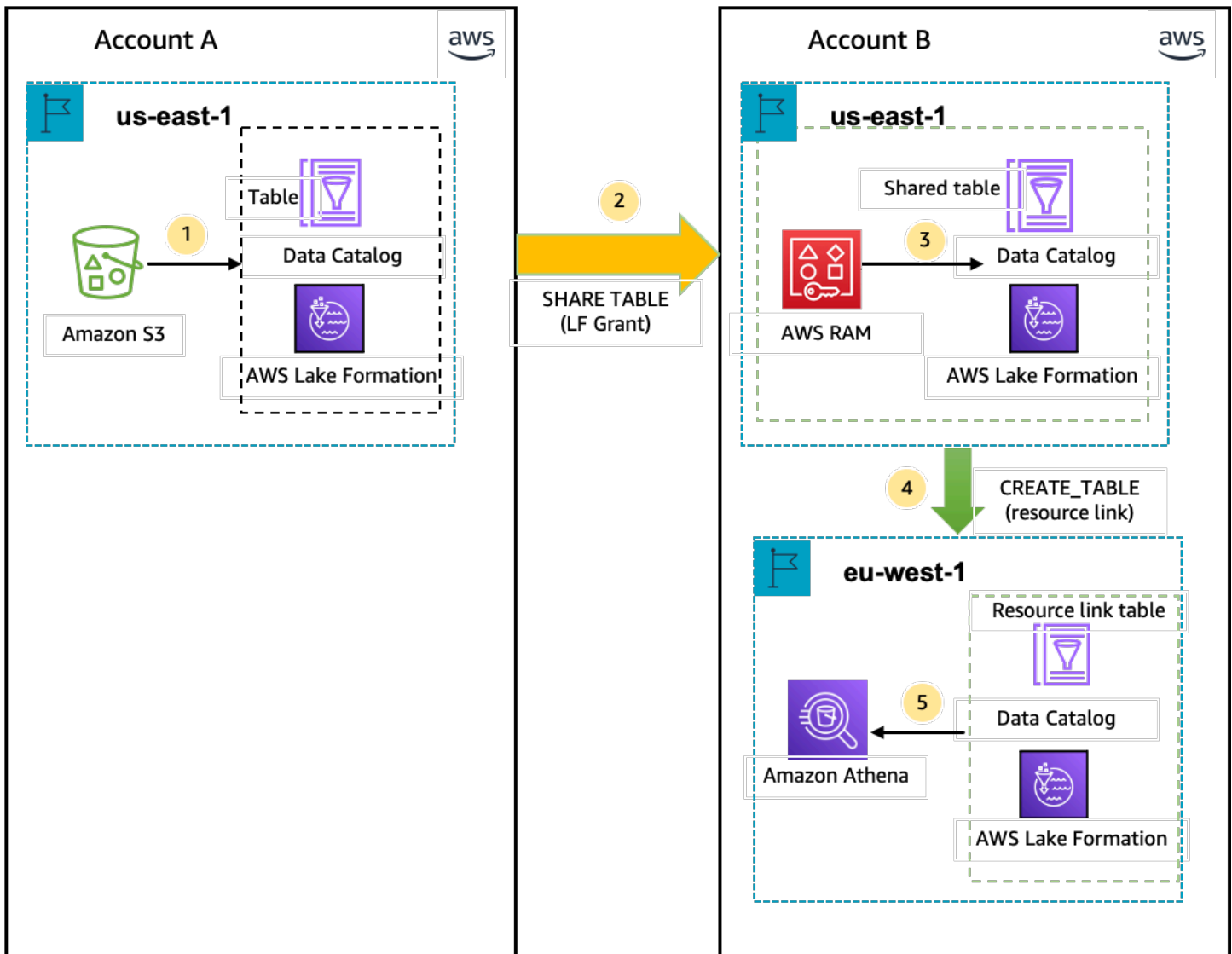
1. Um administrador de data lake configura uma conta da AWS com os bancos de dados e tabelas do Catálogo de Dados e registra um local de dados do Amazon S3 no Lake Formation na região Leste dos EUA (Norte da Virgínia).

Conceder a permissão `Select` em um recurso do Catálogo de Dados (tabela de produtos no diagrama) a uma entidade principal (usuário) na mesma conta.

2. Cria um link de recurso na região da Europa (Irlanda) apontando para a tabela de origem na região Leste dos EUA (Norte da Virgínia). Conceder a permissão `DESCRIBE` no link do recurso da região da Europa (Irlanda) à entidade principal.
3. O usuário consulta a tabela da região da Europa (Irlanda) usando Athena.

Fluxo de trabalho para acessar tabelas compartilhadas com uma conta externa da AWS.

No diagrama abaixo, a conta do produtor (Conta A) hospeda o bucket Amazon S3, registra o local dos dados e compartilha uma tabela do Catálogo de Dados com uma conta de consumidor (Conta B) na região Leste dos EUA (Norte da Virgínia) e um usuário da conta do consumidor (Conta B) consulta a tabela da região Europa (Irlanda).



1. Um administrador do data lake configura uma conta da AWS (conta do produtor) com os recursos do Catálogo de Dados e um local de dados do Amazon S3 registrado no Lake Formation na região Leste dos EUA (Norte da Virgínia).
2. O administrador do data lake da conta do produtor compartilha uma tabela do Catálogo de Dados com uma conta de consumidor.
3. O administrador do data lake da conta do consumidor aceita o convite de compartilhamento de dados na região Leste dos EUA (Norte da Virgínia) e concede a permissão `Select` na tabela compartilhada a uma entidade principal da mesma região.
4. O administrador do data lake da conta do consumidor cria um link de recurso na região da Europa (Irlanda) apontando para a tabela compartilhada de destino na região Leste dos EUA (Norte

da Virgínia) e concede ao usuário a permissão DESCRIBE no link do recurso da região Europa (Irlanda).

5. O usuário consulta os dados da região da Europa (Irlanda) usando o Athena.

## Como configurar o acesso à tabela entre regiões

Para acessar dados de uma região diferente, você precisa primeiro configurar os bancos de dados e as tabelas do Catálogo de Dados na região em que você registra seu local de dados do Amazon S3. Você pode compartilhar os bancos de dados e tabelas do Catálogo de Dados com as entidades principais na sua conta ou em outra. Em seguida, você precisa criar administradores de data lake que possam criar links de recursos apontando para o local de destino dos dados compartilhados nas regiões onde os usuários consultam os dados.

Como consultar dados compartilhados na mesma conta de uma região diferente

Nesta seção, a região da tabela compartilhada de destino é chamada de Região A e os usuários executam consultas na Região B.

1. Configuração da conta na Região A (onde você cria e compartilha os dados)

O administrador do data lake precisa concluir as ações a seguir:

a. Registre um local de dados do Amazon S3.

Para obter mais informações, consulte [Adicionar uma localização do Amazon S3 ao seu data lake](#).

b. Crie bancos de dados e tabelas na conta. Isso também pode ser feito por um usuário não administrativo que tenha permissões para criar bancos de dados e tabelas.

c. Conceda permissões de dados em uma tabela às entidades principais com `Grantable permissions`.

Para obter mais informações, consulte, [Concedendo e revogando permissões nos recursos do Catálogo de dados](#).

2. Configuração da conta na Região B (onde você acessa os dados)

O administrador do data lake precisa concluir as ações a seguir:

a. Crie um link de recurso na Região B apontando para a tabela compartilhada de destino na Região A. Especifique a Região proprietária da tabela compartilhada na tela Criar tabela.

## Create table

**Table details**  
Create a table in the AWS Glue Data Catalog.

**Table**  
Create a table in my account.

**Resource link**  
Create a resource link to a shared table.

**Resource link name**  
  
Name may contain letters (A-Z), numbers (0-9), hyphens (-), or underscores (\_), and must be less than 256 characters long.

**Database**  
Resource link will be contained in this database.

**Shared table owner region**  
Select the region where the table is shared

**Shared table**  
Enter or choose a shared table.

**Shared table's database**  
Enter the database containing the shared table.

**Shared table's owner ID**  
Enter the AWS account ID of the shared table owner.

**Cancel** **Create**

Para obter instruções sobre como criar links de recursos para bancos de dados e tabelas, consulte [Criação de links de recursos](#).

- b. Conceda a permissão `Describe` às entidades principais do IAM no link do recurso na Região B.

Para obter mais informações sobre a concessão de permissões em links de recursos, consulte [Como conceder permissões de links de recursos](#).

As entidades principais do IAM na Região B podem consultar a tabela de destino por meio do link usando o Athena.

Como acessar dados de várias contas de uma região diferente

## 1. Configuração da conta do produtor/concedente

O administrador do data lake precisa concluir as ações a seguir:

- a. Configure a conta do produtor/concedente na Região A.
- b. Registre um local de dados do Amazon S3 na Região A.
- c. Criar bancos de dados e tabelas. Isso pode ser feito por um usuário não administrativo que tenha permissões para criar tabelas.
- d. Conceda permissões de dados para a conta do consumidor/beneficiário em uma tabela na Região A com `Grantable permissions`.

Para obter mais informações, consulte [Compartilhamento de tabelas e bancos de dados do Catálogo de Dados entre Contas da AWS e entidades principais do IAM a partir de contas externas](#).

## 2. Configuração da conta do consumidor/beneficiário

O administrador do data lake precisa concluir as ações a seguir:

- a. Aceite o convite de compartilhamento de recursos do AWS RAM da Região A.
- b. Crie um link de recurso na Região B apontando para a tabela compartilhada. A região B é onde os usuários desejam consultar a tabela.
- c. Conceda permissões de dados na tabela compartilhada às entidades principais do IAM na Região A.

### Note

Você deve conceder permissões para a tabela compartilhada na mesma região em que a tabela foi compartilhada.

- d. Conceda permissões às entidades principais no link do recurso na Região B.

As entidades principais da conta do consumidor na Região B então consultam a tabela compartilhada da Região B usando o Athena.



# Compartilhamento de dados no AWS Lake Formation

Você pode usar o recurso de compartilhamento de AWS Lake Formation dados para conceder e gerenciar permissões sobre dados armazenados em locais diferentes do Amazon S3 e metadados armazenados em locais diferentes do. AWS Glue Data Catalog Com o recurso de compartilhamento de dados, você pode configurar e gerenciar permissões em conjuntos de dados no Amazon Redshift sem migrar os dados para o Amazon S3. Você também pode usar o recurso de federação do catálogo de dados para se conectar a metástores externos.

Depois, você pode usar o Lake Formation para gerenciar dados e permissões de acesso em um Catálogo de Dados central, definindo políticas de controle de acesso refinadas. Os administradores do data lake podem conceder permissões a outras entidades principais do IAM dentro da conta ou entre contas nos recursos do Catálogo de Dados. As entidades principais do IAM podem consultar os dados compartilhados usando o Amazon Redshift Spectrum e o Amazon Athena.

O Lake Formation fornece os seguintes métodos para compartilhar dados e gerenciar permissões em conjuntos de dados externos e repositórios externos:

- Como integrar o Lake Formation com a unidade de compartilhamento de dados do Amazon Redshift — Use o Lake Formation para gerenciar centralmente as permissões de acesso por banco de dados, tabela, coluna e linha das unidades de compartilhamento de dados do [Amazon Redshift](#) e restringir o acesso dos usuários a objetos em uma unidade de compartilhamento de dados.
- Conexão AWS Glue Data Catalog a metástores externos — Conecte-se a metástores externos AWS Glue Data Catalog para gerenciar permissões de acesso em conjuntos de dados no Amazon S3 usando o Lake Formation. Não é necessária nenhuma migração de metadados para o AWS Glue Data Catalog.
- Como integrar o Lake Formation com o AWS Data Exchange: o Lake Formation oferece suporte ao licenciamento de acesso aos seus dados por meio de AWS Data Exchange. Se você estiver interessado em licenciar seus dados do Lake Formation, consulte [O que está AWS Data Exchange](#) no Guia do Usuário AWS Data Exchange.

## Tópicos

- [Gerenciamento de permissões para dados em uma unidade de compartilhamento de dados do Amazon Redshift.](#)
- [Gerenciamento de permissões em conjuntos de dados que usam repositórios de dados externos](#)

# Gerenciamento de permissões para dados em uma unidade de compartilhamento de dados do Amazon Redshift.

Com AWS Lake Formation, você pode gerenciar dados com segurança em um compartilhamento de dados do Amazon Redshift. O Amazon Redshift é um serviço de data warehouse totalmente gerenciado e em escala de petabytes na nuvem do AWS. Usando o recurso de compartilhamento de dados, o Amazon Redshift ajuda você a compartilhar dados entre Contas da AWS. Para obter mais informações sobre o compartilhamento de dados do Amazon Redshift, consulte [Visão geral do compartilhamento de dados no Amazon Redshift](#).

No Amazon Redshift, o administrador do cluster produtor cria um compartilhamento de dados e o compartilha com o administrador do data lake. Para step-by-step obter instruções sobre como criar um administrador de data lake, consulte [Crie um administrador de data lake](#).

Depois que você (administrador do data lake) aceitar a unidade compartilhamento de dados, deverá criar um banco de dados do AWS Glue Data Catalog para a unidade de compartilhamento de dados específica. Isso serve para que você possa controlar o acesso a ele usando as permissões do Lake Formation. O Lake Formation mapeia cada unidade de compartilhamento de dados para um banco de dados correspondente do Catálogo de Dados. Estes aparecem como bancos de dados federados no Catálogo de Dados.

Um banco de dados é chamado de banco de dados federado quando aponta para uma entidade fora do Catálogo de Dados. As tabelas e exibições na unidade de compartilhamento de dados do Amazon Redshift são listadas como tabelas individuais no Catálogo de Dados. É possível compartilhar o banco de dados federado com diretores IAM e usuários SAML selecionados na mesma conta ou em outra conta com Lake Formation. Você também pode incluir expressões de filtro de linha e coluna para restringir o acesso a determinados dados. Para ter mais informações, consulte [Visão geral da filtragem de dados](#).

Para fornecer aos usuários acesso a um compartilhamento de dados do Amazon Redshift, você deve fazer o seguinte:

1. Atualize as Configurações do Catálogo de Dados para ativar as permissões do Lake Formation.
2. Aceite o convite da unidade de compartilhamento de dados do administrador do cluster produtor do Amazon Redshift e registre a unidade de compartilhamento de dados no Lake Formation.

Depois de concluir essa etapa, você pode gerenciar o compartilhamento de dados no Catálogo de Dados do Lake Formation.

3. Crie um banco de dados federado e defina permissões neste banco de dados.
4. Conceda permissões aos usuários em bancos de dados e tabelas. Você pode compartilhar o banco de dados inteiro ou um subconjunto de tabelas com usuários na mesma conta ou em outra conta.

Para conhecer as limitações, consulte [Limitações do compartilhamento de dados do Amazon Redshift](#).

### Tópicos

- [Requisitos para configurar permissões em unidades de compartilhamento de dados do Amazon Redshift](#)
- [Configuração de permissões para unidades de compartilhamento de dados do Amazon Redshift](#)
- [Consultar bancos de dados federados](#)

## Requisitos para configurar permissões em unidades de compartilhamento de dados do Amazon Redshift

### Atualizar as configurações padrão do Catálogo de Dados

Para habilitar as permissões do Lake Formation para os recursos do Catálogo de Dados, recomendamos que você desative as Configurações-padrão do Catálogo de Dados no Lake Formation. Para ter mais informações, consulte [Altere o modelo de permissão padrão ou use o modo de acesso híbrido](#).

### Permissões atualizadas

Além das permissões de administrador do data lake (AWSLakeFormationDataAdmin), as seguintes permissões também são necessárias para aceitar um compartilhamento de dados do Amazon Redshift no Lake Formation:

- `glue:PassConnection on aws:redshift`
- `redshift:AssociateDataShareConsumer`
- `redshift:DescribeDataSharesForConsumer`
- `redshift:DescribeDataShares`

O usuário do IAM administrador do data lake tem as seguintes permissões implicitamente.

- Acesso à localização de dados
- Criar banco de dados
- Lake Fomation: recurso de registro

## Configuração de permissões para unidades de compartilhamento de dados do Amazon Redshift

Este tópico descreve as etapas que você precisa seguir para aceitar um convite de unidade de compartilhamento de dados, criar um banco de dados federado, e conceder permissões. Você pode usar o console Lake Formation ou o AWS Command Line Interface (AWS CLI). Os exemplos neste tópico mostram o cluster produtor, o Catálogo de Dados, e o consumidor de dados na mesma conta.

Para saber mais sobre os recursos de entre contas do Lake Formation, consulte [Compartilhamento de dados entre contas no Lake Formation](#).

Como configurar permissões para uma unidade de compartilhamento de dados

1. Analise e aceite um convite de unidade de compartilhamento de dados.

### Console

1. Faça login no console do Lake Formation como administrador do data lake em <https://console.aws.amazon.com/lakeformation/>. Navegue até a página Compartilhamento de dados.
2. Analise os compartilhamentos de dados que você está autorizado a acessar. A coluna Status indica seu status atual de participação na unidade de compartilhamento de dados. O status Pendente indica que você foi adicionado a uma unidade de compartilhamento de dados, mas ainda não o aceitou ou rejeitou o convite.
3. Para responder a um convite de compartilhamento de dados, selecione o nome do compartilhamento de dados e escolha Revisar convite. Em Aceitar ou rejeitar o compartilhamento de dados, revise os detalhes do convite. Selecione Aceitar para aceitar o convite ou Rejeitar para recusar o convite. Você não terá acesso ao compartilhamento de dados se rejeitar o convite.

## AWS CLI

Os exemplos a seguir mostram como exibir, aceitar e registrar o convite. Substitua o Conta da AWS ID por um Conta da AWS ID válido. Substitua `data-share-arn` pelo nome de recurso da Amazon (ARN) real que faz referência à unidade de compartilhamento de dados.

### 1. Veja um convite pendente.

```
aws redshift describe-data-shares \  
  --data-share-arn 'arn:aws:redshift:us-  
east-1:111122223333:datashare:abcd1234-1234-ab12-cd34-1a2b3c4d5e6f/  
federatedds' \  
  --output text
```

### 2. Aceitar uma unidade de compartilhamento de dados.

```
aws redshift associate-data-share-consumer \  
  --data-share-arn 'arn:aws:redshift:us-  
east-1:111122223333:datashare:abcd1234-1234-ab12-cd34-1a2b3c4d5e6f/  
federatedds' \  
  --consumer-arn 'arn:aws:glue:us-east-1:111122223333:catalog  
consumer'
```

### 3. Registre a unidade de compartilhamento de dados na conta do Lake Formation. Use a operação [RegisterResource](#) da API para registrar o compartilhamento de dados no Lake Formation. `DataShareArn` é o parâmetro de entrada para `ResourceArn`.

#### Note

Este é uma etapa obrigatória.

```
aws lakeformation register-resource \  
  --resource-arn 'arn:aws:redshift:us-  
east-1:111122223333:datashare:abcd1234-1234-ab12-cd34-1a2b3c4d5e6f/  
federatedds'
```

### 2. Crie um banco de dados.

Depois de aceitar um convite de unidade de compartilhamento de dados, você precisa criar um banco de dados que aponte para o banco de dados do Amazon Redshift associado à unidade

de compartilhamento de dados. Você deve ser administrador de data lake para criar um banco de dados.

## Console

1. Selecione a unidade de compartilhamento de dados no painel Convites e selecione Definir informações do banco de dados.
2. Em Definir informações do banco de dados, insira um nome e um identificador exclusivos para a unidade de compartilhamento de dados. Você usa esse identificador para mapear o compartilhamento de dados internamente na hierarquia de metadados (DBName.schema.table).
3. Selecione Avançar para conceder permissões a outros usuários no banco de dados e nas tabelas compartilhados.

## AWS CLI

Use o código de exemplo a seguir para criar um banco de dados que aponta para o banco de dados do Amazon Redshift compartilhado com o Lake Formation usando o AWS CLI

```
aws glue create-database --cli-input-json \  
  
'{  
  "CatalogId": "111122223333",  
  "DatabaseInput": {  
    "Name": "tahoedb",  
    "FederatedDatabase": {  
      "Identifier": "arn:aws:redshift:us-  
east-1:111122223333:datashare:abcd1234-1234-ab12-cd34-1a2b3c4d5e6f/federatedds",  
      "ConnectionName": "aws:redshift"  
    }  
  }  
}'
```

### 3. Conceder permissões

Depois de criar o banco de dados, você pode conceder permissões aos usuários da sua conta ou a organizações externas Contas da AWS e externas. Você não poderá conceder permissões de gravação de dados (inserir, excluir) e permissões de metadados (alterar, eliminar, criar) no banco de dados federado que está mapeado para um compartilhamento de dados do Amazon

Redshift. Para obter mais informações sobre como conceder permissões, consulte [Gerenciando permissões do Lake Formation](#).

#### Note

Como administrador de data lake, você só pode visualizar tabelas nos bancos de dados federados. Para realizar qualquer outra ação, você precisa conceder a si mesmo mais permissões nessas tabelas.

## Console

1. Na tela Conceder permissões, selecione os usuários aos quais planeja conceder permissões.
2. Selecione Conceder.

## AWS CLI

Use os exemplos a seguir para conceder permissões de banco de dados e tabelas usando o AWS CLI:

```
aws lakeformation grant-permissions --input-cli-json file://input.json

{
  "Principal": {
    "DataLakePrincipalIdentifier": "arn:aws:iam::111122223333:user/non-admin"
  },
  "Resource": {
    "Database": {
      "CatalogId": "111122223333",
      "Name": "tahoedb"
    }
  },
  "Permissions": [
    "DESCRIBE"
  ],
  "PermissionsWithGrantOption": [
  ]
}
```

```
}
```

```
aws lakeformation grant-permissions --input-cli-json file://input.json

{
    "Principal": {
        "DataLakePrincipalIdentifier":
"arn:aws:iam::111122223333:user/non-admin"
    },
    "Resource": {
        "Table": {
            "CatalogId": "111122223333",
            "DatabaseName": "tahoedb",
            "Name": "public.customer"
        }
    },
    "Permissions": [
        "SELECT"
    ],
    "PermissionsWithGrantOption": [
        "SELECT"
    ]
}
```

## Consultar bancos de dados federados

Após conceder permissões, os usuários podem fazer login e começar a consultar o banco de dados federado usando o Amazon Redshift. Agora, os usuários podem usar o nome do banco de dados local para referenciar a unidade de compartilhamento de dados do Amazon Redshift em consultas SQL. No Amazon Redshift, a tabela do cliente no esquema público compartilhada por meio da unidade de compartilhamento de dados terá uma tabela correspondente criada como `public.customer` no Catálogo de Dados.

1. Antes de consultar o banco de dados federado usando o Amazon Redshift, o administrador do cluster cria um banco de dados a partir do banco de dados do Catálogo de Dados usando o seguinte comando:



```
CREATE DATABASE sharedcustomerdb FROM ARN
'arn:aws:glue:<region>:111122223333:database/tahoedb' WITH DATA CATALOG SCHEMA
tahoedb
```

2. O administrador do cluster concede permissões de uso no banco de dados.

```
GRANT USAGE ON DATABASE sharedcustomerdb TO IAM:user;
```

3. Agora você (o usuário federado) pode fazer login nas ferramentas SQL para consultar a tabela.

```
Select * from sharedcustomerdb.public.customer limit 10;
```

Para obter mais informações, consulte [Consulta no AWS Glue Data Catalog](#) no Guia de Gerenciamento do Amazon Redshift.

## Gerenciamento de permissões em conjuntos de dados que usam repositórios de dados externos

Com a federação de metadados do AWS Glue Data Catalog (federação do Catálogo de Dados), você pode conectar o Catálogo de Dados a repositórios externos que armazenam metadados para seus dados do Amazon S3 e gerenciar com segurança as permissões de acesso aos dados usando o AWS Lake Formation. Você não precisa migrar os metadados do repositório externo para o Catálogo de Dados.

O Catálogo de Dados fornece um repositório centralizado de metadados que facilita o gerenciamento e a descoberta de dados em sistemas diferentes. Quando sua organização gerencia dados no Catálogo de Dados, você pode usar AWS Lake Formation para controlar o acesso aos seus conjuntos de dados no Amazon S3.

### Note

Atualmente, oferecemos suporte somente à federação de repositórios do Apache Hive (versão 3 e superior).

Para configurar a federação do Catálogo de Dados, fornecemos um aplicativo AWS Serverless Application Model (AWS SAM) chamado [GlueDataCatalogFederation- HiveMetastore](#) no AWS Serverless Application Repository.

A implementação de referência é fornecida no GitHub como um projeto de código aberto na [AWS Glue Data Catalog Federation - Hive Metastore](#).

O aplicativo AWS SAM cria e implanta os seguintes recursos que são necessários para conectar o Catálogo de Dados ao repositório do Hive:

- Uma AWS Lambda função — Hospeda a implementação do serviço de federação que se comunica entre o Catálogo de Dados e o metastore Hive. AWS Glue invoca essa função Lambda para recuperar objetos de metadados do metastore Hive.
- Amazon API Gateway — O endpoint de conexão do seu repositório do Hive que atua como um proxy para rotear todas as invocações para a função do Lambda.
- Uma função do IAM — uma função com as permissões necessárias para criar a conexão entre o catálogo de dados e o metastore do Hive.
- AWS Glue conexão — Um Amazon API Gateway tipo de AWS Glue conexão que armazena o Amazon API Gateway endpoint e uma função do IAM para invocá-lo.

Quando você consulta tabelas, o serviço AWS Glue faz uma chamada de tempo de execução para o repositório do Hive e busca os metadados. A função do Lambda atua como um tradutor entre o Repositório do Hive e o Catálogo de Dados.

Após estabelecer a conexão, para sincronizar os metadados no repositório do Hive com o Catálogo de Dados, você precisa criar um banco de dados federado no Catálogo de Dados usando as informações da conexão do repositório do Hive, e mapear esse banco de dados para o banco de dados do Hive. Um banco de dados é chamado de banco de dados federado quando aponta para uma entidade fora do Catálogo de Dados.

Você pode aplicar as permissões do Lake Formation usando o controle de acesso baseado em tags e o método de recurso nomeado no banco de dados federado e compartilhá-lo entre várias Contas da AWS unidades organizacionais (OUs). AWS Organizations Você também pode compartilhar o banco de dados federado diretamente com as entidades principais do IAM de outra conta.

Você pode definir permissões refinadas em nível de coluna, nível de linha e nível de célula usando filtros de dados do Lake Formation nas tabelas externas do Hive. Você pode usar o Amazon Athena,

o Amazon Redshift ou o Amazon EMR para consultar as tabelas externas gerenciadas pelo Lake Formation do Hive.

Para obter mais informações sobre compartilhamento de dados entre contas e filtragem de dados, consulte:

- [Compartilhamento de dados entre contas no Lake Formation](#)
- [Filtragem de dados e segurança em nível de célula no Lake Formation](#)

### Etapas de alto nível da federação de metadados do Catálogo de Dados

1. Você cria usuários e funções do IAM que têm as permissões apropriadas para implantar o AWS SAM aplicativo e criar bancos de dados federados.
2. Você registra o local dos dados do Amazon S3 com o Lake Formation selecionando a opção `Enable Data Catalog federation` para conjuntos de dados que usam um repositório externo do Hive.
3. Você define as configurações do aplicativo do AWS SAM (nome da conexão do AWS Glue, URL para o repositório do Hive e parâmetros da função do Lambda) e implanta o aplicativo do AWS SAM.
4. O aplicativo do AWS SAM implanta os recursos necessários para conectar o repositório externo do Hive ao Catálogo de Dados.
5. Para aplicar as permissões do Lake Formation no banco de dados e nas tabelas do Hive, você cria um banco de dados no Catálogo de Dados usando os detalhes da conexão do Hive Metastore e mapeia esse banco de dados para o banco de dados do Hive.
6. Conceda permissões nos bancos de dados federados às entidades principais da sua conta ou de outra conta.

#### Note

Você pode conectar o Data Catalog a um repositório externo do Hive, criar bancos de dados federados e executar consultas e scripts do ETL em bancos de dados e tabelas do Hive sem aplicar as permissões do Lake Formation. Para dados de origem no Amazon S3 que não estão registrados no Lake Formation, o acesso é determinado pelas políticas de permissões do IAM para ações do Amazon S3 e do AWS Glue.

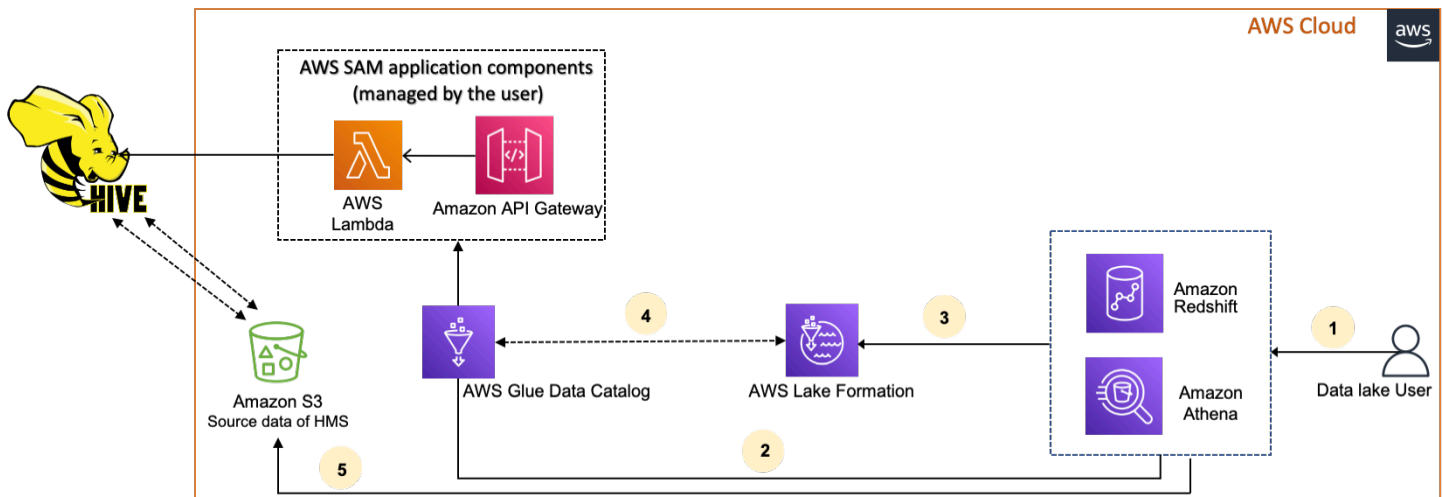
Para conhecer as limitações, consulte [Considerações e limitações do compartilhamento de dados de armazenamento de metadados do Hive](#).

## Tópicos

- [Fluxo de trabalho](#)
- [Requisitos para conectar o Catálogo de Dados ao Repositório do Hive](#)
- [Conectando o Catálogo de Dados a um repositório externo do Hive](#)
- [Recursos adicionais do](#)

## Fluxo de trabalho

O diagrama a seguir mostra o fluxo de trabalho para conectar o AWS Glue Data Catalog a um repositório externo do Hive.



1. Uma entidade principal envia uma consulta usando um serviço integrado, como Athena ou Redshift Spectrum.
2. O serviço integrado faz uma chamada para o Catálogo de Dados para obter os metadados, que por sua vez chama o endpoint do metastore Hive disponível por trás do Amazon API Gateway e recebe respostas às solicitações de metadados.
3. O serviço integrado envia a solicitação ao Lake Formation para verificar as informações e credenciais da tabela para acessar a tabela.
4. O Lake Formation autoriza a solicitação e fornece credenciais temporárias para o aplicativo integrado, que permite o acesso aos dados.

5. Usando as credenciais temporárias recebidas do Lake Formation, o serviço integrado lê os dados do Amazon S3 e compartilha os resultados com a entidade principal.

## Requisitos para conectar o Catálogo de Dados ao Repositório do Hive

Para conectar o AWS Glue Data Catalog a um repositório externo do Apache Hive e configurar as permissões de acesso aos dados, você precisa preencher os seguintes requisitos:

### Note

Recomendamos que um administrador do Lake Formation implante o aplicativo do AWS SAM e que somente um usuário privilegiado use a conexão de repositório do Hive para criar os bancos de dados federados correspondentes.

1. Crie perfis do IAM.

Como implantar o aplicativo AWS SAM

- Crie um perfil que tenha as permissões necessárias para implantar os recursos (função do Lambda, Amazon API Gateway, perfil do IAM e conexão AWS Glue) necessários para criar uma conexão com a repositório do Hive.

Como criar bancos de dados federados

As seguintes permissões são necessárias nos recursos:

- `glue:CreateDatabase` on resource `arn:aws:glue:region:account-id:database/gluedatabasename`
- `glue:PassConnection` on resource `arn:aws:glue:region:account-id:connection/hms_connection`

2. Registre o local do Amazon S3 com o Lake Formation.

Para usar o Lake Formation para gerenciar e proteger os dados em seu data lake, você deve registrar o local do Amazon S3 que tem os dados das tabelas na repositório do Hive com o Lake Formation. Ao fazer isso, a Lake Formation pode fornecer credenciais para serviços AWS analíticos como Athena, Redshift Spectrum e Amazon EMR.

Para obter mais informações sobre o registro de um local do Amazon S3, consulte [Adicionar uma localização do Amazon S3 ao seu data lake](#).

Ao registrar a localização do Amazon S3, marque a caixa de seleção Enable Data Catalog Federation para permitir que o Lake Formation assuma a função de acessar tabelas em um banco de dados federado.

AWS Lake Formation > Data lake locations > Register location

## Register location


### Amazon S3 location

Register an Amazon S3 path as the storage location for your data lake.

**Amazon S3 path**  
Choose an Amazon S3 path for your data lake.

**Review location permissions - strongly recommended**  
Registering the selected location may result in your users gaining access to data already at that location. Before registering a location, we recommend that you review existing location permissions on resources in that location.

**IAM role**  
To add or update data, Lake Formation needs read/write access to the chosen Amazon S3 path. Choose a role that you know has permission to do this, or choose the **AWSServiceRoleForLakeFormationDataAccess** service-linked role. When you register the first Amazon S3 path, the service-linked role and a new inline policy are created on your behalf. Lake Formation adds the first path to the inline policy and attaches it to the service-linked role. When you register subsequent paths, Lake Formation adds the path to the existing policy.

 Do not select the service linked role if you plan to use EMR.

**Enable Data Catalog Federation**  
Checking this box will allow Lake Formation to assume a role to access tables in a federated database.

Para obter mais informações sobre como registrar um local de dados no Lake Formation, consulte [Como configurar um local no Amazon S3 para o data lake](#).

### 3. Use a versão correta do Amazon EMR.

Para usar o Amazon EMR com os bancos de dados federados do metastore Hive, você precisa ter o Hive versão 3.x ou superior e o Amazon EMR versão 6.x ou superior.

## Conectando o Catálogo de Dados a um repositório externo do Hive

[Para conectá-lo AWS Glue Data Catalog a uma metastore do Hive, você precisa implantar um AWS SAM aplicativo chamado - GlueDataCatalogFederation HiveMetastore](#) Ele cria os recursos necessários para conectar o repositório externo do Hive ao Catálogo de Dados. Você pode acessar o aplicativo do AWS SAM no AWS Serverless Application Repository.

O aplicativo do AWS SAM cria a conexão para o Repositório do Hive por trás do Amazon API Gateway usando uma função do Lambda. O aplicativo do AWS SAM usa um identificador uniforme de recursos (URI) como entrada do usuário e conecta o repositório externo do Hive ao Catálogo de Dados. Quando um usuário executa uma consulta nas tabelas do Hive, o Catálogo de Dados chama o endpoint do API Gateway. O endpoint invoca a função do Lambda para recuperar os metadados das tabelas do Hive.

Como conectar o Catálogo de Dados ao repositório do Hive e configurar permissões

### 1. Implante a aplicação AWS SAM.

1. Faça login no AWS Management Console e abra a AWS Serverless Application Repository.
2. No painel de navegação, escolha Aplicativos disponíveis.
3. Selecione Aplicativos públicos.
4. Selecione a opção Show apps that create custom IAM roles or resource policies (Mostrar aplicações que criam funções personalizadas do IAM ou políticas de recursos).
5. Na caixa de pesquisa, digite o nome GlueDataCatalogFederation- HiveMetastore.
6. Escolha o HiveMetastore aplicativo GlueDataCatalogFederation-.
7. Em Configurações do aplicativo, digite as seguintes configurações mínimas necessárias para sua função do Lambda:
  - Nome do aplicativo — Um nome para seu aplicativo AWS SAM.
  - GlueConnectionName- Um nome para a conexão.
  - HiveMetastoreURIs - O URI do seu host de metastore Hive.

- LambdaMemory- A quantidade de memória Lambda em MB de 128 a 10240. O padrão é 1024.
  - LambdaTimeout- O tempo máximo de execução de invocação do Lambda em segundos. O padrão é 30.
  - VPC e SecurityGroupIds VPC SubnetIds - Informações para a VPC onde existe a metastore Hive.
8. Selecione Reconheço que este aplicativo cria perfis personalizadas do IAM e políticas de recursos. Para obter mais informações, escolha o link Informações.
  9. Na parte inferior direita da página Configurações da aplicação selecione Implantar. Quando a implantação for concluída, a função do Lambda será exibida seção Recursos no console do Lambda.

O aplicativo é implantado no Lambda. Seu nome é prefixado com serverlessrepo- para indicar que o aplicativo foi implantado a partir do. AWS Serverless Application Repository Selecionar o aplicativo leva você à página Recursos, na qual cada um dos recursos do aplicativo que foram implantados está listado. Os recursos incluem a função do Lambda, que permite a comunicação entre o Catálogo de Dados e o Repositório do Hive, a conexão do AWS Glue e outros recursos necessários para a federação do banco de dados.

2. Crie um banco de dados federado no Catálogo de Dados.

Depois de criar uma conexão com o metastore do Hive, você pode criar bancos de dados federados no Catálogo de Dados que apontam para os bancos de dados externos do metastore do Hive. Você precisa criar um banco de dados correspondente no Catálogo de Dados para cada banco de dados de metastore do Hive que você está conectando ao Catálogo de Dados.

### Lake Formation console

1. Na página Compartilhamento de dados, selecione a guia Bancos de dados compartilhados e, em seguida, selecione Criar banco de dados.
2. Em Nome da conexão, escolha o nome da sua conexão de metastore do Hive no menu suspenso.
3. Digite um nome de banco de dados exclusivo e o identificador de origem da federação para o banco de dados. Esse é o nome que você usa em suas instruções SQL ao consultar tabelas. O nome pode ter no máximo 255 caracteres e deve ser exclusivo em sua conta.



#### 4. Selecione Criar banco de dados.

##### AWS CLI

```
aws glue create-database \  
'{  
  "CatalogId": "<111122223333>",  
  "database-input": {  
    "Name": "<fed_glue_db>",  
    "FederatedDatabase": {  
      "Identifier": "<hive_db_on_emr>",  
      "ConnectionName": "<hms_connection>"  
    }  
  }  
'
```

#### 3. Visualize tabelas no banco de dados federado.

Após criar o banco de dados federado, você pode exibir a lista de tabelas em seu repositório do Hive usando o console Lake Formation ou o AWS CLI.

##### Lake Formation console

1. Selecione o nome do banco de dados na guia Bancos de dados compartilhados.
2. Na página Bancos de dados, selecione Exibir tabelas.

##### AWS CLI

Os exemplos a seguir mostram como recuperar a definição da conexão, o nome do banco de dados e algumas ou todas as tabelas no banco de dados. Substitua a ID do Catálogo de Dados pela Conta da AWS ID válida que você usou para criar o banco de dados. Substitua `hms_connection` pelo nome da conexão.

```
aws glue get-connection \  
--name <hms_connection> \  
--catalog-id 111122223333
```

```
aws glue get-database \  
--catalog-id 111122223333
```

```
--name <fed_glu_db> \  
--catalog-id 111122223333
```

```
aws glue get-tables \  
--database-name <fed_glue_db> \  
--catalog-id 111122223333
```

```
aws glue get-table \  
--database-name <fed_glue_db> \  
--name <hive_table_name> \  
--catalog-id 111122223333
```

#### 4. Conceder permissões

Depois de criar o banco de dados, você pode conceder permissões a outros usuários e funções do IAM em sua conta ou a organizações externas Contas da AWS e externas. Você não poderá conceder permissões de gravação de dados (inserir, excluir) e permissões de metadados (alterar, eliminar, criar) nos bancos de dados federados. Para obter mais informações sobre como conceder permissões, consulte [Gerenciando permissões do Lake Formation](#).

#### 5. Consulte os bancos de dados federados.

Após conceder permissões, os usuários podem fazer login e começar a consultar o banco de dados federado usando o Athena e o Amazon Redshift. Agora, os usuários podem usar o nome do banco de dados local para referenciar o banco de dados do Hive em consultas SQL.

Exemplo de sintaxe de Amazon Athena consulta

fed\_glue\_dbSubstitua pelo nome do banco de dados local que você criou anteriormente.

```
Select * from fed_glue_db.customers limit 10;
```

## Recursos adicionais do

A postagem do blog a seguir contém instruções detalhadas para configurar as permissões do Lake Formation em um banco de dados e tabelas de repositório do Hive e consultá-los usando o Athena. Também ilustramos um caso de uso de compartilhamento entre contas, em que um diretor de Lake Formation na conta de produtor A compartilha um banco de dados federado do Hive e tabelas usando a tag LF na conta do consumidor B.

- [Consulte seu repositório do Apache Hive com permissões do AWS Lake Formation](#)

# Segurança em AWS Lake Formation

A segurança para com a nuvem na AWS é a nossa maior prioridade. Como cliente da AWS, você se contará com um datacenter e uma arquitetura de rede criados para atender aos requisitos das organizações com as maiores exigências de segurança.

A segurança é uma responsabilidade compartilhada entre a AWS e você. O [modelo de responsabilidade compartilhada](#) descreve isto como segurança da nuvem e segurança na nuvem:

- **Segurança da nuvem:** a AWS é responsável pela proteção da infraestrutura que executa produtos da AWS na Nuvem AWS. A AWS também fornece serviços que podem ser usados com segurança. Auditores de terceiros testam e verificam regularmente a eficácia da nossa segurança como parte dos [programas de conformidade da AWS](#). Para saber mais sobre os programas de conformidade que se aplicam ao AWS Lake Formation, consulte [Serviços da AWS no escopo pelo programa de conformidade](#).
- **Segurança na nuvem:** sua responsabilidade é determinada pelo serviço da AWS que você usa. Você também é responsável por outros fatores, incluindo a confidencialidade de seus dados, os requisitos da sua empresa e as leis e regulamentos aplicáveis.

Esta documentação ajuda você a entender como aplicar o modelo de responsabilidade compartilhada ao usar o Lake Formation. Os tópicos a seguir mostram como configurar o Lake Formation para atender aos seus objetivos de segurança e conformidade. Você também aprenderá como usar outros serviços AWS que ajudam a monitorar e proteger seus recursos do Lake Formation.

## Tópicos

- [Proteção de dados no Lake Formation](#)
- [Segurança da infraestrutura no AWS Lake Formation](#)
- [Prevenção contra o ataque do “substituto confuso” em todos os serviços](#)
- [Registro em log de eventos de segurança no AWS Lake Formation](#)

## Proteção de dados no Lake Formation

O [modelo de responsabilidade compartilhada AWS](#) se aplica à proteção de dados no Lake Formation AWS. Conforme descrito nesse modelo, a AWS é responsável por proteger a infraestrutura global

que executa toda a Nuvem AWS. Você é responsável por manter o controle sobre seu conteúdo hospedado nessa infraestrutura. Você também é responsável pelas tarefas de configuração e gerenciamento de segurança dos Serviços da AWS que usa. Para obter mais informações sobre a privacidade de dados, consulte as [Perguntas frequentes sobre privacidade de dados](#). Para mais informações sobre a proteção de dados na Europa, consulte o artigo [AWS Shared Responsibility Model and GDPR](#) no Blog de segurança da AWS.

Para fins de proteção de dados, recomendamos que você proteja as Conta da AWS credenciais da e configure as contas de usuário individuais com o AWS IAM Identity Center ou o AWS Identity and Access Management (IAM). Dessa maneira, cada usuário receberá apenas as permissões necessárias para cumprir suas obrigações de trabalho. Recomendamos também que você proteja seus dados das seguintes formas:

- Use uma autenticação multifator (MFA [multi-factor authentication]) com cada conta.
- Use SSL/TLS para se comunicar com os atributos da AWS. Exigimos TLS 1.2 e recomendamos TLS 1.3.
- Configure o registro em log das atividades da API e do usuário com o .AWS CloudTrail
- Use AWS as soluções de criptografia da , juntamente com todos os controles de segurança padrão dos Serviços da AWS.
- Use serviços gerenciados de segurança avançada, como o Amazon Macie, que ajuda a localizar e proteger dados sigilosos armazenados no Amazon S3.
- Se você precisar de módulos criptográficos validados pelo FIPS 140-2 ao acessar a AWS por meio de uma interface de linha de comandos ou uma API, use um endpoint do FIPS. Para ter mais informações sobre endpoints do FIPS, consulte [Federal Information Processing Standard \(FIPS\) 140-2](#).

É altamente recomendável que nunca sejam colocadas informações de identificação confidenciais, como endereços de e-mail dos seus clientes, em marcações ou campos de formato livre, como um campo Name (Nome). Isso inclui quando você trabalha com Lake Formation ou outro Serviços da AWS usando o console, API, AWS CLI ou SDKs AWS. Quaisquer dados inseridos em tags ou campos de texto de formato livre usados para nomes podem ser usados para logs de faturamento ou de diagnóstico. Se você fornecer um URL para um servidor externo, recomendamos fortemente que não sejam incluídas informações de credenciais no URL para validar a solicitação a esse servidor.

## Criptografia em repouso

AWS Lake Formation oferece suporte à criptografia de dados nas seguintes áreas:

- Dados no data lake do Amazon Simple Storage Service (Amazon S3).

O Lake Formation oferece suporte à criptografia de dados com [AWS Key Management Service](#) (AWS KMS). Os dados geralmente são gravados no data lake por meio de trabalhos de extração, transformação e carregamento (ETL) de AWS Glue. Para obter informações sobre como criptografar dados gravados por trabalhos AWS Glue, consulte [Criptografar dados gravados por crawlers, trabalhos e endpoints de desenvolvimento](#) no Guia do desenvolvedor AWS Glue

- O AWS Glue Data Catalog, que é onde o Lake Formation armazena tabelas de metadados que descrevem os dados no data lake.

Para obter mais informações, consulte [Criptografando seu Catálogo de dados](#) no AWS Glue Guia do desenvolvedor .

Para adicionar um local do Amazon S3 como armazenamento em seu data lake, você registra o local com AWS Lake Formation. Em seguida, você pode usar as permissões do Lake Formation para um controle de acesso refinado aos objetos AWS Glue Data Catalog que apontam para esse local e aos dados subjacentes no local.

O Lake Formation suporta o registro de uma localização do Amazon S3 que contém dados criptografados. Para obter mais informações, consulte [Registrando uma localização criptografada do Amazon S3](#).

## Segurança da infraestrutura no AWS Lake Formation

Como um serviço gerenciado, o AWS Lake Formation é protegido pelos procedimentos de segurança de rede global da AWS que estão descritos no whitepaper [Amazon Web Services: Overview of Security Processes](#).

Você usa chamadas de API publicadas AWS para acessar o Lake Formation por meio da rede. Os clientes devem oferecer suporte a Transport Layer Security (TLS) 1.0 ou posterior. Recomendamos TLS 1.2 ou posterior. Os clientes também devem ter suporte a conjuntos de criptografia com perfect forward secrecy (PFS) como Ephemeral Diffie-Hellman (DHE) ou Ephemeral Elliptic Curve Diffie-Hellman (ECDHE). A maioria dos sistemas modernos como Java 7 e versões posteriores oferece suporte a esses modos.

Além disso, as solicitações devem ser assinadas usando um ID da chave de acesso e uma chave de acesso secreta associada a uma entidade principal do IAM. Ou você pode usar o [AWS](#)

[Security Token Service](#) (AWS STS) para gerar credenciais de segurança temporárias para assinar solicitações.

## Prevenção contra o ataque do “substituto confuso” em todos os serviços

"Substituto confuso" é um problema de segurança em que uma entidade que não tem permissão para executar uma ação pode coagir uma entidade mais privilegiada a executá-la. Em AWS, a personificação entre serviços pode resultar no problema do "substituto confuso". A imitação entre serviços pode ocorrer quando um serviço (o serviço de chamada) chama outro serviço (o serviço chamado). O serviço de chamada pode ser manipulado de modo a usar suas permissões para atuar nos recursos de outro cliente de uma forma na qual ele não deveria ter permissão para acessar. Para evitar isso, o AWS fornece ferramentas que ajudam você a proteger seus dados para todos os serviços com entidades principais de serviço que receberam acesso aos recursos em sua conta.

Recomendamos o uso das chaves de contexto de condição global [aws:SourceArn](#) e [aws:SourceAccount](#) nas políticas de recursos para limitar as permissões que AWS Lake Formation concede a outro serviço para o recurso. Se você usar as duas chaves de contexto de condição global, o valor `aws:SourceAccount` e a conta no valor `aws:SourceArn` deverão usar o mesmo ID de conta quando usados na mesma declaração de política.

Atualmente, o Lake Formation só é compatível `aws:SourceArn` com o seguinte formato:

```
arn:aws:lakeformation:aws-region:account-id:*
```

O exemplo a seguir mostra como você pode usar as teclas de contexto de condição global `aws:SourceArn` e `aws:SourceAccount` no Lake Formation para evitar o problema do "substituto confuso".

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "ConfusedDeputyPreventionExamplePolicy",
      "Effect": "Allow",
      "Principal": {
        "Service": "lakeformation.amazonaws.com"
      },
      "Action": [
```

```
    "sts:AssumeRole"
  ],
  "Condition": {
    "StringEquals": {
      "aws:SourceAccount": "account-id"
    },
    "ArnEquals": {
      "aws:SourceArn": "arn:aws:lakeformation:aws-region:account-id:*"
    }
  }
}
```

## Registro em log de eventos de segurança no AWS Lake Formation

O Lake Formation AWS está integrado com AWS CloudTrail, um serviço que fornece um registro das ações realizadas por um usuário, função ou serviço AWS no Lake Formation. O CloudTrail captura todas as chamadas de API do Lake Formation como eventos. As chamadas capturadas incluem chamadas do console do Lake Formation, do AWS Command Line Interface e de chamadas de código para as operações da API do Lake Formation.

Para obter mais informações sobre o log de eventos no Lake Formation, consulte [Como fazer log de chamados de API do Lake Formation do AWS usando AWS CloudTrail](#).

### Note

GetTableObjects, UpdateTableObjects e GetWorkUnitResults são operações de plano de dados de alto volume. No momento, as chamadas para essas APIs não estão registradas no CloudTrail. Para obter mais informações sobre operações de plano de dados no CloudTrail, consulte [Registrando eventos de dados para trilhas](#) no Guia do usuário AWS CloudTrail

As mudanças no Lake Formation para dar suporte a eventos adicionais do CloudTrail serão documentadas em [Histórico da documentação do AWS Lake Formation](#).



# Integração de serviços de terceiros com o Lake Formation

A integração com o AWS Lake Formation permite que serviços de terceiros acessem com segurança dados em seus data lakes baseados no Amazon S3. Você pode usar o Lake Formation como seu mecanismo de autorização para gerenciar ou aplicar permissões ao seu data lake com serviços da integrados da AWS, como Amazon Athena, Amazon EMR e Redshift Spectrum. O Lake Formation oferece duas opções para integrar serviços:

1. As configurações de integração do aplicativo Lake Formation: O Lake Formation pode fornecer credenciais temporárias com escopo reduzido na forma de tokens STS da AWS para locais registrados do Amazon S3 com base nas permissões efetivas, para que aplicativos autorizados possam acessar dados em nome dos usuários.
2. Aplicação central: as operações de [API de consulta](#) do Lake Formation recuperam dados do Amazon S3 e filtram os resultados com base nas permissões efetivas. O mecanismo ou aplicativo que se integra à operação da API de consulta pode depender do Lake Formation para avaliar as permissões da identidade de chamada e filtrar com segurança os dados com base nessas permissões. Mecanismos de consulta de terceiros só veem e operam com dados filtrados.

## Tópicos

- [Como usar a integração de aplicativos Lake Formation](#)

## Como usar a integração de aplicativos Lake Formation

O Lake Formation permite que serviços terceirizados se integrem ao Lake Formation e obtenham acesso temporário aos dados do Amazon S3 em nome de seus usuários por meio do uso [GetTemporaryGlueTableCredentials](#) [GetTemporaryGluePartitionCredentials](#) das operações. Isso permite que serviços de terceiros usem o mesmo atributo de autorização e fornecimento de credenciais usado pelos demais serviços analíticos da AWS. Esta seção descreve como usar essas operações de API para integrar um mecanismo de consulta de terceiros com o Lake Formation.

Por padrão, essas operações de API estão desativadas. Há duas opções para autorizar a integração de aplicativos com o Lake Formation:

- Configure tags de sessão do IAM que são validadas sempre que as operações da API de integração de aplicativos são chamadas

Para ter mais informações, consulte [Como habilitar permissões para que um mecanismo de consulta de terceiros chame operações de API de integração de aplicativos](#).

- Ative a opção que permite que mecanismos externos acessem dados em locais do Amazon S3 com acesso total à tabela.

Essa opção permite que mecanismos de consulta e aplicativos obtenham credenciais sem tags de sessão do IAM se o usuário tiver acesso total à tabela. Ele fornece benefícios de desempenho para mecanismos de consulta e aplicativos, além de simplificar o acesso aos dados. O Amazon EMR no Amazon EC2 é capaz de utilizar essa configuração.

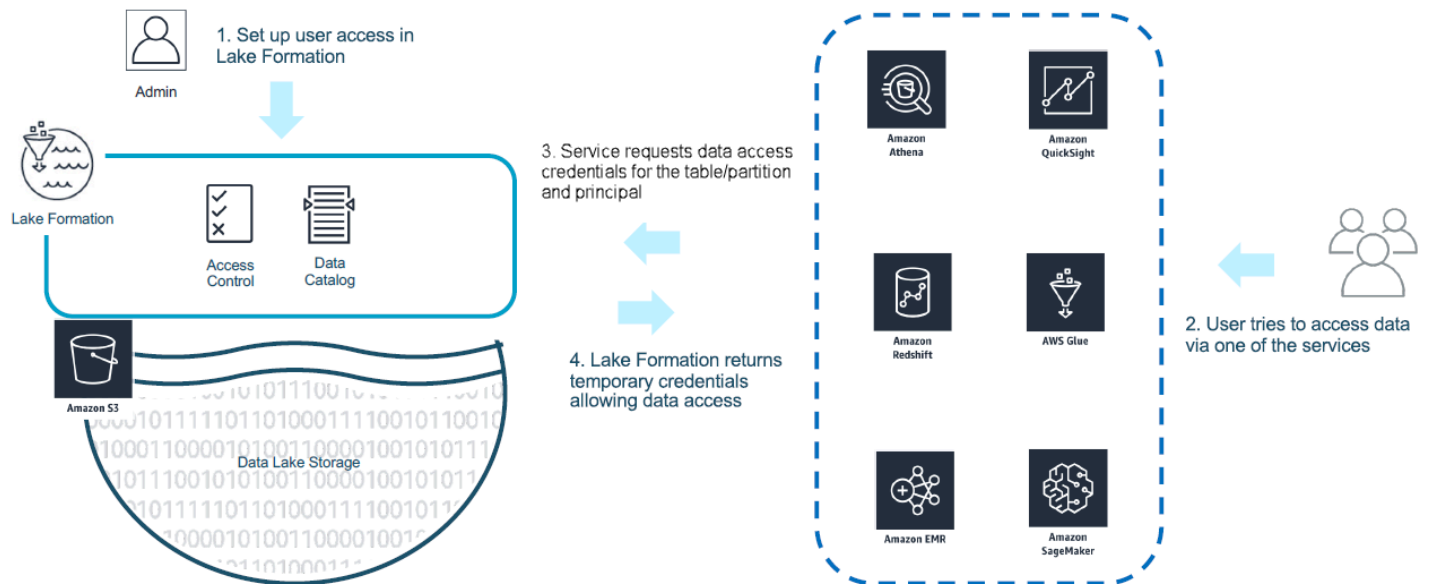
Para ter mais informações, consulte [Integração de aplicativos para acesso total à tabela](#).

## Tópicos

- [Como funciona a integração de aplicações do Lake Formation](#)
- [Perfis e responsabilidades na integração de aplicativos do Lake Formation](#)
- [Fluxo de trabalho do Lake Formation para operações de API de integração de aplicativos](#)
- [Como registrar um mecanismo de consulta de terceiros](#)
- [Como habilitar permissões para que um mecanismo de consulta de terceiros chame operações de API de integração de aplicativos](#)
- [Integração de aplicativos para acesso total à tabela](#)

## Como funciona a integração de aplicações do Lake Formation

Esta seção descreve como usar as operações da API de integração de aplicativos para integrar um aplicativo de terceiros (mecanismo de consulta) com o Lake Formation.



## 1. O administrador do Lake Formation executa as seguintes atividades:

- Registra um local do Amazon S3 no Lake Formation fornecendo um perfil do IAM (usado para fornecimento de credenciais) que tem permissões apropriadas para acessar dados dentro do local do Amazon S3
- Registra um aplicativo de terceiros para poder chamar as operações de API de fornecimento de credenciais do Lake Formation. Consulte [the section called “Como registrar um mecanismo de consulta de terceiros”](#)
- Concede permissões para que os usuários acessem bancos de dados e tabelas

Por exemplo, se você quiser publicar um conjunto de dados de sessões de usuário que inclua algumas colunas contendo informações de identificação pessoal (PII), para restringir o acesso, atribua a essas colunas uma [LF-TBAC](#) chamada “classificação” com o valor “sensível”. Em seguida, defina uma permissão para um analista de negócios acessar os dados das sessões do usuário, mas exclui as colunas marcadas com classificação = sensível.

2. Uma entidade principal (usuário) envia uma consulta para um serviço integrado.
3. O aplicativo integrado solicita ao Lake Formation informações e credenciais da tabela para acessar a tabela.
4. Se o responsável pela consulta estiver autorizado a acessar a tabela, o Lake Formation retornará as credenciais para o aplicativo integrado, que permite o acesso aos dados.

**Note**

O Lake Formation não acessa os dados subjacentes ao vender credenciais.

5. O serviço integrado lê dados do Amazon S3, filtra as colunas com base nas políticas recebidas e retorna os resultados à entidade principal.

**Important**

As operações de API do Lake Formation de fornecimento de credenciais permitem um modelo de fiscalização distribuída com negação explícita em caso de falha (fail-close). Isso introduz um modelo de segurança tripartido entre clientes, serviços terceirizados e Lake Formation. Os serviços integrados são confiáveis para aplicar adequadamente as permissões do Lake Formation (fiscalização distribuída).

O serviço integrado é responsável por filtrar os dados lidos do Amazon S3 com base nas políticas retornadas do Lake Formation antes que os dados filtrados sejam devolvidos ao usuário. Os serviços integrados seguem um modelo de "fail-close", o que significa que eles devem falhar na consulta se não conseguirem aplicar as permissões do Lake Formation necessárias.

## Perfis e responsabilidades na integração de aplicativos do Lake Formation

Função	Responsabilidade
O consumidor	<ul style="list-style-type: none"><li>• Ative a configuração de integração do aplicativo Lake Formation (consulte <a href="#">the section called “Como registrar um mecanismo de consulta de terceiros”</a>).</li><li>• Registra explicitamente terceiros aprovados no Lake Formation (consulte <a href="#">the section called “Como registrar um mecanismo de consulta de terceiros”</a>).</li><li>• Testa e valida soluções de terceiros com as permissões do Lake Formation.</li><li>• Monitora e audita o uso de terceiros das operações da API de fornecimento automática de credenciais do Lake Formation.</li></ul>

Função	Responsabilidade
Terceiros	<ul style="list-style-type: none"> <li>• Documenta publicamente o recurso suportado para cada revisão de software e fornece instruções para ativá-lo corretamente.</li> <li>• Anuncia com precisão os recursos suportados ao chamar as operações da API de fornecimento de credenciais do Lake Formation (de acordo com a documentação).</li> <li>• Armazena e gerencia com segurança as credenciais fornecidas para evitar vazamentos de credenciais e aumento de privilégios.</li> <li>• Impõe permissões com base nos recursos suportados e retorna somente dados filtrados aos usuários</li> <li>• Falha na consulta quando não é possível aplicar adequadamente as permissões necessárias</li> </ul>
AWS Lake Formation	<ul style="list-style-type: none"> <li>• Deriva e retorna corretamente as permissões efetivas para uma determinada entidade principal.</li> <li>• Valida os recursos suportados por terceiros com call-by-call base na operação da API.</li> <li>• Retorna credenciais do IAM com escopo reduzido somente quando os recursos anunciados do mecanismo correspondem aos definidos nos recursos do catálogo, caso contrário, retorna um erro.</li> </ul>

## Fluxo de trabalho do Lake Formation para operações de API de integração de aplicativos

A seguir está o fluxo de trabalho das operações da API de integração de aplicativos:

1. Um usuário envia uma consulta ou solicitação de dados usando um mecanismo de consulta integrado de terceiros. O mecanismo de consulta assume um perfil do IAM que representa o usuário ou um grupo de usuários e recupera credenciais confiáveis para serem usadas ao chamar as operações da API de integração de aplicativos.
2. O mecanismo de consulta chama `GetUnfilteredTableMetadata` e, se for uma tabela particionada, o mecanismo de consulta chama `GetUnfilteredPartitionsMetadata` para recuperar metadados e informações de política do Catálogo de Dados.

3. O Lake Formation realiza a autorização para a solicitação. Se o usuário não tiver as permissões apropriadas na mesa, ele será `AccessDeniedException` descartado.
4. Como parte da solicitação, o mecanismo de consulta envia a filtragem compatível. Há dois sinalizadores que podem ser enviados em uma matriz: `COLUMN_PERMISSIONS` e `CELL_FILTER_PERMISSION`. Se o mecanismo de consulta não oferecer suporte a nenhum desses recursos e existir uma política na tabela para o recurso, a será lançado e a consulta falhará. `PermissionTypeMismatchException` Isso acontece para evitar o vazamento de dados.
5. A resposta obtida contém o seguinte:
  - O esquema inteiro da tabela para que os mecanismos de consulta possam usá-lo para analisar os dados do armazenamento.
  - Uma lista de colunas autorizadas que o usuário tem acesso. Se a lista de colunas autorizadas estiver vazia, isso indica que o usuário tem permissões `DESCRIBE`, mas não tem permissões `SELECT`, e a consulta falha.
  - Um alerta, `IsRegisteredWithLakeFormation`, que indica se o Lake Formation pode fornecer credenciais para esses dados de recursos. Se isso retornar falso, as credenciais dos clientes devem ser usadas para acessar o Amazon S3.
  - Uma lista de `CellFilters`, se houver, que deve ser aplicada às linhas de dados. Essa lista contém colunas e uma expressão para avaliar cada linha. Isso só deve ser preenchido se `CELL_FILTER_PERMISSION` for enviado como parte da solicitação e houver um filtro de dados na tabela do usuário chamador.
6. Depois que os metadados são recuperados, o mecanismo de consulta chama `GetTemporaryGlueTableCredentials` ou `GetTemporaryGluePartitionCredentials` para obter credenciais da AWS para recuperar dados do local do Amazon S3.
7. O mecanismo de consulta lê objetos relevantes do Amazon S3, filtra os dados com base nas políticas recebidas na etapa 2 e retorna os resultados ao usuário.

As operações da API de integração de aplicativos para o Lake Formation contêm conteúdo adicional para configurar a integração com mecanismos de consulta de terceiros. Você pode ver os informações da operação na seção [Operações da API de fornecimento de credenciais](#).

## Como registrar um mecanismo de consulta de terceiros

Antes que um mecanismo de consulta de terceiros possa usar as operações da API de integração de aplicativos, você precisa habilitar explicitamente as permissões para que o mecanismo de consulta chame as operações da API em seu nome. Isso é feito em algumas etapas:

1. Você precisa especificar as contas da AWS e as tags de sessão do IAM que exigem permissão para chamar as operações da API de integração de aplicativos por meio do console do Lake Formation AWS ou AWS CLI da API/SDK.
2. Quando o mecanismo de consulta de terceiros assume o perfil de execução em sua conta, o mecanismo de consulta deve anexar uma tag de sessão registrada no Lake Formation representando o mecanismo de terceiros. O Lake Formation usa essa tag para validar se a solicitação for proveniente de um mecanismo aprovado. Para obter mais informações sobre tags de sessão, consulte [Tags de sessão](#) no Guia do usuário do IAM.
3. Ao configurar um perfil de execução de mecanismo de consulta de terceiros, você deve ter o seguinte conjunto mínimo de permissões na política do IAM:

```
{
  "Version": "2012-10-17",
  "Statement": [{"Effect": "Allow",
    "Action": [
      "lakeformation:GetDataAccess",
      "glue:GetTable",
      "glue:GetTables",
      "glue:GetDatabase",
      "glue:GetDatabases",
      "glue>CreateDatabase",
      "glue:GetUserDefinedFunction",
      "glue:GetUserDefinedFunctions",
      "glue:GetPartition",
      "glue:GetPartitions"
    ],
    "Resource": "*"
  }
}
```

4. Configure uma política de confiança de perfil no perfil de execução do mecanismo de consulta para ter um controle de acesso preciso sobre qual par de chave-valor de tag de sessão pode ser anexado a esse perfil. No exemplo a seguir, esse perfil só pode ter a chave "LakeFormationAuthorizedCaller" da tag da sessão e o valor "engine1" da tag da sessão a serem anexados, e nenhum outro par de chave e valor da tag da sessão é permitido.

```
{
  "Sid": "AllowPassSessionTags",
  "Effect": "Allow",
  "Principal": {
```

```
    "AWS": "arn:aws:iam::111122223333:role/query-execution-role"
  },
  "Action": "sts:TagSession",
  "Condition": {
    "StringLike": {
      "aws:RequestTag/LakeFormationAuthorizedCaller": "engine1"    }
    }
  }
}
```

Quando `LakeFormationAuthorizedCaller` chama a operação STS: AssumeRole API para buscar credenciais para o mecanismo de consulta usar, a tag da sessão deve ser incluída na `AssumeRole` solicitação. A credencial temporária retornada pode ser usada para fazer solicitações de API de integração de aplicativos do Lake Formation.

As operações da API de integração de aplicativos do Lake Formation exigem que a entidade principal da chamada tenha um perfil do IAM. O perfil do IAM deve incluir uma tag de sessão com um valor predeterminado que tenha sido registrado com o Lake Formation. Essa tag permite ao Lake Formation verificar se o perfil usada para chamar as operações da API de integração de aplicativos tem permissão para fazer isso.

## Como habilitar permissões para que um mecanismo de consulta de terceiros chame operações de API de integração de aplicativos

Siga estas etapas para permitir que um mecanismo de consulta de terceiros chame operações de API de integração de aplicativos por meio do console do Lake Formation da AWS, do AWS CLI, ou da API/SDK.

### Console

Como registrar sua conta para filtragem externa de dados:

1. Faça login no AWS Management Console e abra o console do Lake Formation em <https://console.aws.amazon.com/lakeformation/>.
2. Na navegação à esquerda, expanda Permissões e selecione Configuração de integração de aplicativos.
3. Na página de configuração de Integração do aplicativo, selecione a opção Permitir que mecanismos externos filtrem dados em locais do Amazon S3 registrados com o Lake Formation.



4. Digite as tags de sessão que você criou para o mecanismo de terceiros. Para obter informações sobre tags de sessão, consulte [Passagem de tags de sessão no STS da AWS](#) no Guia do Usuário do AWS Identity and Access Management,
5. Digite os IDs de conta dos usuários que podem usar o mecanismo de terceiros para acessar informações de metadados não filtradas e as credenciais de acesso aos dados dos recursos na conta atual.

Você também pode usar o campo ID da conta da AWS para configurar o acesso entre contas.

## Application integration settings [Learn more](#)

### Application integration settings

Use the options below to control which third-party engines are allowed to read and filter data in Amazon S3 locations registered with Lake Formation.

**Allow external engines to filter data in Amazon S3 locations registered with Lake Formation**  
 Check this box to allow third-party engines to access data in Amazon S3 locations that are registered with Lake Formation.

**Session tag values**  
 Enter one or more strings that match the LakeFormationAuthorizedCaller session tag defined for third-party engines.

engine 1 ✕

engine 2 ✕

session 1 ✕

Enter one or several string values separated by comma.

**AWS account IDs**  
 Enter the external AWS account IDs from where third-party engines are allowed to access locations registered with Lake Formation.

111111111111 ✕  
Account

222222222222 ✕  
Account

Enter one or more AWS account IDs. Press enter after each ID.

**Allow external engines to access data in Amazon S3 locations with full table access.**  
 When you enable this option, Lake Formation will return credentials to the integrated application directly without IAM session tag validation.

## CLI

Use o comando `put-data-lake-settings` do CLI para definir os parâmetros a seguir.

Há três campos a serem configurados ao usar esse comando AWS CLI:

- `allow-external-data-filtering` — (boolean) Indica que um mecanismo de terceiros pode acessar informações de metadados não filtradas e credenciais de acesso a dados de recursos na conta corrente.
- `external-data-filtering-allow-list` — (matriz) Uma lista de IDs de conta que podem acessar informações de metadados não filtradas e credenciais de acesso a dados de recursos na conta atual ao usar um mecanismo de terceiros.
- `authorized-sessions-tag-value-list` — (matriz) Uma lista de valores de tags de sessão autorizados (cadeias). Se uma credencial de perfil do IAM tiver sido anexada a um par de chave e valor autorizado, se a tag da sessão for incluída na lista, a sessão terá acesso a informações de metadados não filtrados e credenciais de acesso a dados em recursos na conta configurada. A chave da tag de sessão autorizada é definida como `*LakeFormationAuthorizedCaller*`.
- `AllowFullTableExternalDataAccess` - (booleano) Permitir ou não que um mecanismo de consulta de terceiros obtenha credenciais de acesso a dados sem tags de sessão quando um chamador tiver permissões completas de acesso a dados.

Por exemplo: .

```
aws lakeformation put-data-lake-settings --cli-input-json file://
datalakesettings.json

{
  "DataLakeSettings": {
    "DataLakeAdmins": [
      {
        "DataLakePrincipalIdentifier": "arn:aws:iam::111111111111:user/lakeAdmin"
      }
    ],
    "CreateDatabaseDefaultPermissions": [],
    "CreateTableDefaultPermissions": [],
    "TrustedResourceOwners": [],
    "AllowExternalDataFiltering": true,
    "ExternalDataFilteringAllowList": [
      {"DataLakePrincipalIdentifier": "111111111111"}
    ],
    "AuthorizedSessionTagValueList": ["engine1"]
  }
}
```

```
"AllowFullTableExternalDataAccess": false
}
```

## API/SDK

Use a operação `PutDataLakeSetting` da API para definir os seguintes parâmetros.

Há três campos a serem configurados ao usar essa operação de API:

- `AllowExternalDataFiltering` — (boolean) Indica se um mecanismo de terceiros pode acessar informações de metadados não filtradas e credenciais de acesso a dados de recursos nesta conta .
- `ExternalDataFilteringAllowList` — (matriz) Uma lista de IDs de conta que podem acessar informações de metadados não filtradas e as credenciais de acesso aos dados dos recursos na conta atual usando um mecanismo de terceiros.
- `AuthorizedSectionsTagValueList` — (matriz) Uma lista de valores de tag autorizados (cadeias). Se uma credencial de perfil do IAM tiver sido anexada a uma tag autorizada, a sessão terá acesso às informações de metadados não filtradas e às credenciais de acesso aos dados nos recursos da conta configurada. A chave da tag de sessão autorizada é definida como `*LakeFormationAuthorizedCaller*`.
- `AllowFullTableExternalDataAccess` - (booleano) Permitir ou não que um mecanismo de consulta de terceiros obtenha credenciais de acesso a dados sem tags de sessão quando um chamador tiver permissões completas de acesso a dados.

Por exemplo: .

```
//Enable session tag on existing data lake settings
public void sessionTagSetUpForExternalFiltering(AWSLakeFormationClient
lakeformation) {
    GetDataLakeSettingsResult getDataLakeSettingsResult =
    lfClient.getDataLakeSettings(new GetDataLakeSettingsRequest());
    DataLakeSettings dataLakeSettings =
    getDataLakeSettingsResult.getDataLakeSettings();

    //set account level flag to allow external filtering
    dataLakeSettings.setAllowExternalDataFiltering(true);

    //set account that are allowed to call credential vending or Glue
    GetFilteredMetadata API
```

```
List<DataLakePrincipal> allowlist = new ArrayList<>();
allowlist.add(new
DataLakePrincipal().withDataLakePrincipalIdentifier("111111111111"));
dataLakeSettings.setWhitelistedForExternalDataFiltering(allowlist);

//set registered session tag values
List<String> registeredTagValues = new ArrayList<>();
registeredTagValues.add("engine1");
dataLakeSettings.setAuthorizedSessionTagValueList(registeredTagValues);

lakeformation.putDataLakeSettings(new
PutDataLakeSettingsRequest().withDataLakeSettings(dataLakeSettings));
}
```

## Integração de aplicativos para acesso total à tabela

Siga estas etapas para permitir que mecanismos de consulta de terceiros acessem dados sem a validação da tag de sessão do IAM:

### Console

1. Faça login no console do Lake Formation em <https://console.aws.amazon.com/lakeformation/>.
2. Na navegação à esquerda, expanda Permissões e selecione Configurações de integração de aplicativos.
3. Na página de configurações de integração de aplicativos, selecione a caixa de seleção Permitir que mecanismos externos acessem dados em locais do Amazon S3 com acesso total à tabela.

Quando você ativa essa opção, o Lake Formation retornará as credenciais diretamente para o aplicativo de consulta, sem a validação da tag de sessão do IAM.

# Application integration settings [Learn more](#)

## Application integration settings

Use the options below to control which third-party engines are allowed to read and filter data in Amazon S3 locations registered with Lake Formation.

Allow external engines to filter data in Amazon S3 locations registered with Lake Formation

Check this box to allow third-party engines to access data in Amazon S3 locations that are registered with Lake Formation.

### Session tag values

Enter one or more strings that match the LakeFormationAuthorizedCaller session tag defined for third-party engines.

Clear all

engine 1 ✕ engine 2 ✕ session 1 ✕

Enter one or several string values separated by comma.

### AWS account IDs

Enter the external AWS account IDs from where third-party engines are allowed to access locations registered with Lake Formation.

Clear all

111111111111 ✕ 222222222222 ✕  
Account Account

Enter one or more AWS account IDs. Press enter after each ID.

Allow external engines to access data in Amazon S3 locations with full table access.

When you enable this option, Lake Formation will return credentials to the integrated application directly without IAM session tag validation.

Cancel

Save

## AWS CLI

Use o comando `put-data-lake-settings` do CLI para definir os parâmetros em `AllowFullTableExternalDataAccess`.

```
aws lakeformation put-data-lake-settings --cli-input-json file://put-data-lake-
settings.json --region ap-northeast-1
{
  "DataLakeSettings": {
    "DataLakeAdmins": [
      {
        "DataLakePrincipalIdentifier": "arn:aws:iam::111111111111:user/
lakeAdmin"
      }
    ]
  }
}
```

```
    ],  
    "AllowFullTableExternalDataAccess": true  
  }  
}
```

# Trabalhando com outros AWS serviços

AWS serviços como Amazon Athena AWS Glue, Amazon Redshift Spectrum e Amazon EMR podem usar o Lake Formation para acessar com segurança dados em locais do Amazon S3 registrados no Lake Formation. Com o Lake Formation, você pode definir e gerenciar permissões refinadas de controle de acesso (FGAC) para seus dados no AWS Glue Data Catalog. Cada um desses AWS serviços é um chamador confiável para o Lake Formation, e o Lake Formation fornece acesso aos dados armazenados no Amazon S3 por meio de credenciais temporárias. Para ter mais informações, consulte [Como funciona a integração de aplicações do Lake Formation](#).

Para aproveitar esses recursos, o Lake Formation exige que você primeiro registre a localização do Amazon S3 e atribua as permissões apropriadas à entidade principal do IAM para acessar a tabela, o banco de dados e a localização do Amazon S3. Para obter mais informações, consulte, [Gerenciando permissões do Lake Formation](#).

## Tópicos


- [Usando AWS Lake Formation com o Amazon Athena](#)
- [Usando AWS Lake Formation com o Amazon Redshift Spectrum](#)
- [Usando AWS Lake Formation com AWS Glue](#)
- [Usando AWS Lake Formation com o Amazon EMR](#)
- [Usando AWS Lake Formation com a Amazon QuickSight](#)
- [Usando AWS Lake Formation com o AWS CloudTrail Lake](#)

## Usando AWS Lake Formation com o Amazon Athena

[Amazon Athena](#) é um serviço de consulta sem servidor que ajuda a analisar dados estruturados, semiestruturados e não estruturados armazenados no Amazon S3. O Athena oferece suporte à consulta de dados dos formatos de dados CSV, JSON, Parquet e Avro. O Athena também oferece suporte a formatos de tabela como tabelas controladas [Apache Hive](#), [Apache Hudi](#), [Apache Iceberg](#) e Lake Formation. O Athena se integra ao AWS Glue Data Catalog para armazenar metadados de seus conjuntos de dados no Amazon S3. O Athena pode usar o Lake Formation para definir e manter políticas de controle de acesso nesses conjuntos de dados.


Aqui estão alguns casos de uso comuns em que você pode usar o Lake Formation com o Athena.

- Use as permissões do Lake Formation para acessar os recursos do Catálogo de Dados (banco de dados e tabelas) do Athena. Você pode usar o método de recurso nomeado ou as LF-tags para definir permissões no banco de dados e nas tabelas. Para obter mais informações, consulte:
  - [Conceder permissões de banco de dados usando o método de recurso nomeado](#)
  - [Controle de acesso baseado em tags do Lake Formation](#)

 Note

As permissões do Lake Formation são válidas somente ao usar o Athena para consultar dados de origem do Amazon S3 e metadados no Catálogo de Dados.


As permissões do Lake Formation oferecem suporte a operações de leitura e gravação em bancos de dados e tabelas.

 Note

Você não pode aplicar filtros de dados ao usar LF-tags para gerenciar permissões nos recursos do Catálogo de Dados.

- Controle os resultados da consulta usando [Filtros de dados no Lake Formation](#) para proteger tabelas em seus data lakes do Amazon S3, concedendo permissões nos níveis de coluna, linha e célula. Veja a [limitação na projeção de partições](#) no Guia do usuário do Amazon Athena.
- Aplique um controle de acesso refinado dos dados disponíveis para o usuário do Athena baseado em SAML ao executar consultas federadas.

Os drivers Athena JDBC e ODBC oferecem suporte à configuração do acesso federado à sua fonte de dados usando o provedor de identidades (IdP) baseado em SAML. Use a Amazon QuickSight integrada ao Lake Formation com sua função existente do IAM ou usuários ou grupos do SAML para visualizar os resultados da consulta do Athena.

 Note

As permissões do Lake Formation para usuários e grupos SAML serão aplicadas somente quando você enviar consultas ao Athena usando o driver JDBC ou ODBC.



Para obter mais informações, consulte [Como usar o Lake Formation e drivers JDBC e ODBC do Athena para acesso federado ao Athena](#).

#### Note

Atualmente, não há suporte para autorizar o acesso às identidades SAML no Lake Formation nas seguintes regiões:

- Oriente Médio (Bahrein): me-south-1
- Ásia-Pacífico (Hong Kong): ap-east-1
- África (Cidade do Cabo): af-south-1
- China (Ningxia): cn-northwest-1
- Asia Pacific (Osaka): ap-northeast-3

- Use [Compartilhamento de dados entre contas no Lake Formation](#) para consultar tabelas em outra conta.

#### Note

Para obter mais informações sobre limitações ao usar as permissões do Lake Formation para Views, consulte [Considerações e limitações](#).

## Suporte a formatos de tabelas transacionais

A aplicação das permissões do Lake Formation permite que você proteja seus dados transacionais em seus data lakes baseados no Amazon S3. A tabela abaixo lista os formatos de tabela transacional compatíveis com as permissões do Athena e do Lake Formation. O Lake Formation impõe essas permissões quando os usuários do Athena executam suas consultas.

Formato da tabela	Descrição e operações permitidas	Permissões do Lake Formation possíveis no Athena
Apache Hudi	Um formato usado para simplificar o processamento	Use <a href="#">Filtragem de dados e segurança em nível de célula</a>

Formato da tabela	Descrição e operações permitidas	Permissões do Lake Formation possíveis no Athena
	<p>incremental de dados e o desenvolvimento de pipelines de dados.</p> <p>O Athena oferece suporte a operações de criação e leitura usando formatos de tabela Apache Hudi em conjuntos de dados do Amazon S3 para os tipos de tabela Hudi Copiar na Gravação (CoW) e Mesclar na Leitura (MoR). O Athena não suporta operações de gravação em tabelas Hudi.</p> <p>Use o <a href="#">Athena para consultar conjuntos de dados Hudi</a>.</p>	<p><a href="#">no Lake Formation</a> para proteger a tabela Hudi com permissões no nível de tabela, coluna, linha e célula.</p>
Apache Iceberg	<p>Um formato de tabela aberto que gerencia grandes coleções de arquivos como tabelas e oferece suporte a operações analíticas modernas de data lake, como inserção, atualização, exclusão e consultas de viagem no tempo em nível de registro.</p> <p>Para obter mais informações sobre o suporte do Athena para tabelas Iceberg, consulte <a href="#">Como usar tabelas Iceberg</a>.</p>	<p>Suporte para permissões em nível de tabela, coluna, linha e célula. Atualmente, o Lake Formation não oferece suporte ao gerenciamento de permissões em operações de gravação como VACUUM, MERGE, UPDATE e OPTIMIZE em tabelas em formatos de tabela aberta.</p>

Formato da tabela	Descrição e operações permitidas	Permissões do Lake Formation possíveis no Athena
Linux Foundation Delta Lake	<p>O Delta Lake é um projeto de código aberto que ajuda a implementar arquiteturas modernas de data lake, geralmente construídas no Amazon S3 ou no HDFS (Sistema de Arquivos Distribuídos do Hadoop).</p> <p>O Athena é compatível com tabelas Delta Lake criadas usando uma definição de tabela de manifesto baseada em links simbólicos a AWS Glue Data Catalog partir de uma tabela Delta Lake.</p> <p>Para obter mais informações, consulte <a href="#">Rastrear tabelas do Delta Lake usando AWS Glue rastreadores</a>.</p> <p>O Athena (motor de versão 3) suporta a leitura de tabelas nativas do Delta Lake.</p> <p>Para obter mais informações, consulte <a href="#">Apresentando o suporte de mesa nativo do Delta Lake com AWS Glue rastreadores</a>.</p>	Suporte para permissões em nível de tabela, coluna, linha e célula para tabelas de links simbólicos e tabelas nativas do Delta Lake.

## Recursos adicionais do

Publicações em blogs, vídeos e oficinas

- [Como consultar um conjunto de dados do Apache Hudi em um data lake do Amazon S3 com o Amazon Athena](#)
- [Crie um data lake Apache Iceberg usando o Amazon Athena, o Amazon EMR e AWS Glue](#)
- [Insira, atualize e exclua no Amazon S3 com Athena e Apache Iceberg](#)
- Oficina de [Controle de acesso baseado em LF-Tag](#) do Lake Formation sobre como consultar um data lake.

## Usando AWS Lake Formation com o Amazon Redshift Spectrum

O [Amazon Redshift Spectrum](#) permite que você consulte e recupere dados em data lakes do Amazon S3 sem que seja necessário carregar dados em nós de cluster do Amazon Redshift.

O Redshift Spectrum oferece suporte a duas formas de registrar um catálogo de AWS Glue dados externo habilitado com o Lake Formation.

- Como usar um perfil do IAM anexado ao cluster que tenha permissão para acessar o Catálogo de Dados

Para criar um perfil do IAM, siga as etapas descritas no procedimento abaixo.

[Para criar uma função do IAM para o Amazon Redshift usando um AWS Glue Data Catalog habilitado para AWS Lake Formation](#)

- Como usar identidade federada do IAM configurada para gerenciar o acesso a recursos AWS Glue Data Catalog externos

O Redshift Spectrum suporta a consulta de tabelas do Lake Formation usando identidades federadas do IAM. As identidades do IAM podem ser um usuário do IAM ou um perfil do IAM. Para obter mais informações sobre a federação de identidades IAM no Redshift Spectrum, consulte [Como usar uma identidade federada para gerenciar o acesso do Amazon Redshift a recursos locais e tabelas externas do Redshift Spectrum](#).

Com a integração do Lake Formation com o Redshift Spectrum, você pode definir permissões de controle de acesso em nível de linha, coluna e célula nas tabelas depois que seus dados forem registrados no Lake Formation.

Para obter mais informações, consulte [Usando o Redshift Spectrum](#) com AWS Lake Formation

O Redshift Spectrum suporta leituras ou consultas SELECT nas tabelas de esquema externo gerenciadas pelo Lake Formation.

Para obter mais informações, confira [Como criar esquemas externos para Redshift Spectrum](#).

## Suporte para tipos de tabelas transacionais

Esta tabela lista os formatos de tabela transacional suportados no Redshift Spectrum e as permissões aplicáveis do Lake Formation.

Formatos de tabela compatíveis

Formato da tabela	Descrição e operações permitidas	Permissões do Lake Formation compatíveis com o Redshift Spectrum
Apache Hudi	<p>Um formato usado para simplificar o processamento incremental de dados e o desenvolvimento de pipelines de dados.</p> <p>O Redshift Spectrum suporta operações de gravação de inserção, exclusão e inserção usando o formato de tabela Apache Hudi <a href="#">Copiar na Gravação (CoW)</a> no Amazon S3.</p> <p>Para obter mais informações, consulte <a href="#">Criação de tabelas externas para dados gerenciados no Apache Hudi</a>.</p>	<p>Use <a href="#">Filtragem de dados e segurança em nível de célula no Lake Formation</a> para proteger a tabela Hudi com permissões no nível de tabela, coluna, linha e célula.</p>

Formato da tabela	Descrição e operações permitidas	Permissões do Lake Formation compatíveis com o Redshift Spectrum
Apache Iceberg	<p>Um formato de tabela aberto que gerencia grandes coleções de arquivos como tabelas e oferece suporte a operações analíticas modernas de data lake, como inserção, atualização, exclusão e consultas de viagem no tempo em nível de registro.</p> <p>Para obter mais informações, confira <a href="#">Como usar tabelas do Apache Iceberg com o Amazon Redshift</a>.</p>	O Redshift Spectrum oferece suporte a tabelas do Apache Iceberg para consultas.
Linux Foundation Delta Lake	<p>O Delta Lake é um projeto de código aberto que ajuda a implementar arquiteturas modernas de data lake, geralmente construídas no Amazon S3 ou no HDFS (Sistema de Arquivos Distribuídos do Hadoop).</p> <p>O Redshift Spectrum oferece suporte à consulta de tabelas Delta Lake. Para obter mais informações, consulte <a href="#">Criação de tabelas externas para dados gerenciados no Delta Lake</a>.</p>	Suporte para permissões em nível de tabela, coluna, linha e célula.

## Recursos adicionais do

Publicações em blogs e oficinas

- [Centralize a governança do seu data lake usando, AWS Lake Formation ao mesmo tempo, uma arquitetura de dados moderna com o Amazon Redshift Spectrum](#)
- [Use o Redshift Spectrum para consultar as tabelas do Apache HUDI Copiar na Gravação \(CoW\) no data lake do Amazon S3](#)

## Usando AWS Lake Formation com AWS Glue

Engenheiros e DevOps profissionais de dados usam AWS Glue o Extract, Transform and Load (ETL) com o Apache Spark para realizar transformações em seus conjuntos de dados no Amazon S3 e carregar os dados transformados em lagos de dados e armazéns de dados para análise, aprendizado de máquina e desenvolvimento de aplicativos. Com equipes diferentes acessando o mesmo conjunto de dados no Amazon S3, é imperativo conceder e restringir permissões com base em seus perfis.

AWS Lake Formation é construído e AWS Glue os serviços interagem das seguintes maneiras:

- Lake Formation e AWS Glue compartilham o mesmo Catálogo de Dados.
- Os seguintes atributos do console Lake Formation invocam o console AWS Glue:
  - Trabalhos — Para obter mais informações, consulte [Como adicionar trabalhos](#) no Guia do desenvolvedor AWS Glue .
  - Crawlers - Para obter mais informações, consulte [Catalogação de tabelas com um Crawler](#) Guia do desenvolvedor AWS Glue .
- Os fluxos de trabalho gerados quando você usa um esquema do Lake Formation são fluxos de trabalho AWS Glue. Você pode visualizar e gerenciar esses fluxos de trabalho no console do Lake Formation e no console AWS Glue.
- As transformações de machine learning são fornecidas com o Lake Formation e são baseadas em operações de API do AWS Glue. Você cria e gerencia transformações de machine learning no console AWS Glue. Para obter mais informações, consulte [Transformações de machine learning](#) no Guia do desenvolvedor AWS Glue .

Você pode usar o controle de acesso refinado do Lake Formation para gerenciar seus recursos existentes do Catálogo de Dados e os locais de dados do Amazon S3.

**Note**

AWS Glue O ETL exige acesso total à tabela inteira enquanto busca dados da localização subjacente do Amazon S3. AWS Glue O trabalho de ETL falhará se você aplicar permissões em nível de coluna em uma tabela. No entanto, você pode criar segurança em nível de coluna e linha definindo filtros de dados. Para obter mais informações, consulte [Notas e restrições para filtragem em nível de coluna](#) Lake Formation avalia o filtro de dados definido na tabela e recupera somente os dados filtrados do Amazon S3 necessários para o trabalho de ETL. AWS Glue

## Suporte para tipos de tabelas transacionais

A aplicação das permissões do Lake Formation permite que você proteja seus dados transacionais em seus data lakes baseados no Amazon S3. A tabela abaixo lista os formatos de tabela transacional suportados AWS Glue e as permissões do Lake Formation. Lake Formation impõe essas permissões para AWS Glue operações.

### Formatos de tabela compatíveis

Formato da tabela	Descrição e operações permitidas	Permissões do Lake Formation suportadas em AWS Glue
Apache Hudi	Um formato de tabela aberta usado para simplificar o processamento incremental de dados e o desenvolvimento de pipelines de dados.  Para exemplos, consulte <a href="#">Usando a estrutura Hudi em AWS Glue</a> .	As permissões em nível de tabela estão disponíveis para tabelas do Hudi.  Para obter mais informações, consulte <a href="#">Limitações</a> .
Apache Iceberg	Um formato de tabela aberta que gerencia grandes coleções de arquivos como tabelas.	As permissões em nível de tabela estão disponíveis para tabelas do Iceberg.



Formato da tabela	Descrição e operações permitidas	Permissões do Lake Formation suportadas em AWS Glue
	Para obter exemplos, consulte <a href="#">Usando a estrutura Iceberg em AWS Glue</a> .	Para obter mais informações, consulte <a href="#">Limitações</a> .
Linux Foundation Delta Lake	<p>O Delta Lake é um projeto de código aberto que ajuda a implementar arquiteturas modernas de data lake, geralmente construídas no Amazon S3 ou no HDFS (Sistema de Arquivos Distribuídos do Hadoop).</p> <p>Para ver exemplos, consulte <a href="#">Usando a estrutura Delta Lake em AWS Glue</a>.</p>	<p>As permissões em nível de tabela estão disponíveis para tabelas do Delta Lake.</p> <p>Para obter mais informações, consulte <a href="#">Limitações</a>.</p>

## Recursos adicionais do

### Publicações em blogs e repositórios

- [Use o AWS Glue conector para ler e gravar tabelas Apache Iceberg com transações ACID e realizar viagens no tempo](#)
- [Escrevendo em tabelas do Apache Hudi usando conector personalizado AWS Glue](#)
- AWS repositório do [modelo Cloudformation e amostra de código do pyspark](#) para analisar dados de streaming usando o Apache Hudi e o AWS Glue Amazon S3.

## Usando AWS Lake Formation com o Amazon EMR

O Amazon EMR é uma plataforma de cluster AWS gerenciada flexível na qual você pode executar qualquer código personalizado em estruturas de big data compatíveis, como Hadoop Map-Reduce, Spark, Hive, Presto etc. As organizações também usam o Amazon EMR para executar aplicativos

de processamento de dados em lote e streaming em um cluster altamente distribuído. Usando o Amazon EMR, você pode executar suas transformações de dados e código personalizado em bancos de dados e tabelas cujas permissões são gerenciadas pelo Lake Formation.

Existem três opções para implementar o Amazon EMR:

- EMR no EC2
- Tecnologia sem servidor do EMR
- Amazon EMR no EKS

Para obter mais informações, consulte [Integrar o Amazon EMR com o Lake Formation](#) ou Usar o [EMR Serverless](#) com para um controle de acesso refinado AWS Lake Formation

## Suporte a formatos de tabelas transacionais

As versões 6.15.0 e posteriores do Amazon EMR incluem suporte para permissões de controle de acesso em nível de tabela, linha, coluna e célula do Lake Formation nos formatos [Apache Hudi](#), [Apache Iceberg](#) e [Delta Lake](#) ao ler e gravar dados com o Spark SQL.

Formatos de tabela compatíveis

Formato da tabela	Descrição e operações permitidas	Permissões do Lake Formation aceitas no Amazon EMR
Apache Hudi	Um formato de tabela aberta usado para simplificar o processamento incremental de dados e o desenvolvimento de pipelines de dados.  Para obter uma lista das operações compatíveis, consulte <a href="#">Apache Hudi e Lake Formation</a> .	O Amazon EMR oferece suporte ao controle de acesso no nível de tabela, linha, coluna e célula com o Apache Hudi.
Apache Iceberg	Um formato de tabela aberta que gerencia grandes	O Amazon EMR oferece suporte ao controle de acesso no nível de tabela, linha,

Formato da tabela	Descrição e operações permitidas	Permissões do Lake Formation aceitas no Amazon EMR
	<p>coleções de arquivos como tabelas.</p> <p>Para obter uma lista das operações compatíveis, consulte <a href="#">Apache Iceberg e Lake Formation</a>.</p>	coluna e célula com o Apache Iceberg.
Linux Foundation Delta Lake	<p>O Delta Lake é um projeto de código aberto que ajuda a implementar arquiteturas modernas de data lake, geralmente construídas no Amazon S3 ou no HDFS (Sistema de Arquivos Distribuídos do Hadoop).</p> <p>Para obter uma lista das operações compatíveis, consulte <a href="#">Delta Lake e Lake Formation</a>.</p>	O Amazon EMR oferece suporte ao controle de acesso em nível de tabela, linha, coluna e célula com tabelas Delta Lake.

## Recursos adicionais do

Guia do usuário, postagens de blog e oficinas

- [Integração com o Amazon EMR usando perfis de runtime](#)
- [Comece a usar Apache Hudi, Apache Iceberg e Delta Lake com o Amazon EMR no EKS](#)
- [Usar o Delta Lake OSS com o EMR Sem Servidor](#)

## Usando AWS Lake Formation com a Amazon QuickSight

A Amazon QuickSight oferece suporte à exploração de conjuntos de dados gerenciados pelas permissões do Lake Formation no Amazon S3 usando o Athena.

Os usuários das edições Standard e Enterprise da Amazon QuickSight se integram ao Lake Formation, mas de forma um pouco diferente.

- Edição Enterprise — conceda permissões refinadas de controle de acesso (FGAC) a QuickSight usuários individuais, grupos e funções do IAM da Amazon para acessar bancos de dados e tabelas.
- Edição padrão — conceda permissões às perfis do IAM para acessar bancos de dados e tabelas.

### Note

Por padrão, a Amazon QuickSight usa uma função chamada `aws-quicksight-service-role-v0`. Você também pode definir funções personalizadas com as permissões necessárias que permitem que QuickSight a Amazon acesse o Athena.

Para obter mais informações, consulte [Autorização de conexões por meio de AWS Lake Formation](#)

## Recursos adicionais do

Publicações no blog

- [Habilite permissões refinadas para autores da Amazon em QuickSight AWS Lake Formation](#)
- [Analise seus dados com segurança com AWS Lake Formation a Amazon QuickSight](#)

## Usando AWS Lake Formation com o AWS CloudTrail Lake

AWS CloudTrail O Lake suporta a exploração de armazenamentos de dados Amazon Athena de eventos usando permissões refinadas em. AWS Lake Formation

### Note

CloudTrail O lago só pode ser consultado. Amazon Athena

Para registrar seu armazenamento de dados de eventos do CloudTrail Lake no Lake Formation, consulte [Federar um armazenamento de dados de eventos](#).

# Como fazer log de chamados de API do Lake Formation do AWS usando AWS CloudTrail

O AWS Lake Formation é integrado ao AWS CloudTrail, um serviço que fornece um registro das ações realizadas por um usuário, uma função ou um serviço do AWS no Lake Formation. O CloudTrail captura todas as chamadas de API do Lake Formation como eventos. Os chamados capturados incluem chamados do console do Lake Formation do AWS Command Line Interface, o e chamados de código para as ações da API do Lake Formation. Se você criar uma trilha, poderá ativar a entrega contínua de eventos do CloudTrail em um bucket do Amazon S3, incluindo eventos para o Lake Formation. Se você não configurar uma trilha, ainda poderá visualizar os eventos mais recentes no console do CloudTrail em Histórico de eventos. Usando as informações coletadas pelo CloudTrail, é possível determinar a solicitação feita ao Lake Formation, o endereço IP a partir do qual a solicitação foi feita, quem fez a solicitação, quando ela foi feita e outros detalhes.

Para saber mais sobre o CloudTrail, consulte o [Guia do usuário do AWS CloudTrail](#).

## Informações sobre o Lake Formation no CloudTrail

O CloudTrail é habilitado por padrão quando você cria uma nova conta da AWS. Quando ocorre uma atividade no Lake Formation, essa atividade é registrada como um evento do CloudTrail junto com outros eventos do serviço AWS no Histórico de eventos. Um evento representa uma única solicitação de qualquer origem e inclui informações sobre a ação solicitada, a data e a hora da ação e os parâmetros de solicitação. Cada entrada de log ou evento contém informações sobre quem gerou a solicitação. As informações de identidade ajudam a determinar:

- Se a solicitação foi feita com credenciais de usuário raiz ou usuário do IAM AWS Identity and Access Management
- Se a solicitação foi feita com credenciais de segurança temporárias de um perfil ou de um usuário federado.
- Se a solicitação foi feita por outro serviço da AWS.

Para obter mais informações, consulte [Elemento userIdentity do CloudTrail](#).

Você pode visualizar, pesquisar e fazer download de eventos recentes para a sua conta da AWS. Para obter mais informações, consulte [Como visualizar eventos com o histórico de eventos do CloudTrail](#).

Para obter um registro contínuo de eventos na sua conta da AWS, incluindo eventos para Lake Formation, crie uma trilha. Uma trilha permite que o CloudTrail entregue arquivos de log a um bucket do Amazon S3. Por padrão, quando você cria uma trilha no console, ela é aplicada a todas as regiões do AWS. A trilha registra em log eventos de todas as regiões na partição da AWS e entrega os arquivos de log para o bucket do Amazon S3 especificado por você. Além disso, você pode configurar outros serviços do AWS, como o Amazon Athena, para analisar e agir com base nos dados de eventos coletados nos logs do CloudTrail. O CloudTrail também pode entregar arquivos de log para o Amazon CloudWatch Logs e o CloudWatch Events.

Para obter mais informações, consulte:

- [Visão geral da criação de uma trilha](#)
- [Serviços e integrações compatíveis com o CloudTrail](#)
- [Configuração notificações do Amazon SNS para o CloudTrail](#)
- [Como receber arquivos de log do CloudTrail de várias regiões](#) e [Como receber arquivos de log do CloudTrail de várias contas](#)

## Como compreender os eventos do Lake Formation

Todas as ações da API do Lake Formation são registradas pelo CloudTrail e documentadas no Guia do Desenvolvedor do AWS Lake Formation. Por exemplo, as chamadas para as APIs `PutDataLakeSettings`, `GrantPermissions` e `RevokePermissions` geram entradas nos arquivos de log do CloudTrail.

O exemplo a seguir mostra um evento do CloudTrail para a ação `GrantPermissions`. A entrada inclui o usuário que concedeu a permissão (`datalake_admin`), a entidade principal à qual a permissão foi concedida (`datalake_user1`) e a permissão que foi concedida (`CREATE_TABLE`). A entrada também mostra que a concessão falhou porque o banco de dados de destino não foi especificado no argumento `resource`.

```
{
  "eventVersion": "1.08",
  "userIdentity": {
    "type": "IAMUser",
    "principalId": "AIDAZKE67KM3P775X74U2",
    "arn": "arn:aws:iam::111122223333:user/datalake_admin",
    "accountId": "111122223333",
    "accessKeyId": "...",
```

```

    "userName": "datalake_admin"
  },
  "eventTime": "2021-02-06T00:43:21Z",
  "eventSource": "lakeformation.amazonaws.com",
  "eventName": "GrantPermissions",
  "awsRegion": "us-east-1",
  "sourceIPAddress": "72.21.198.65",
  "userAgent": "aws-cli/1.19.0 Python/3.6.12
Linux/4.9.230-0.1.ac.223.84.332.metal1.x86_64 boto3/1.20.0",
  "errorCode": "InvalidInputException",
  "errorMessage": "Resource must have one of the have either the catalog, table or
database field populated.",
  "requestParameters": {
    "principal": {
      "dataLakePrincipalIdentifier": "arn:aws:iam::111122223333:user/
datalake_user1"
    },
    "resource": {},
    "permissions": [
      "CREATE_TABLE"
    ]
  },
  "responseElements": null,
  "requestID": "b85e863f-e75d-4fc0-9ff0-97f943f706e7",
  "eventID": "8d2ccef0-55f3-42d3-9ede-3a6faedaa5c1",
  "readOnly": false,
  "eventType": "AwsApiCall",
  "managementEvent": true,
  "eventCategory": "Management",
  "recipientAccountId": "111122223333"
}

```

Este exemplo mostra uma entrada de log do CloudTrail para a ação `GetDataAccess`. As entidades principais não chamam essa API diretamente. Em vez disso, `GetDataAccess` é registrado sempre que uma entidade principal ou serviço integrado da AWS solicita credenciais temporárias para acessar dados em um local de data lake registrado no Lake Formation.

```

{
  "eventVersion": "1.05",
  "userIdentity": {
    "type": "AWSAccount",
    "principalId": "AROAQGFTBBBG0BWV2EMZA:GlueJobRunnerSession",
    "accountId": "111122223333"
  }
}

```



```
    },  
    "eventSource": "lakeformation.amazonaws.com",  
    "eventName": "GetDataAccess",  
    ...  
    ...  
    "additionalEventData": {  
      "requesterService": "GLUE_JOB",  
      "lakeFormationPrincipal": "arn:aws:iam::111122223333:role/ETL-Glue-Role",  
      "lakeFormationRoleSessionName": "AWSLF-00-GL-111122223333-G13T0Rmng2"  
    },  
    ...  
  }  
}
```

### Consulte também

- [Registro em várias contas CloudTrail](#)

# Práticas recomendadas, considerações e limitações do Lake Formation

Use esta seção para encontrar rapidamente as práticas recomendadas, considerações e limitações no AWS Lake Formation.

Consulte [Service Quotas](#) para conhecer o número máximo de recursos de serviço ou operações da Conta da AWS.

## Tópicos

- [Práticas recomendadas e considerações sobre compartilhamento de dados entre contas](#)
- [Limitações de acesso aos dados entre regiões](#)
- [Considerações e limitações das visualizações do catálogo de dados](#)
- [Limitações de filtragem de dados](#)
- [Considerações e limitações do modo de acesso híbrido](#)
- [Considerações e limitações do compartilhamento de dados de armazenamento de metadados do Hive](#)
- [Limitações do compartilhamento de dados do Amazon Redshift](#)
- [Limitações da integração com o Centro de Identidade do IAM](#)
- [Considerações e práticas recomendadas de controle de acesso com base em tags do Lake Formation](#)
- [Formatos e limitações compatíveis para compactação gerenciada de dados](#)

## Práticas recomendadas e considerações sobre compartilhamento de dados entre contas

Os recursos de várias contas do Lake Formation permitem que os usuários compartilhem com segurança lagos de dados distribuídos em várias AWS organizações ou diretamente com os diretores do IAM em outra conta Contas da AWS, fornecendo acesso refinado aos metadados do Catálogo de Dados e aos dados subjacentes.

Pense nas seguintes práticas recomendadas ao usar o compartilhamento de dados entre contas do Lake Formation:

- Não há limite para o número de concessões de permissão do Lake Formation que você pode conceder aos diretores em sua própria AWS conta. No entanto, o Lake Formation usa a capacidade AWS Resource Access Manager (AWS RAM) para concessões entre contas que sua conta pode fazer com o método de recurso nomeado. Para maximizar a AWS RAM capacidade, siga estas práticas recomendadas para o método de recurso nomeado:
  - Use o novo modo de concessão entre contas (versão 3 e superior em Configurações de versão entre contas) para compartilhar um recurso com um externo Conta da AWS. Para ter mais informações, consulte [Como atualizar as configurações da versão de compartilhamento de dados entre contas](#).
  - Organize AWS contas em organizações e conceda permissões a organizações ou unidades organizacionais. Uma concessão para uma organização ou unidade organizacional conta como apenas uma concessão.

A concessão para organizações ou unidades organizacionais também elimina a necessidade de aceitar um convite AWS Resource Access Manager (AWS RAM) de compartilhamento de recursos para a concessão. Para ter mais informações, consulte [Acessar e visualizar tabelas e bancos de dados compartilhados do Catálogo de Dados](#).

- Em vez de conceder permissões em várias tabelas individuais do banco de dados, use o curinga especial Todas as tabelas para conceder permissões em todas as tabelas do banco de dados. A concessão em Todas as tabelas conta como uma única concessão. Para ter mais informações, consulte [Concedendo e revogando permissões nos recursos do Catálogo de dados](#).

#### Note

Para obter mais informações sobre como solicitar um limite maior para o número de compartilhamentos de recursos em AWS RAM, consulte [cotas AWS de serviço](#) no Referência geral da AWS

- Você deve criar um link de recurso para um banco de dados compartilhado para que esse banco de dados apareça nos editores de consulta Amazon Athena e no Amazon Redshift Spectrum. Da mesma forma, para poder consultar tabelas compartilhadas usando o Athena e o Redshift Spectrum, você deve criar links de recursos para as tabelas. Em seguida, os links de recursos aparecem na lista de tabelas dos editores de consulta.

Em vez de criar links de recursos para várias tabelas individuais para consulta, você pode usar o curinga Todas as tabelas para conceder permissões em todas as tabelas em um banco de dados. Em seguida, ao criar um link de recurso para esse banco de dados e selecionar esse

link de recurso de banco de dados no editor de consultas, você terá acesso a todas as tabelas desse banco de dados para sua consulta. Para ter mais informações, consulte [Criação de links de recursos](#).

- Quando você compartilha recursos diretamente com principais em outra conta, a entidade principal do IAM na conta do destinatário pode não ter permissão para criar links de recursos para poder consultar as tabelas compartilhadas usando o Athena e o Amazon Redshift Spectrum. Em vez de criar um link de recurso para cada tabela compartilhada, o administrador do data lake pode criar um banco de dados provisório e conceder a permissão `CREATE_TABLE` ao grupo `ALLIAMPPrincipals`. Em seguida, todas as entidades principais do IAM na conta do destinatário podem criar links de recursos no banco de dados de espaços reservados, e começar a consultar as tabelas compartilhadas.

Veja o exemplo de comando CLI para conceder permissões para `ALLIAMPPrincipals` em [Conceder permissões de banco de dados usando o método de recurso nomeado](#).

- O Athena e o Redshift Spectrum oferecem suporte ao controle de acesso em nível de coluna, mas somente para inclusão, não exclusão. O controle de acesso em nível de coluna não é suportado em tarefas de ETL AWS Glue.
- Quando um recurso é compartilhado com sua AWS conta, você pode conceder permissões sobre o recurso somente aos usuários da sua conta. Você não pode conceder permissões sobre o recurso para outras AWS contas, para organizações (nem mesmo para sua própria organização) ou para o `IAMAllowedPrincipals` grupo.
- Não é possível conceder `DROP` ou `Super` em um banco de dados a uma conta externa.
- Revogue as permissões entre contas antes de excluir um banco de dados ou uma tabela. Caso contrário, você deverá excluir compartilhamentos de recursos órfãos em AWS Resource Access Manager

#### Consulte também

- [Considerações e práticas recomendadas de controle de acesso com base em tags do Lake Formation](#)
- [CREATE\\_TABLE](#) na seção [Referência de permissões do Lake Formation](#) para obter mais regras e limitações de acesso entre contas.

## Limitações de acesso aos dados entre regiões

O Lake Formation aceita a consulta de tabelas do catálogo de dados entre Regiões da AWS. Você pode acessar dados em uma região de outras regiões usando o Amazon Athena Amazon EMR e o AWS Glue ETL criando links de recursos em outras regiões apontando para os bancos de dados e tabelas de origem. Com o acesso à tabela entre regiões, você pode acessar dados entre regiões sem copiar os dados subjacentes ou os metadados no Catálogo de Dados.

As limitações a seguir se aplicam ao acesso a tabelas entre regiões.

- O Lake Formation não suporta a consulta de tabelas do Catálogo de Dados de outra região usando o Amazon Redshift Spectrum.
- No console do Lake Formation, as visualizações do banco de dados e da tabela não mostram os nomes dos bancos de dados/tabelas da região de origem.
- Para exibir a lista de tabelas em um banco de dados compartilhado de outra região, você precisa primeiro criar um link de recurso para o banco de dados compartilhado, depois selecionar o link do recurso e escolher Exibir tabelas.
- O recurso de acesso a tabelas entre regiões não funciona quando você cria links de recursos Regiões da AWS nesse ponto para bancos de dados compartilhados e tabelas criadas em regiões opcionais.

Para obter mais informações, consulte [Optar por regiões na página de suporte Regiões da AWS e serviços](#).

- O Lake Formation não é compatível com chamadas de links de recursos entre regiões.

## Considerações e limitações das visualizações do catálogo de dados

Em AWS Glue Data Catalog, uma exibição é uma tabela virtual na qual o conteúdo é definido por uma consulta que faz referência a uma ou mais tabelas. É possível criar uma visualização que faça referência a até dez tabelas usando editores SQL para Amazon Athena, Amazon Redshift ou Amazon EMR. As tabelas de referência subjacentes de uma visualização podem pertencer ao mesmo banco de dados ou a bancos de dados diferentes na mesma Conta da AWS.

Estas são considerações e limitações que se aplicam às visualizações do catálogo de dados.

- O Amazon Redshift sempre cria visualizações com colunas varchar com base em tabelas com strings. É necessário converter colunas de string em varchar com um tamanho explícito ao adicionar dialetos de outros mecanismos.
- Conceder permissões de data lake a All views em um banco de dados resultará em permissões do favorecido em todas as tabelas e visualizações do banco de dados.
- Não é possível criar visualizações:
  - Isso faz referência a outras visualizações.
  - Quando a referência a uma tabela é um link de recurso.
  - Quando as tabelas de referência têm permissões de entidade principal IAM\_ALLOWED\_GROUP.
  - Quando a tabela de referência está em outra conta.
  - De metastores externos do Hive.

## Limitações de filtragem de dados

Ao conceder permissões do Lake Formation em uma tabela do Catálogo de Dados, você pode incluir especificações de filtragem de dados para restringir o acesso a determinados dados nos resultados da consulta e nos mecanismos integrados ao Lake Formation. O Lake Formation usa a filtragem de dados para obter segurança em nível de coluna, segurança em nível de linha e segurança de nível de célula. Será possível definir e aplicar filtros de dados em colunas aninhadas se os dados de origem contiverem estruturas aninhadas.

Tenha em mente as seguintes notas e restrições para filtragem em nível de linha e de célula:

- A segurança em nível de célula não é aceita em colunas aninhadas.
- Todas as expressões aceitas em colunas de nível superior também são aceitas em colunas aninhadas. No entanto, os campos aninhados em colunas de partição NÃO devem ser referenciados ao definir expressões aninhadas em nível de linha.
- A segurança em nível de célula está disponível em todas as regiões ao usar o Athena Engine versão 3 ou o Amazon Redshift Spectrum. Para outros serviços, a segurança em nível de célula só está disponível nas regiões mencionadas em [Regiões compatíveis](#).
- As instruções SELECT INTO não são compatíveis.
- Os tipos de dados array e map não são compatíveis com expressões de filtro de linha. O tipo de dados struct é aceito.

- Para executar operações de consulta em tabelas que usam filtragem em nível de linha e em nível de célula, você deve usar um grupo de trabalho especial chamado `AmazonAthenaLakeFormation`. Para obter mais informações sobre grupos de trabalho, consulte [Usar grupos de trabalho para executar consultas](#) no Manual do Usuário do Amazon Athena.
- Não há limite para o número de filtros de dados que podem ser definidos em uma tabela, mas há um limite de 100 permissões `SELECT` de filtro de dados para um único entidade principal em uma tabela.
- O número máximo de filtros de dados que podem ser incluídos em uma concessão em uma tabela é dez.
- Para aplicar um filtro de dados com uma expressão de filtro de linha, você deve ter a opção `SELECT` de concessão em todas as colunas da tabela. Essa restrição não se aplica a administradores em contas externas quando a concessão foi feita para a conta externa.
- Se uma entidade principal for membro de um grupo e tanto a entidade principal quanto o grupo receberem permissões em um subconjunto de linhas, as permissões de linha efetivas da entidade principal são a união das permissões do entidade principal e das permissões do grupo.
- Os seguintes nomes de colunas são restritos em uma tabela para filtragem em nível de linha e de célula:
  - `ctid`
  - `oid`
  - `xmin`
  - `cmin`
  - `xmax`
  - `cmax`
  - `ttableoid`
  - `insertxid`
  - `deletexid`
  - `importoid`
  - `redcatuniqueid`
- Se você aplicar a expressão de filtro de todas as linhas em uma tabela simultaneamente com outras expressões de filtro com predicados, a expressão de todas as linhas prevalecerá sobre todas as outras expressões de filtro.
- Quando as permissões em um subconjunto de linhas são concedidas a uma AWS conta externa e o administrador do data lake da conta externa concede essas permissões a um principal nessa

conta, o predicado de filtro efetivo do principal é a interseção do predicado da conta com qualquer predicado que tenha sido concedido diretamente ao principal.

Por exemplo, se a conta tiver permissões de linha com o predicado `dept='hr'` e a entidade principal tiver recebido a permissão `country='us'` separadamente, a entidade principal terá acesso somente às linhas com `dept='hr'` e `country='us'`.

Para obter mais informações sobre a filtragem em nível de célula, consulte [Filtragem de dados e segurança em nível de célula no Lake Formation](#).

## Considerações e limitações do modo de acesso híbrido

O modo de acesso híbrido oferece a flexibilidade de habilitar seletivamente as permissões do Lake Formation para bancos de dados e tabelas em seu AWS Glue Data Catalog.

Com o modo de acesso híbrido, agora você tem um caminho incremental que permite definir permissões do Lake Formation para um conjunto específico de usuários sem interromper as políticas de permissão de outros usuários ou workload existentes.

As considerações e limitações a seguir se aplicam ao modo de acesso híbrido.

### Limitações

- Atualizar o registro de localização do Amazon S3 – Você não pode editar parâmetros de um local registrado no Lake Formation usando uma função vinculada ao serviço.
- Opção de ativação ao usar LF-Tags – Quando você pode conceder permissões do Lake Formation usando LF-Tags, você pode optar por entidades principais para aplicar as permissões do Lake Formation em uma etapa consecutiva, escolhendo bancos de dados e tabelas com LF-Tags anexadas.
- Optar por entidades principais – Atualmente, somente uma função de administrador de data lake pode optar por entidades principais para recursos.
- Aceitar todas as tabelas em um banco de dados: em concessões entre contas, ao conceder permissões e aceitar todas as tabelas em um banco de dados, é necessário aceitar o banco de dados também para que as permissões funcionem.



## Considerações

- Atualização da localização do Amazon S3 registrada no Lake Formation para o modo de acesso híbrido – Não recomendamos converter uma localização de dados do Amazon S3 que já esteja registrada no Lake Formation para o modo de acesso híbrido, embora isso possa ser feito.
- Comportamentos da API quando um local de dados é registrado no modo de acesso híbrido
  - CreateTable — O local é considerado registrado no Lake Formation, independentemente da bandeira do modo de acesso híbrido e do status de opt-in. Assim, o usuário precisa da permissão de localização dos dados para criar uma tabela.
  - CreatePartition/BatchCreatePartitions/UpdatePartitions (quando a localização da partição é atualizada para apontar para a localização registrada com híbrido) — A localização do Amazon S3 é considerada registrada no Lake Formation, independentemente da bandeira do modo de acesso híbrido e do status de aceitação. Assim, o usuário precisa da permissão de localização de dados para criar ou atualizar um banco de dados.
  - CreateDatabase/UpdateDatabase (quando a localização do banco de dados é atualizada para apontar para a localização registrada no modo de acesso híbrido) — A localização é considerada registrada no Lake Formation, independentemente da bandeira do modo de acesso híbrido e do status de ativação. Assim, o usuário precisa da permissão de localização de dados para criar ou atualizar um banco de dados.
  - UpdateTable (quando a localização de uma tabela é atualizada para apontar para a localização registrada no modo de acesso híbrido) — A localização é considerada registrada no Lake Formation, independentemente da bandeira do modo de acesso híbrido e do status de ativação. Assim, o usuário precisa de permissão de localização de dados para atualizar a tabela. Se a localização da tabela não for atualizada ou estiver apontando para uma localização que não esteja registrada no Lake Formation, o usuário não precisará da permissão de localização de dados para atualizar a tabela.

## Considerações e limitações do compartilhamento de dados de armazenamento de metadados do Hive

Com a federação de AWS Glue Data Catalog metadados (federação do catálogo de dados), você pode conectar o catálogo de dados a metastores externos que armazenam metadados para seus dados do Amazon S3 e gerenciar com segurança as permissões de acesso aos dados usando AWS Lake Formation

As seguintes considerações e limitações se aplicam aos bancos de dados federados criados a partir dos bancos de dados do Hive:

### Considerações

- **AWS SAM suporte de aplicativos** — Você é responsável pela disponibilidade dos recursos do aplicativo que são AWS SAM implantados (Amazon API Gateway e pela função Lambda). Certifique-se de que a conexão entre o metastore AWS Glue Data Catalog e o Hive esteja funcionando quando os usuários executam consultas.
- **Requisito da versão do metastore do Hive:** é possível criar bancos de dados federados somente usando o Apache Hive versão 3 e posterior.
- **Requisito de banco de dados mapeado** — Todo banco de dados do Hive deve ser mapeado para um novo banco de dados no Lake Formation.
- **Suporte à federação em nível de banco de dados** — Você pode se conectar ao repositório do Hive somente no nível do banco de dados.
- **Permissões em bancos de dados federados** — As permissões aplicadas em um banco de dados federado ou tabelas em um banco de dados federado persistem mesmo quando uma tabela de origem ou um banco de dados é excluído. Quando o banco de dados ou tabela de origem é recriado, você não precisa reconceder as permissões. Quando uma tabela federada com permissões do Lake Formation é excluída na fonte, as permissões do Lake Formation ainda estão visíveis e você pode revogá-las se necessário.

Se um usuário excluir um banco de dados federado, todas as permissões correspondentes serão perdidas. Recriar o mesmo banco de dados com o mesmo nome não recuperará as permissões do Lake Formation. Os usuários precisarão configurar novas permissões novamente.

- **Permissões de `AllowedPrincipal` grupo do IAM em bancos de dados federados** — Com base no `DataLakeSettings`, Lake Formation pode definir permissões para todos os bancos de dados e tabelas para um grupo virtual chamado `IAMAllowedPrincipal`. O `IAMAllowedPrincipal` se refere a todos os diretores do IAM que têm acesso aos recursos do catálogo de dados por meio das políticas principais e políticas de AWS Glue recursos do IAM. Se essas permissões existirem em um banco de dados ou tabela, todos as entidades principais terão acesso ao banco de dados ou à tabela.

No entanto, o Lake Formation não aceita permissões `IAMAllowedPrincipal` em tabelas em bancos de dados federados. Ao criar bancos de dados federados, certifique-se de passar o parâmetro `CreateTableDefaultPermissions` como uma lista vazia.

Para ter mais informações, consulte [Alterando as configurações padrão do seu data lake](#).

- Unir tabelas em consultas — Você pode unir tabelas de repositório do Hive com tabelas nativas do Catálogo de Dados para executar consultas.

## Limitações

- Limitação na sincronização de metadados entre o AWS Glue Data Catalog e o metastore do Hive — Depois de estabelecer a conexão do metastore do Hive, você precisa criar um banco de dados federado para sincronizar os metadados no metastore do Hive com o. AWS Glue Data Catalog As tabelas no banco de dados federado são sincronizadas em runtime quando os usuários executam consultas.
- Limitação na criação de novas tabelas em um banco de dados federado — Você não poderá criar novas tabelas em bancos de dados federados.
- Limitação de permissão de dados — O suporte para permissões nas visualizações de tabela do Repositório do Hive não está disponível.

## Limitações do compartilhamento de dados do Amazon Redshift

AWS Lake Formation permite que você gerencie dados com segurança em um compartilhamento de dados do Amazon Redshift. O Amazon Redshift é um serviço de armazém de dados totalmente gerenciado em escala de petabytes na nuvem. AWS Usando o recurso de compartilhamento de dados, o Amazon Redshift ajuda você a compartilhar dados entre Contas da AWS. Para obter mais informações sobre o compartilhamento de dados do Amazon Redshift, consulte [Visão geral do compartilhamento de dados no Amazon Redshift](#).

As seguintes observações e restrições aplicam-se a bancos de dados federados criados a partir de compartilhamentos de dados do Amazon Redshift:

- Requisito de banco de dados mapeado — Toda unidade de compartilhamento de dados do Amazon Redshift deve ser mapeada em um novo banco de dados no Lake Formation. Isso é necessário para manter nomes de tabela exclusivos quando a representação dos objetos da unidade de compartilhamento de dados é nivelada no banco de dados do Catálogo de Dados.
- Limitação na criação de novas tabelas em um banco de dados federado — Você não poderá criar novas tabelas em bancos de dados federados.

- **Permissões nos bancos de dados federados** — As permissões aplicadas em um banco de dados federado ou tabelas em um banco de dados federado persistem mesmo quando uma tabela de origem ou um banco de dados é excluído. Quando o banco de dados ou a tabela de origem são recriados, você não precisa conceder as permissões novamente. Quando uma tabela federada com permissões do Lake Formation é excluída na fonte, as permissões do Lake Formation ainda estarão visíveis e você poderá revogá-las se necessário.

Se um usuário excluir um banco de dados federado, todas as permissões correspondentes serão perdidas. Recriar o mesmo banco de dados com o mesmo nome não recuperará as permissões do Lake Formation. Os usuários precisarão configurar novas permissões novamente.

- **Permissões de `AllowedPrincipal` grupo do IAM em bancos de dados federados** — Com base no `DataLakeSettings`, Lake Formation pode definir permissões para todos os bancos de dados e tabelas para um grupo virtual chamado `IAMAllowedPrincipal`. O `IAMAllowedPrincipal` se refere a todos os diretores do IAM que têm acesso aos recursos do catálogo de dados por meio das políticas principais e políticas de AWS Glue recursos do IAM. Se essas permissões existirem em um banco de dados ou tabela, todos as entidades principais terão acesso ao banco de dados ou à tabela.

No entanto, o Lake Formation não aceita permissões `IAMAllowedPrincipal` em tabelas em bancos de dados federados. Ao criar bancos de dados federados, certifique-se de passar o parâmetro `CreateTableDefaultPermissions` como uma lista vazia.

Para ter mais informações, consulte [Alterando as configurações padrão do seu data lake](#).

- **Filtragem de dados** — No Lake Formation, você pode conceder permissões em uma tabela em um banco de dados federado com filtragem em nível de coluna e em nível de linha. No entanto, você não pode combinar a filtragem em nível de coluna e em nível de linha para restringir o acesso na granularidade em nível de célula em tabelas em bancos de dados federados.
- **Identificador de distinção entre maiúsculas e minúsculas** — Os objetos de unidades de compartilhamento de dados do Amazon Redshift gerenciados pelo Lake Formation suportarão nomes de tabelas e nomes de colunas somente em minúsculas. Não ative o identificador de diferenciação de maiúsculas e minúsculas para bancos de dados, tabelas e colunas nos compartilhamentos de dados do Amazon Redshift, caso eles sejam compartilhados e gerenciados usando o Lake Formation.

Para obter mais informações sobre limitações ao trabalhar com unidades de compartilhamento de dados no Amazon Redshift, [consulte Limitações para a unidade de compartilhamento de dados](#) no Guia do desenvolvedor de banco de dados do Amazon Redshift.

## Limitações da integração com o Centro de Identidade do IAM

Com AWS IAM Identity Center, você pode se conectar a provedores de identidade (IdPs) e gerenciar centralmente o acesso de usuários e grupos em todos os serviços de AWS análise. Você pode configurar AWS Lake Formation como um aplicativo habilitado no IAM Identity Center, e os administradores do data lake podem conceder permissões refinadas a usuários e grupos autorizados sobre recursos. AWS Glue Data Catalog

As seguintes limitações se aplicam à integração do Lake Formation com o Centro de Identidade do IAM:

- Não é possível atribuir usuários e grupos do Centro de Identidade do IAM como administradores de data lake ou administradores somente leitura no Lake Formation.
- Usuários e grupos do IAM Identity Center não podem consultar tabelas do catálogo de dados criptografadas usando chaves AWS Key Management Service (AWS KMS). AWS KMS não oferece suporte à propagação de identidade confiável.
- Usuários e grupos do Centro de Identidade do IAM só podem invocar operações de API listadas na política `AWSIAMIdentityCenterAllowListForIdentityContext` fornecida pelo Centro de Identidade do IAM.

## Considerações e práticas recomendadas de controle de acesso com base em tags do Lake Formation

É possível criar, manter e atribuir tags do LF para controlar o acesso a bancos de dados, tabelas e colunas do catálogo de dados.

Pense nas seguintes práticas recomendadas ao usar o controle de acesso com base em tags do Lake Formation:

- Todas as LF-Tags devem ser predefinidas antes de poderem ser atribuídas aos recursos do Catálogo de Dados ou concedidas às entidades principais.

O administrador do data lake pode delegar tarefas de gerenciamento de tags gerando criadores de LF-Tags com as permissões necessárias do IAM. Os engenheiros e analistas de dados decidem sobre as características e os relacionamentos das LF-Tags. Os criadores do LF-Tag então criam e mantêm os LF-Tags no Lake Formation.

- Você pode atribuir várias LF-Tags aos recursos do Catálogo de Dados. Somente um valor para uma chave específica pode ser atribuído a um recurso específico.

Por exemplo, você pode atribuir `module=Orders`, `region=West` e `division=Consumer` e assim por diante a um banco de dados, uma tabela ou uma coluna. Você não pode atribuir `module=Orders,Customers`.

- Você não pode atribuir LF-Tags aos recursos ao criá-los. Você só pode adicionar LF-Tags aos recursos existentes.
- Você pode conceder expressões de LF-Tag, não apenas LF-Tags únicas, a uma entidade principal.

Uma expressão de LF-Tag se parece com a seguinte (em pseudocódigo).

```
module=sales AND division=(consumer OR commercial)
```

Uma entidade principal que recebe essa expressão de LF-Tag pode acessar somente os recursos do Catálogo de Dados (bancos de dados, tabelas e colunas) que receberam `module=sales` e `division=consumer` ou `division=commercial`. Se você quiser que a entidade principal possa acessar recursos que tenham `module=sales` ou `division=commercial`, não inclua ambos na mesma concessão. Faça duas concessões, uma para `module=sales` e outra para `division=commercial`.

A expressão LF-tag mais simples consiste em apenas uma LF-Tag, como `module=sales`.

- Uma entidade principal que recebe permissões em uma LF-Tag com vários valores pode acessar os recursos do Catálogo de Dados com qualquer um deles. Por exemplo, se um usuário receber uma LF-Tag com `key=module` e `values=orders,customers`, o usuário terá acesso aos recursos atribuídos `module=orders` ou `module=customers`.
- Você precisa ter permissão `Grant with LF-Tag expressions` para conceder permissões de dados nos recursos do Catálogo de Dados usando o método LF-TBAC. O administrador do data lake e o criador da tag do LF recebem implicitamente essa permissão. Uma entidade principal que tenha a permissão `Grant with LFTag expressions` pode conceder permissões de dados sobre os recursos usando:
  - o método de recurso nomeado
  - o método LF-TBAC, mas usando apenas a mesma expressão LF-tag

Por exemplo, suponha que o administrador do data lake faça a seguinte concessão (em pseudocódigo).

```
GRANT (SELECT ON TABLES) ON TAGS module=customers, region=west,south TO user1 WITH GRANT OPTION
```

Nesse caso, `user1` pode conceder `SELECT` em tabelas a outras entidades principais usando o método `LF-TBAC`, mas somente com a expressão `LF-tag` completa `module=customers, region=west, south`.

- Se uma entidade principal receber permissões em um recurso com o método `LF-TBAC` e o método de recurso nomeado, as permissões que a entidade principal tem sobre o recurso são a união das permissões concedidas pelos dois métodos.
- O Lake Formation oferece suporte à concessão de `DESCRIBE` e `ASSOCIATE` em `LF-Tags` em todas as contas e à concessão de permissões nos recursos do Catálogo de Dados em todas as contas usando o método `LF-TBAC`. Em ambos os casos, o principal é o ID AWS da conta.

#### Note

O Lake Formation aceita concessões entre contas para organizações e unidades organizacionais usando o método `LF-TBAC`. Para usar esse recurso, você precisa atualizar as Configurações de versão entre contas para a Versão 3.

Para ter mais informações, consulte [Compartilhamento de dados entre contas no Lake Formation](#).

- Os recursos do Catálogo de Dados criados em uma conta só podem ser marcados usando `LF-Tags` criadas na mesma conta. As `LF-Tags` criadas em uma conta não podem ser associadas a recursos compartilhados de outra conta.
- Usar o controle de acesso baseado em tags do Lake Formation (`LF-TBAC`) para conceder acesso entre contas aos recursos do Catálogo de Dados requer acréscimos à política de recursos do Catálogo de Dados para sua conta. AWS Para ter mais informações, consulte [Pré-requisitos](#).
- As chaves da `LF-Tag` e os valores da `LF-Tag` não podem exceder 50 caracteres de comprimento.
- O número máximo de `LF-Tags` que podem ser atribuídas a um recurso do Catálogo de Dados é 50.
- Os seguintes limites são limites flexíveis:
  - O número máximo de `LF-tags` que podem ser criados é 1000.
  - O número máximo de valores que podem ser definidos para uma `LF-Tag` é 1000.
- As chaves e valores das tags são convertidos em letras minúsculas quando são armazenados.

- Somente um valor para uma LF-Tag pode ser atribuído a um recurso específico.
- Se várias LF-Tags forem concedidas a uma entidade principal com uma única concessão, a entidade principal poderá acessar somente os recursos do Catálogo de Dados que tenham todas as LF-Tags.
- Os trabalhos de ETL do AWS Glue exigem acesso total à tabela. Os trabalhos falharão se a função ETL do AWS Glue não tiver acesso a todas as colunas em uma tabela. É possível aplicar tags LF em nível de coluna, mas isso pode fazer com que as funções de AWS Glue ETL percam o acesso total à tabela e façam com que os trabalhos falhem. O uso de filtros de dados para filtragem de colunas e/ou linhas não é afetado por essa limitação.
- Se uma avaliação da expressão LF-tag resultar em acesso somente a um subconjunto de colunas da tabela, mas a permissão Lake Formation concedida quando há uma correspondência for uma das permissões que exigiram acesso total à coluna, ou seja, `Alter`, `Drop`, `Insert` ou `Delete`, nenhuma dessas permissões será concedida. Em vez disso, somente `Describe` é concedido. Se a permissão concedida for `All (Super)`, somente `Select` e `Describe` serão concedidas.
- Curingas não são usados com tags do LF. Para atribuir uma tag do LF a todas as colunas de uma tabela, atribua a tag do LF à tabela e todas as colunas na tabela herdam a tag do LF. Para atribuir uma tag do LF a todas as tabelas em um banco de dados, atribua a tag do LF ao banco de dados, e todas as tabelas no banco de dados herdam essa tag do LF.

## Formatos e limitações compatíveis para compactação gerenciada de dados

Para melhor desempenho de leitura por serviços de AWS análise, como Amazon Athena, Amazon EMR e trabalhos de AWS Glue ETL, AWS Glue Data Catalog fornece compactação gerenciada (um processo que compacta pequenos objetos do Amazon S3 em objetos maiores) para tabelas Iceberg no Data Catalog.

A compactação de dados aceita uma variedade de tipos de dados e formatos de compactação para leitura e gravação de dados, incluindo a leitura de dados de tabelas criptografadas.

A compactação de dados suporta:

- Tipos de dados: Booleano, Inteiro, Longo, Flutuante, Duplo, String, Decimal, Data, Hora, Timestamp, String, UUID, Binário
- Compressão: `zstd`, `gzip`, `snappy`, não compactado



- Criptografia: a compactação de dados suporta somente a criptografia padrão do Amazon S3 (SSE-S3) e a criptografia KMS do lado do servidor (SSE-SKMS).
- Compactação do compartimento
- Evolução do esquema
- Tabelas com tamanho de arquivo de destino (gravação). `target-file-size-bytes` propriedade na configuração do iceberg) dentro da faixa inclusiva de 128 MB a 512 MB.
- Regiões
  - Ásia-Pacífico (Tóquio)
  - Ásia-Pacífico (Seul)
  - Ásia-Pacífico (Mumbai)
  - Europa (Irlanda)
  - Europa (Frankfurt)
  - Leste dos EUA (Norte da Virgínia)
  - Leste dos EUA (Ohio)
  - Oeste dos EUA (N. da Califórnia)
  - América do Sul (São Paulo)
- Você pode executar a compactação a partir da conta em que o catálogo de dados reside quando o bucket do Amazon S3 que armazena os dados subjacentes estiver em outra conta. Para fazer isso, a função de compactação exige acesso ao bucket do Amazon S3.

Atualmente, a compactação de dados não oferece suporte a:

- Tipos de dados: Fixo
- Compressão: brotli, lz4
- Compactação de arquivos enquanto a especificação da partição evolui.
- Classificação regular ou classificação por ordem z
- Mesclar ou excluir arquivos: o processo de compactação ignora os arquivos de dados que têm arquivos excluídos associados a eles.
- Compactação em tabelas de contas cruzadas: você não pode executar a compactação em tabelas de contas cruzadas.
- Compactação de tabelas entre regiões: não é possível executar a compactação de tabelas entre regiões.

- **Habilitando a compactação em links de recursos**
- **Endpoints da VPC para o Amazon S3**

# Solução de problemas do Lake Formation

Se você encontrar problemas ao trabalhar com o AWS Lake Formation, consulte os tópicos desta seção.

## Tópicos

- [Solução de problemas gerais](#)
- [Resolução de problemas de acesso entre contas](#)
- [Solução de problemas em esquemas e fluxos de trabalho](#)
- [Problemas conhecidos do AWS Lake Formation](#)
- [Mensagem de erro atualizada](#)

## Solução de problemas gerais

Use as informações aqui para ajudá-lo a diagnosticar e corrigir vários problemas do Lake Formation.

### Erro: permissões insuficientes do Lake Formation em <Amazon S3 location>

Foi feita uma tentativa de criar ou alterar um recurso do Catálogo de Dados sem permissões de localização de dados na localização do Amazon S3 apontada pelo recurso.

Se um banco de dados ou tabela do Data Catalog apontar para uma localização do Amazon S3, ao conceder as permissões do Lake Formation CREATE\_TABLE ou ALTER, você também deverá conceder a permissão DATA\_LOCATION\_ACCESS no local. Ao conceder permissões a contas ou organizações externas, é necessário incluir a opção de concessão.

Depois que essas permissões forem concedidas a uma conta externa, o administrador do data lake dessa conta deverá conceder as permissões às entidades principais (usuários ou funções) na conta. Ao conceder a permissão DATA\_LOCATION\_ACCESS recebida de outra conta, você deve especificar a ID do catálogo (ID da conta AWS) da conta do proprietário. A conta do proprietário é a conta que registrou o local.

Para obter mais informações, consulte [Controle de acesso a dados subjacente](#) e [Concedendo permissões de localização de dados](#).

## Erro: “permissões de chave de criptografia insuficientes para a API Glue”

Foi feita uma tentativa de conceder permissões do Lake Formation sem permissões do AWS Identity and Access Management (IAM) na chave de criptografia AWS KMS de um catálogo de dados criptografado.

## Minha consulta do Amazon Athena ou do Amazon Redshift que usa manifestos está falhando

O Lake Formation não suporta consultas que usam manifestos.

## Erro: "permissão(ões) do Lake Formation insuficiente(s): necessária a criação de tag no catálogo"

O usuário/função deve ser administrador do data lake.

## Erro ao excluir administradores de data lake inválidos

É necessário excluir todos os administradores de data lake inválidos (perfis do IAM excluídos que são definidos como administradores de data lake) simultaneamente. Se você tentar excluir separadamente administradores de data lake inválidos, o Lake Formation gerará um erro de entidade principal inválida.

## Resolução de problemas de acesso entre contas

Use estas informações para ajudar a diagnosticar e corrigir problemas de acesso entre contas.

### Tópicos

- [Eu concedi uma permissão para várias contas do Lake Formation, mas o destinatário não consegue ver o recurso](#)
- [As entidades principais da conta do destinatário podem ver o recurso do Catálogo de Dados, mas não podem acessar os dados subjacentes](#)
- [Erro: “falha na associação porque o chamador não foi autorizado” ao aceitar um convite de compartilhamento de recursos AWS RAM.](#)
- [Erro: “não autorizado a conceder permissões para o recurso”](#)
- [Erro: “acesso negado para recuperar informações da organização da AWS”](#)
- [Erro: “organização <organization-ID> não encontrada”](#)

- [Erro: "permissões insuficientes do Lake Formation: combinação ilegal"](#)
- [ConcurrentModificationException em solicitações de concessão/revogação para contas externas](#)
- [Erro ao usar o Amazon EMR para acessar dados compartilhados por meio de várias contas](#)

Eu concedi uma permissão para várias contas do Lake Formation, mas o destinatário não consegue ver o recurso

- O usuário na conta do destinatário é administrador do data lake? Somente administradores de data lake podem ver o recurso no momento do compartilhamento.
- Você está compartilhando com uma conta externa à sua organização usando o método de recurso nomeado? Nesse caso, o administrador do data lake da conta do destinatário deve aceitar um convite de compartilhamento de recursos em AWS Resource Access Manager (AWS RAM).

Para obter mais informações, consulte [the section called "Aceitar um convite de compartilhamento de recursos do AWS RAM"](#).

- Você está usando políticas de recursos em nível de conta (Catálogo de dados) em AWS Glue? Se sim, se você usar o método de recursos nomeados, deverá incluir uma declaração especial na política que autorize a AWS RAM o compartilhamento de políticas em seu nome.

Para obter mais informações, consulte [the section called "Gerenciamento de permissões entre contas usando A e Lake Formation"](#).

- Você tem as permissões do AWS Identity and Access Management (IAM) necessárias para conceder acesso entre contas?

Para obter mais informações, consulte [the section called "Pré-requisitos"](#).

- O recurso para o qual você concedeu permissões não deve ter nenhuma permissão do Lake Formation concedida ao grupo IAMAllowedPrincipals.
- Há uma declaração deny sobre o recurso na política em nível de conta?

As entidades principais da conta do destinatário podem ver o recurso do Catálogo de Dados, mas não podem acessar os dados subjacentes

As entidades principais da conta do destinatário devem ter as permissões do AWS Identity and Access Management (IAM) necessárias. Para obter mais detalhes, consulte [Como acessar os dados subjacentes de uma tabela compartilhada](#).

## Erro: “falha na associação porque o chamador não foi autorizado” ao aceitar um convite de compartilhamento de recursos AWS RAM.

Depois de conceder acesso a um recurso em uma conta diferente, quando a conta receptora tenta aceitar o convite de compartilhamento de recursos, a ação falha.

```
$ aws ram get-resource-share-associations --association-type PRINCIPAL --resource-
share-arns arn:aws:ram:aws-region:444444444444:resource-share/e1d1f4ba-xxxx-xxxx-xxxx-
xxxxxxxx5d8d
{
  "resourceShareAssociations": [
    {
      "resourceShareArn": "arn:aws:ram:aws-region:444444444444:resource-share/
e1d1f4ba-xxxx-xxxx-xxxx-xxxxxxxx5d8d
",
      "resourceShareName": "LakeFormation-MMCC0XQBH3Y",
      "associatedEntity": "5815803XXXXX",
      "associationType": "PRINCIPAL",
      "status": "FAILED",
      "statusMessage": "Association failed because the caller was not
authorized.",
      "creationTime": "2021-07-12T02:20:10.267000+00:00",
      "lastUpdatedTime": "2021-07-12T02:20:51.830000+00:00",
      "external": true
    }
  ]
}
```

O erro ocorre porque `glue:PutResourcePolicy` é invocado pelo AWS Glue quando a conta receptora aceita o convite de compartilhamento de recursos. Para resolver o problema, permita a ação `glue:PutResourcePolicy` pela função assumida usada pela conta do produtor/concedente.

## Erro: “não autorizado a conceder permissões para o recurso”

Foi feita uma tentativa de conceder permissões entre contas em um banco de dados ou tabela pertencente a outra conta. Quando um banco de dados ou tabela é compartilhado com sua conta, como administrador do data lake, você pode conceder permissões sobre ele somente aos usuários da sua conta.

## Erro: “acesso negado para recuperar informações da organização da AWS”

Sua conta é uma conta de gerenciamento das organizações da AWS e você não tem as permissões necessárias para recuperar informações da organização, como unidades organizacionais na conta.

Para obter mais informações, consulte [Required permissions for cross-account grants](#).

## Erro: “organização <organization-ID> não encontrada”

Foi feita uma tentativa de compartilhar um recurso com uma organização, mas o compartilhamento com organizações não está habilitado. Habilitar o compartilhamento de recursos com organizações.

Para obter mais informações, consulte [Habilitar compartilhamento com organizações da AWS](#) no Guia do usuário do AWS RAM.

## Erro: "permissões insuficientes do Lake Formation: combinação ilegal"

Um usuário compartilhou um recurso do Catálogo de Dados enquanto as permissões do Lake Formation foram concedidas ao grupo IAMAllowedPrincipals para o recurso. O usuário deve revogar todas as permissões do Lake Formation de IAMAllowedPrincipals antes de compartilhar o recurso.

## ConcurrentModificationException em solicitações de concessão/revogação para contas externas

Quando os usuários fazem várias solicitações simultâneas de concessão e/ou revogação de permissão para uma entidade principal nas políticas de tag do LF, o Lake Formation lança ConcurrentModificationException. Os usuários precisam capturar a exceção e tentar novamente a solicitação de concessão/revogação que falhou. Usar versões em lote das operações GrantPermissions/RevokePermissionsAPI - [batchGrantPermissions](#) e [batchRevokePermissions](#) alivia esse problema até certo ponto, reduzindo o número de solicitações simultâneas de concessão/revogação.

## Erro ao usar o Amazon EMR para acessar dados compartilhados por meio de várias contas

Quando você usa o Amazon EMR para acessar dados de outra conta compartilhados com você, algumas bibliotecas do Spark tentarão chamar a operação de API

`Glue:GetUserDefinedFunctions`. Como as versões 1 e 2 das permissões gerenciadas pelo AWS RAM não oferecem suporte a essa ação, você receberá esta mensagem de erro:

```
"ERROR: User: arn:aws:sts::012345678901:assumed-role/my-spark-role/i-06ab8c2b59299508a is not authorized to perform: glue:GetUserDefinedFunctions on resource: arn:exampleCatalogResource because no resource-based policy allows the glue:GetUserDefinedFunctions action"
```

Para resolver esse erro, o administrador do data lake que criou o compartilhamento de recursos deve atualizar as permissões gerenciadas pelo AWS RAM anexadas ao compartilhamento de recursos. A versão 3 das permissões gerenciadas pelo AWS RAM permite que as entidades principais executem a ação `glue:GetUserDefinedFunctions`.

Se você criar um novo compartilhamento de recursos, o Lake Formation aplicará a versão mais recente da permissão gerenciada pelo AWS RAM por padrão, e você não precisará realizar nenhuma ação. Para habilitar o acesso a dados entre contas para compartilhamentos de recursos atuais, é necessário atualizar as permissões gerenciadas pelo AWS RAM para a versão 3.

Visualize as permissões do AWS RAM atribuídas aos recursos compartilhados com você no AWS RAM. As permissões incluídas na versão 3 são estas:

```
Databases
  AWSRAMPermissionGlueDatabaseReadWriteForCatalog
  AWSRAMPermissionGlueDatabaseReadWrite

Tables
  AWSRAMPermissionGlueTableReadWriteForCatalog
  AWSRAMPermissionGlueTableReadWriteForDatabase

AllTables
  AWSRAMPermissionGlueAllTablesReadWriteForCatalog
  AWSRAMPermissionGlueAllTablesReadWriteForDatabase
```

Para atualizar a versão de permissões gerenciadas pelo AWS RAM dos compartilhamentos de recursos atuais

Você (administrador do data lake) pode [atualizar as permissões gerenciadas pelo AWS RAM para uma versão mais recente](#) seguindo as instruções no Guia do usuário do AWS RAM ou revogar todas as permissões atuais para o tipo de recurso e concedê-las novamente. Se você revogar



as permissões, o AWS RAM excluirá o compartilhamento de recursos do AWS RAM associado ao tipo de recurso. Quando você concede permissões novamente, o AWS RAM cria novos compartilhamentos de recursos anexando a versão mais recente das permissões gerenciadas pelo AWS RAM.

## Solução de problemas em esquemas e fluxos de trabalho

Use as informações contidas aqui para ajudar a diagnosticar e corrigir problemas no esquema e no fluxo de trabalho.

### Tópicos

- [O meu esquema falhou com "usuário: <user-ARN> não está autorizado a realizar: iam:PassRole no recurso: <role-ARN>"](#)
- [O meu esquema falhou com "usuário: <user-ARN> não está autorizado a realizar: iam:PassRole no recurso: <role-ARN>"](#)
- [Um crawler no meu fluxo de trabalho falhou com "o recurso não existe ou o solicitante não está autorizado a acessar as permissões solicitadas"](#)
- [Um crawler no meu fluxo de trabalho falhou com "ocorreu um erro \(AccessDeniedException\) ao chamar a operação CreateTable..."](#)

### O meu esquema falhou com "usuário: <user-ARN> não está autorizado a realizar: iam:PassRole no recurso: <role-ARN>"

Foi feita uma tentativa de criar um esquema por um usuário que não tem permissões suficientes para passar a função escolhida.

Atualize a política do IAM do usuário para poder transmitir a função ou peça que ele escolha uma função diferente com as permissões de senha necessárias.

Para obter mais informações, consulte [the section called "Referência de personas e permissões do IAM do Lake Formation"](#).

### O meu esquema falhou com "usuário: <user-ARN> não está autorizado a realizar: iam:PassRole no recurso: <role-ARN>"

A função que você especificou para o fluxo de trabalho não tinha uma política em linha que permitisse que a função se transmitisse sozinha.

Para obter mais informações, consulte [the section called “\(Opcional\) Crie uma função do IAM para fluxos de trabalho”](#).

Um crawler no meu fluxo de trabalho falhou com “o recurso não existe ou o solicitante não está autorizado a acessar as permissões solicitadas”

Uma possível causa é que a função passada não tinha permissões suficientes para criar uma tabela no banco de dados de destino. Conceda à função a permissão CREATE\_TABLE no banco de dados.

Um crawler no meu fluxo de trabalho falhou com “ocorreu um erro (AccessDeniedException) ao chamar a operação CreateTable...”

Uma possível causa é que a função do fluxo de trabalho não tinha permissões de localização de dados no local de armazenamento de destino. Conceda permissões de localização de dados para a função.

Para obter mais informações, consulte [the section called “DATA\\_LOCATION\\_ACCESS”](#).

## Problemas conhecidos do AWS Lake Formation

Analise esses problemas conhecidos do AWS Lake Formation.

### Tópicos

- [Limitação na filtragem de metadados da tabela](#)
- [Problema ao renomear uma coluna excluída](#)
- [Problema com a exclusão de colunas em tabelas CSV](#)
- [As partições da tabela devem ser adicionadas em um caminho comum](#)
- [Problema com a criação de um banco de dados durante a criação do fluxo de trabalho](#)
- [Problema com a exclusão e a recriação de um usuário](#)
- [As APIs GetTables e SearchTables não atualizam o valor do parâmetro IsRegisteredWithLakeFormation](#)
- [As operações da API do Catálogo de dados não atualizam o valor do parâmetro IsRegisteredWithLakeFormation](#)
- [As operações do Lake Formation não oferecem suporte ao Registro de esquemas AWS Glue](#)

## Limitação na filtragem de metadados da tabela

Permissões em nível de coluna AWS Lake Formation podem ser usadas para restringir o acesso a colunas específicas em uma tabela. Quando um usuário recupera metadados sobre a tabela usando o console ou uma API como `glue:GetTable`, a lista de colunas no objeto da tabela contém somente os campos aos quais ele tem acesso. É importante entender as limitações dessa filtragem de metadados.

Embora o Lake Formation disponibilize metadados sobre permissões de coluna para serviços integrados, a filtragem real das colunas nas respostas da consulta é de responsabilidade do serviço integrado. Os clientes do Lake Formation que oferecem suporte à filtragem em nível de coluna, incluindo Amazon Athena, Amazon Redshift Spectrum e Amazon EMR, filtram os dados com base nas permissões de coluna registradas no Lake Formation. Os usuários não poderão ler nenhum dado ao qual não devem ter acesso. Atualmente, o ETL AWS Glue não oferece suporte à filtragem de colunas.

### Note

Os clusters do EMR não são totalmente gerenciados por AWS. Portanto, é responsabilidade dos administradores do EMR proteger adequadamente os clusters para evitar o acesso não autorizado aos dados.

Certas aplicações ou formatos podem armazenar metadados adicionais, incluindo nomes e tipos de colunas, no mapa `Parameters` como propriedades da tabela. Essas propriedades são retornadas sem modificações e podem ser acessadas por qualquer usuário com permissão `SELECT` em qualquer coluna.

Por exemplo, o [Avro SerDe](#) armazena uma representação JSON do esquema da tabela em uma propriedade de tabela nomeada como `avro.schema.literal`, que está disponível para todos os usuários com acesso à tabela. Recomendamos que você evite armazenar informações confidenciais nas propriedades da tabela e esteja ciente de que os usuários podem aprender o esquema completo das tabelas no formato Avro. Essa limitação é específica para os metadados sobre uma tabela.

AWS Lake Formation remove qualquer propriedade da tabela, começando com `spark.sql.sources.schema` ao responder a uma solicitação `glue:GetTable` ou similar, se o chamador não tiver permissões `SELECT` em todas as colunas da tabela. Isso impede que os usuários tenham acesso a metadados adicionais sobre tabelas criadas com o Apache Spark.

Quando executadas no Amazon EMR, as aplicações do Apache Spark ainda podem ler essas tabelas, mas certas otimizações podem não ser aplicadas e nomes de colunas com distinção entre maiúsculas e minúsculas não são aceitos. Se o usuário tiver acesso a todas as colunas na tabela, o Lake Formation retornará a tabela sem modificações com todas as propriedades da tabela.

## Problema ao renomear uma coluna excluída

Se você usar permissões em nível de coluna para excluir uma coluna e depois renomeá-la, a coluna não será mais excluída das consultas, assim como `SELECT *`.

## Problema com a exclusão de colunas em tabelas CSV

Se você criar uma tabela do Catálogo de dados com o formato CSV e depois excluir uma coluna do esquema, as consultas poderão retornar dados errados e as permissões em nível de coluna poderão não ser respeitadas.

Solução alternativa: em vez disso, crie uma nova tabela.

## As partições da tabela devem ser adicionadas em um caminho comum

O Lake Formation espera que todas as partições de uma tabela estejam em um caminho comum definido no campo de localização da tabela. Quando você usa o crawler para adicionar partições a um catálogo, isso funciona perfeitamente. Mas se você adicionar partições manualmente e essas partições não estiverem no local definido na tabela principal, o acesso aos dados não funcionará.

## Problema com a criação de um banco de dados durante a criação do fluxo de trabalho

Ao criar um fluxo de trabalho a partir de um esquema usando o console do Lake Formation, você pode criar o banco de dados de destino, caso ele não exista. Quando você faz isso, o usuário que está conectado recebe a permissão `CREATE_TABLE` no banco de dados criado. No entanto, o crawler que o fluxo de trabalho gera assume a função do fluxo de trabalho ao tentar criar uma tabela. Isso falha porque a função não possui a permissão `CREATE_TABLE` no banco de dados.

Solução alternativa: se você criar o banco de dados por meio do console durante a configuração do fluxo de trabalho, antes de executar o fluxo de trabalho, deverá conceder à função associada ao fluxo de trabalho a permissão `CREATE_TABLE` no banco de dados que você acabou de criar.

## Problema com a exclusão e a recriação de um usuário

O cenário a seguir resulta em permissões errôneas do Lake Formation retornadas por `lakeformation:ListPermissions`:

1. Crie um usuário e conceda permissões do Lake Formation.
2. Exclua o usuário.
3. Recrie o usuário com o mesmo nome.

`ListPermissions` retorna duas entradas, uma para o usuário antigo e outra para o novo usuário. Se você tentar revogar as permissões concedidas ao usuário antigo, as permissões serão revogadas do novo usuário.

## As APIs `GetTables` e `SearchTables` não atualizam o valor do parâmetro `IsRegisteredWithLakeFormation`

Há uma limitação conhecida de que as operações da API do Catálogo de dados, como `GetTables` e `SearchTables`, não atualizam o valor de `IsRegisteredWithLakeFormation` parameter e retornam o padrão, que é falso. É recomendável usar a API `GetTable` para visualizar o valor correto para `IsRegisteredWithLakeFormation` parameter.

## As operações da API do Catálogo de dados não atualizam o valor do parâmetro `IsRegisteredWithLakeFormation`

Há uma limitação conhecida de que as operações da API do Catálogo de dados, como `GetTables` e `SearchTables`, não atualizam o valor do parâmetro `IsRegisteredWithLakeFormation` e retornam o padrão, que é falso. É recomendável usar a API `GetTable` para visualizar o valor correto do parâmetro `IsRegisteredWithLakeFormation`.

## As operações do Lake Formation não oferecem suporte ao Registro de esquemas AWS Glue

As operações do Lake Formation não oferecem suporte a tabelas AWS Glue que contenham um `SchemaReference` no `StorageDescriptor` a ser utilizado no [Registro de esquemas](#).

## Mensagem de erro atualizada

AWSO Lake Formation atualizou as exceções específicas do recurso para a mensagem de erro geral `EntityNotFound` das seguintes operações de API para atender aos objetivos de segurança e conformidade.

- `RevokePermissions`
- `GrantPermissions`
- `GetResourceLFTags`
- `GetTables`
- `GetDatabase`

# API AWS Lake Formation

## Note

A [Referência de API](#) atualizada para o serviço AWS Lake Formation já está disponível.

## Sumário

- [APIs de permissões](#)
  - [Operações](#)
  - [Tipos de dados](#)
- [APIs de configurações do data lake](#)
  - [Operações](#)
  - [Tipos de dados](#)
- [APIs de integração com o Centro de Identidade do IAM](#)
  - [Operações](#)
  - [Tipos de dados](#)
- [APIs de modo de acesso híbrido](#)
  - [Operações](#)
  - [Tipos de dados](#)
- [APIs de fornecimento de credenciais](#)
  - [Operações](#)
  - [Tipos de dados](#)
- [APIs de Tags](#)
  - [Operações](#)
  - [Tipos de dados](#)
- [APIs de filtro de dados](#)
  - [Operações](#)
  - [Tipos de dados](#)
- [Tipos de dados comuns](#)
  - [Estrutura ErrorDetail](#)

- [Padrões de string](#)

## APIs de permissões

A seção API de permissões descreve as operações e os tipos de dados necessários para conceder e revogar permissões no AWS Lake Formation. Consulte o [Guia de referência da API do Lake Formation](#) para ver todas as operações e tipos de dados da API do AWS Lake Formation.

## Operações

- [Conceder permissões](#)
- [Revogar permissões](#)
- [Conceder permissões em lote](#)
- [Revogar permissões em lote](#)
- [Obter permissões efetivas para o path](#)
- [Lista de Permissões](#)

## Tipos de dados

- [Recurso](#)
- [Recurso de banco de dados](#)
- [Recurso da tabela](#)
- [Recurso de tabela com colunas](#)
- [Recurso de filtro de células de dados](#)
- [Recurso de local de dados](#)
- [Entidade principal do Data Lake](#)
- [Permissões da entidade principal](#)
- [Permissões de entidades principais](#)
- [Mapa de detalhes](#)
- [Coluna Curinga](#)
- [Entrada de solicitação de permissões em lote](#)
- [Entrada de falha de permissões de lote](#)



## APIs de configurações do data lake

Esta seção contém as operações da API de configurações do data lake e os tipos de dados para gerenciar os administradores do data lake.

### Operações

- [Obter configurações do Data Lake](#)
- [Incluir configurações do Data Lake](#)

### Tipos de dados

- [Configurações do Data Lake](#)

## APIs de integração com o Centro de Identidade do IAM

Esta seção contém as operações para criar e gerenciar a integração do Lake Formation com o Centro de Identidade do IAM.

### Operações

- [CreateLakeFormationIdentityCenterConfiguration](#)
- [DeleteLakeFormationIdentityCenterConfiguration](#)
- [DescribeLakeFormationIdentityCenterConfiguration](#)
- [UpdateLakeFormationIdentityCenterConfiguration](#)

### Tipos de dados

- [ExternalFilteringConfiguration](#)

## APIs de modo de acesso híbrido

A seção API de modo de acesso híbrido descreve as operações e os tipos de dados necessários para configurar o modo de acesso híbrido no AWS Lake Formation. Consulte o [Guia de referência](#)

[da API do Lake Formation](#) para ver todas as operações e tipos de dados da API do AWS Lake Formation.

## Operações

- [Criar opção de inclusão do Lake Formation](#)
- [Excluir opção de inclusão do Lake Formation](#)
- [Listar opções de inclusão do Lake Formation](#)

## Tipos de dados

- [Recurso](#)
- [Recurso de banco de dados](#)
- [Recurso da tabela](#)
- [Informações sobre o recurso](#)
- [Informações sobre as opções de inclusão do Lake Formation](#)
- [Recurso de local de dados](#)

## APIs de fornecimento de credenciais

A seção API de fornecimento de credenciais descreve as operações e os tipos de dados relacionados ao trabalho com o serviço AWS Lake Formation para fornecer credenciais e para registrar e gerenciar um recurso de data lake.

## Operações

- [Registrar recurso](#)
- [Cancelar o registro do recurso](#)
- [Listar recursos](#)
- [Obter metadados de tabela não filtrados](#)
- [Obter metadados de partições não filtradas](#)
- [Obter credenciais temporárias de partição do Glue](#)
- [Obter credenciais temporárias da tabela do Glue](#)
- [Atualizar recurso](#)

## Tipos de dados

- [Condição do filtro](#)
- [Filtro de linha](#)
- [Informações sobre o recurso](#)

## APIs de Tags

A seção API de tags descreve as operações e os tipos de dados relacionados a uma estratégia de autorização que define um modelo de permissões em atributos ou tags de pares de valores-chave.

## Operações

- [Adicionar LF-tags ao recurso](#)
- [Remover LF-tags do recurso](#)
- [Obter recursos de LF-tags](#)
- [Listar LF-tags](#)
- [Criar LF-tags](#)
- [Obter LF-tags](#)
- [Atualizar LF-tags](#)
- [Excluir LF-tags](#)
- [Pesquisar tabelas por LF-tags](#)
- [Pesquisar bancos de dados por LF-tags](#)

## Tipos de dados

- [Recurso de chave de LF-tags](#)
- [Recurso de política de LF-tags](#)
- [Tabelas com tags](#)
- [Banco de com tags](#)
- [LF-tag](#)
- [Par de LF-tag](#)
- [Erro de LF-tag](#)

- [Coluna com LF-tags](#)

## APIs de filtro de dados

As APIs de filtro de dados descrevem como gerenciar filtros de células de dados em AWS Lake Formation.

### Operações

- [CreateDataCellsFilter](#)
- [DeleteDataCellsFilter](#)
- [ListDataCellsFilter](#)
- [GetDataCellsFilter](#)
- [UpdateDataCellsFilter](#)

### Tipos de dados

- [DataCellsFilter](#)
- [RowFilter](#)

## Tipos de dados comuns

Tipos de dados comuns descrevem os diversos tipos de dados comuns no AWS Lake Formation.

### Estrutura ErrorDetail

Contém detalhes sobre um erro.

#### Campos

- **ErrorCode** – String UTF-8, superior a 1 e inferior a 255 bytes de comprimento, correspondente a [Single-line string pattern](#).  
O código associado a este erro.
- **ErrorMessage** – String de descrição, inferior a 2048 bytes de comprimento, correspondente a [URI address multi-line string pattern](#).

Uma mensagem descrevendo o erro.

## Padrões de string

A API usa as seguintes expressões regulares para definir o que é conteúdo válido para vários membros e parâmetros de string:

- Single-line string pattern – "[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]\*"
- Padrão de string com várias linhas de endereço URI – "[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]\*"
- Padrão de string personalizado N.º. 3: "^w+\.w+\.w+\$"
- Padrão de string personalizado N.º. 4: "^w+\.w+\$"
- Padrão de string personalizado N.º. 5: "arn:aws:iam::[0-9]\*:role/.\*"
  - Padrão de string personalizado N.º. 6: "arn:aws:iam::[0-9]\*:user/.\*"
    - Padrão de string personalizado N.º. 7: "arn:aws:iam::[0-9]\*:group/.\*"
      - Padrão de string personalizado N.º. 8 – "arn:aws:iam::[0-9]\*:saml-provider/.\*"
        - Padrão de string personalizado N.º. 9 – "^( [\p{L}\p{Z}\p{N}\_.\:/=+\-@%]\* )\$"
          - Padrão de string personalizado N.º. 10: "^( [\p{L}\p{Z}\p{N}\_.\:/=+\-@%]\* )\$"
            - Padrão de string personalizado N.º. 11: "[\p{L}\p{N}\p{P}]\*"

## Regiões compatíveis

Esta seção contém informações sobre o suporte Regiões da AWS e a funcionalidade do Lake Formation.

### Disponibilidade geral

Para obter o Regiões da AWS suporte de AWS Lake Formation, consulte [Lista de AWS serviços disponíveis por região](#).

Para obter uma lista dos endpoints de serviço do Lake Formation para cada região e as cotas de serviço do Lake Formation, consulte [endpoints e cotas AWS Lake Formation](#).

### AWS GovCloud (US)

Para uma visão geral das diferenças entre AWS GovCloud (US) região e padrão Regiões da AWS, consulte [Como AWS Lake Formation difere para AWS GovCloud \(US\)](#).

## Otimização de transações e armazenamento

As tabelas controladas, o suporte a transações e os recursos de otimização de armazenamento do Lake Formation estão disponíveis no seguinte: Regiões da AWS

Nome da região	Parâmetro da região	Endpoint
Leste dos EUA (Norte da Virgínia)	us-east-1	lakeformation.us-east-1.amazonaws.com lakeformation-fips.us-east-1.amazonaws.com
Leste dos EUA (Ohio)	us-east-2	lakeformation.us-east-2.amazonaws.com lakeformation-fips.us-east-2.amazonaws.com

Nome da região	Parâmetro da região	Endpoint
Oeste dos EUA (Oregon)	us-west-2	lakeformation.us-west-2.amazonaws.com  lakeformation-fips.us-west-2.amazonaws.com
Ásia-Pacífico (Mumbai)	ap-south-1	lakeformation.ap-south-1.amazonaws.com
Ásia-Pacífico (Seul)	ap-northeast-2	lakeformation.ap-northeast-2.amazonaws.com
Ásia-Pacífico (Singapura)	ap-southeast-1	lakeformation.ap-southeast-1.amazonaws.com
Ásia-Pacífico (Sydney)	ap-southeast-2	lakeformation.ap-southeast-2.amazonaws.com
Ásia-Pacífico (Tóquio)	ap-northeast-1	lakeformation.ap-northeast-1.amazonaws.com
Europa (Frankfurt)	eu-central-1	lakeformation.eu-central-1.amazonaws.com
Europa (Irlanda)	eu-west-1	lakeformation.eu-west-1.amazonaws.com
Europa (Londres)	eu-west-2	lakeformation.eu-west-2.amazonaws.com
Europa (Estocolmo)	eu-north-1	lakeformation.eu-north-1.amazonaws.com

Nome da região	Parâmetro da região	Endpoint
Canadá (Central)	ca-central-1	lakeformation.ca-central-1.amazonaws.com
América do Sul (São Paulo)	sa-east-1	lakeformation.sa-east-1.amazonaws.com



# Histórico da documentação do AWS Lake Formation

A tabela a seguir descreve alterações importantes feitas na documentação do AWS Lake Formation.

Alteração	Descrição	Data
<a href="#">Configuração atualizada do Lake Formation</a>	As etapas na AWS Lake Formation seção Configuração foram <a href="#">atualizadas</a> .	7 de fevereiro de 2024
<a href="#">Alteração de política atualizada</a>	Foram adicionadas novas permissões à política embutida da função vinculada ao serviço. Para obter mais informações, consulte <a href="#">Usando funções vinculadas a serviços para Lake Formation</a> .	7 de fevereiro de 2024
<a href="#">Alteração de política atualizada</a>	Documentou a mudança na <a href="#">LakeFormationDataAccessServiceRolePolicy</a> política.	2 de fevereiro de 2022
<a href="#">Limitações consolidadas do Lake Formation</a>	Foi criada uma seção unificada das limitações e considerações do Lake Formation. Para obter mais informações, consulte <a href="#">Lake Formation limitations</a> .	15 de dezembro de 2023
<a href="#">Foi adicionada documentação sobre compactação do Iceberg.</a>	Para obter melhor performance de leitura por serviços de análise da AWS, como Athena e Amazon EMR e trabalhos de ETL do AWS Glue, o AWS Glue Data Catalog oferece compactação gerenciada	25 de novembro de 2023

(um processo que compacta pequenos objetos do Amazon S3 em objetos maiores) para tabelas do Iceberg no catálogo de dados. Para obter mais informações, consulte [Como otimizar tabelas do Iceberg](#).

[Foi adicionada documentação sobre a integração com o Centro de Identidade do IAM.](#)

As integrações com o Centro de Identidade do IAM permitem que usuários e grupos acessem os recursos do catálogo de dados aplicando as permissões do Lake Formation. Para obter mais informações, consulte [IAM Identity Center integration](#).

25 de novembro de 2023

[Foi adicionada documentação sobre visualizações do catálogo de dados.](#)

É possível criar uma visualização no AWS Glue Data Catalog que faça referência a até dez tabelas usando editores SQL para Amazon Athena ou Amazon Redshift. Para obter mais informações, consulte [Creating views](#).

25 de novembro de 2023

[Atualização da mudança de política](#)

Documentou a mudança na [AWSLakeFormationCrossAccountManager](#) política.

25 de outubro de 2023

[Adicionada documentação do modo de acesso híbrido](#)

O modo de acesso híbrido oferece a flexibilidade de habilitar seletivamente as permissões do Lake Formation para bancos de dados e tabelas em seu AWS Glue Data Catalog. Com o modo de acesso híbrido, agora você tem um caminho incremental que permite definir permissões do Lake Formation para um conjunto específico de usuários sem interromper as políticas de permissão de outros usuários ou workload existentes. Para obter mais informações, consulte [Modo de acesso híbrido](#).

26 de setembro de 2023

[Adicionada documentação para a criação de tabelas do Apache Iceberg](#)

Agora você pode criar tabelas do Apache Iceberg que usam o formato de dados do Apache Parquet no AWS Glue Data Catalog com dados residentes no Amazon S3. Para obter mais informações, consulte [Como criar tabelas do Apache Iceberg](#).

16 de agosto de 2023

[Adicionada documentação para acesso a dados entre regiões](#)

O Lake Formation suporta a consulta de tabelas do Catálogo de Dados em todas as regiões do AWS. Você pode acessar dados em uma região de outras regiões usando Athena, Amazon EMR e executar AWS Glue ETL criando links de recursos em outras regiões apontando para os bancos de dados e tabelas de origem. Você pode conectar o Catálogo de Dados a repositórios externos que armazenam metadados para seus dados do Amazon S3 e gerenciar com segurança as permissões de acesso aos dados usando AWS Lake Formation. Para obter mais informações, consulte [Acesso a tabelas entre regiões](#).

30 de junho de 2023

[Reorganização do conteúdo](#)

Capítulos reorganizados no guia para acompanhar a jornada do usuário de Lake Formation.

15 de maio de 2023

[Adicionada a documentação da federação do HMS](#)

Você pode conectar o Catálogo de Dados a repositórios externos que armazenam metadados para seus dados do Amazon S3 e gerenciar com segurança as permissões de acesso aos dados usando AWS Lake Formation. Para obter mais informações, consulte [Gerenciamento de permissões em conjuntos de dados que usam repositórios externos](#).

15 de abril de 2020

[Adicionada a documentação da unidade de compartilhamento de dados do Amazon Redshift](#)

Agora você pode gerenciar dados com segurança em uma unidade de compartilhamento de dados do Amazon Redshift usando as permissões do Lake Formation. O Lake Formation oferece suporte ao licenciamento de acesso aos seus dados por meio do AWS Data Exchange. Para obter mais informações, consulte [Compartilhamento de dados no AWS Lake Formation](#).

30 de novembro de 2022

[Suporte para compartilhamento de dados entre contas diretamente com as entidades principais](#)

Adicionadas informações sobre a unidade de compartilhamento de dados diretamente com as entidades principais do IAM em outra conta. Para obter mais informações, consulte [Compartilhamento de dados entre contas no AWS Lake Formation](#).

10 de novembro de 2022

[Suporte para compartilhamento de dados habilitado para AWS RAM usando TBAC](#)

Adicionadas informações sobre o método LF-TBAC de conceder o uso de permissões do Catálogo de Dados para concessões entre contas. AWS Resource Access Manager

10 de novembro de 2022

[Adicionada uma seção sobre como trabalhar com outros serviços](#)

Adicionadas informações sobre como serviços do AWS como Athena, AWS Glue, Redshift Spectrum e Amazon EMR podem usar o Lake Formation para acessar com segurança dados em locais do Amazon S3 registrados no Lake Formation. Para obter mais informações, consulte [Trabalhando com outros AWS serviços](#).

10 de novembro de 2022

[???](#)

Adicionadas informações sobre como solucionar um erro ao usar o Amazon EMR para acessar dados entre contas. Para ter mais informações, consulte [Erro ao usar o Amazon EMR para acessar dados compartilhados por meio de várias contas](#).

7 de novembro de 2022

[Atualizações no compartilhamento de recursos entre contas](#)

Adicionada uma descrição de como os [compartilhamentos de recursos entre contas](#) funcionam no Lake Formation . Documentou a mudança na [AWSLakeFormationCrossAccountManager](#) política.

6 de maio de 2022

[Novos tutoriais](#)

Adicionados novos tutoriais para criar tabelas controladas, proteger data lakes e compartilhar data lakes. Para obter mais informações, consulte a seção [Começar](#).

20 de abril de 2022

[Nova página inicial do Lake Formation](#)

A página inicial do [Lake Formation](#) foi atualizada para incluir links para tutoriais que fornecem step-by-step instruções sobre como criar um data lake, ingerir dados, compartilhar e proteger data lakes usando o Lake Formation.

20 de abril de 2022

[Suporte para fornecimento de credenciais](#)

Adicionadas informações sobre o fornecimento de credenciais, que oferece suporte ao Lake Formation para permitir que serviços de terceiros integrem-se ao Lake Formation usando operações de API de fornecimento automático de credenciais. Para obter mais informações, consulte [Como funciona o fornecimento de credenciais no Lake Formation](#).

28 de fevereiro de 2022

[Suporte para tabelas controladas e filtragem avançada de dados](#)

Foram adicionadas informações sobre tabelas governadas, que suportam transações ACID, compactação automática de dados e consultas de viagem no tempo. Adicionadas informações sobre a criação de filtros de dados oferecem suporte à segurança no nível da coluna, segurança no nível da linha e segurança no nível da célula. Para obter mais informações, consulte [Tabelas governadas no Lake Formation](#) e [Filtragem de dados e segurança no nível da célula no Lake Formation](#).

30 de novembro de 2021



[Suporte para endpoints de interface VPC](#)

Adicionadas informações sobre a criação de um endpoint de interface de nuvem privada virtual (VPC) para o Lake Formation, para que a comunicação entre sua VPC e o Lake Formation seja conduzida de forma completa e segura na rede AWS. Para obter mais informações, consulte [Como usar o Lake Formation com endpoints da VPC](#).

11 de outubro de 2021

[Suporte para políticas de endpoint da VPC](#)

Foram adicionadas informações sobre o suporte a políticas de endpoint do nuvem privada virtual (VPC) em Lake Formation. Para obter mais informações, consulte [Como usar o Lake Formation com endpoints da VPC](#).

11 de outubro de 2021

[Suporte para controle de acesso baseado em tags](#)

O controle de acesso baseado em tags do Lake Formation fornece uma maneira nova e mais escalável de gerenciar o acesso aos recursos do Catálogo de Dados e aos dados subjacentes usando LF-tags. Para obter mais informações, consulte [Controle de acesso baseado em tags do Lake Formation](#).

7 de maio de 2021

[Novo requisito de aceitação para filtragem de dados no Amazon EMR.](#)

Adicionadas informações sobre a exigência de se cadastrar para permitir que o Amazon EMR filtre dados gerenciados pelo Lake Formation. Para obter mais informações, consulte [Permissão de filtragem de dados no Amazon EMR.](#)

9 de outubro de 2020

[Suporte para conceder permissões completas entre contas nos bancos de dados do Catálogo de Dados](#)

Adicionadas informações sobre a concessão de permissões completas do Lake Formation nos bancos de dados do Catálogo de Dados em todas as contas do AWS, incluindo CREATE\_TABLE . Para obter mais informações, consulte [Compartilhamento de bancos de dados do Data Catalog.](#)

1º de outubro de 2020

[Suporte para autenticação de usuários do Amazon Athena por meio de SAML.](#)

Adicionadas informações sobre o suporte para usuários do Athena que se conectam por meio do driver JDBC ou ODBC, e se autenticam por meio de provedores de identidade SAML, como Okta e Microsoft Active Directory Federation Service (AD FS). Para obter mais informações, consulte [Integrações do serviço AWS com o Lake Formation.](#)

30 de setembro de 2020

[Suporte para acesso entre contas com um Catálogo de Dados criptografado](#)

Adicionadas informações sobre a concessão de permissões entre contas quando o Catálogo de Dados é criptografado. Para obter mais informações, consulte [Requisitos para Acesso Entre Contas](#).

30 de julho de 2020

[Suporte para acesso entre contas ao data lake](#)

Foram adicionadas informações sobre a concessão de permissões AWS Lake Formation em bancos de dados e tabelas do Catálogo de Dados para contas e organizações externas ao AWS, e sobre o acesso a objetos do Catálogo de Dados compartilhados de contas externas. Para obter mais informações, consulte [Acesso entre contas](#).

7 de julho de 2020

[Integração com a Amazon QuickSight](#)

Foram adicionadas informações sobre como conceder permissões do Lake Formation aos usuários do Amazon QuickSight Enterprise Edition para que eles possam acessar conjuntos de dados residentes em locais registrados do Amazon S3. Para obter mais informações, consulte [Concessão de permissões do catálogo de dados](#).

29 de junho de 2020

[Atualizações nos capítulos de configuração e introdução](#)

Reorganizou e aprimorou os capítulos Configuração e Introdução. Atualizou as permissões recomendadas do AWS Identity and Access Management (IAM) para o administrador do data lake.

27 de fevereiro de 2020

[Suporte para AWS Key Management Service](#)

Adicionadas informações sobre como o suporte do Lake Formation para AWS Key Management Service (AWS KMS) simplifica a configuração de serviços integrados para ler e gravar dados criptografados em locais registrados do Amazon Simple Storage Service (Amazon S3). Adicionadas informações sobre o registro de locais do Amazon S3 criptografados com o AWS KMS keys. Para ter mais informações, consulte [the section called “Adicionar uma localização do Amazon S3 ao seu data lake”](#).

27 de fevereiro de 2020

[Atualizações nos esquemas e nas políticas de IAM do administrador do data lake](#)

Parâmetros de entrada esclarecidos para esquemas de banco de dados incrementais. Atualizou as políticas do IAM necessárias para um administrador de data lake.

20 de dezembro de 2019

[Reescrito o capítulo de segurança e revisões do capítulo de atualização](#)

Melhoria dos capítulos de segurança e atualização.

29 de outubro de 2019

[A super permissão substituiu todas as permissões](#)

Os capítulos de Segurança e Atualização foram atualizados para refletir a substituição da permissão All por Super.

10 de outubro de 2019

[Adições, correções e esclarecimentos](#)

Foram feitas adições, correções e esclarecimentos com base nas opiniões. Revisado o capítulo sobre segurança. Os capítulos de Segurança e Atualização foram atualizados para refletir a substituição do grupo Everyone por IAMAllowedPrincipals .

11 de setembro de 2019

[Novo guia](#)

Este é o lançamento inicial do Guia do desenvolvedor do AWS Lake Formation.

8 de agosto de 2019

# Glossário do AWS

Para obter a terminologia mais recente da AWS, consulte o [glossário da AWS](#) na Referência do Glossário da AWS.

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.