

LakeFormation

User Guide

Issue 01
Date 2024-02-01



Copyright © Huawei Cloud Computing Technologies Co., Ltd. 2024. All rights reserved.

No part of this document may be reproduced or transmitted in any form or by any means without prior written consent of Huawei Cloud Computing Technologies Co., Ltd.

Trademarks and Permissions



HUAWEI and other Huawei trademarks are the property of Huawei Technologies Co., Ltd.

All other trademarks and trade names mentioned in this document are the property of their respective holders.

Notice

The purchased products, services and features are stipulated by the contract made between Huawei Cloud and the customer. All or part of the products, services and features described in this document may not be within the purchase scope or the usage scope. Unless otherwise specified in the contract, all statements, information, and recommendations in this document are provided "AS IS" without warranties, guarantees or representations of any kind, either express or implied.

The information in this document is subject to change without notice. Every effort has been made in the preparation of this document to ensure accuracy of the contents, but all statements, information, and recommendations in this document do not constitute a warranty of any kind, express or implied.

Contents

1 Preparations.....	1
1.1 Registering a Huawei Cloud Account.....	1
1.2 Configuring Cloud Service Authorization.....	1
1.3 Managing Permissions.....	3
1.3.1 Creating a User and Assigning Permissions.....	3
1.3.2 Creating a Custom Policy.....	8
2 Instance Management.....	11
2.1 Creating an Instance.....	11
2.2 Configuring an Instance.....	13
2.3 Deleting an Instance.....	14
3 Metadata Management.....	15
3.1 Creating a Metadata Storage Path.....	15
3.2 Managing Catalogs.....	15
3.3 Managing Databases.....	17
3.4 Managing Tables.....	20
3.5 Managing Functions.....	24
3.6 Configuring Metadata Lifecycle.....	25
4 Data Permission Management.....	28
4.1 Data Permission Overview.....	28
4.2 Granting permissions.....	30
4.3 Canceling Authorization.....	33
4.4 Querying Authorization.....	33
4.5 Performing Role-based Authorization.....	34
5 Data Migration Management.....	36
5.1 Granting the Job Management Permission.....	36
5.2 Migrating Metadata.....	37
5.3 Migrating Permissions.....	41
6 Managing Clients.....	44
7 Viewing Audit Logs.....	46
8 Change History.....	48

1 Preparations

1.1 Registering a Huawei Cloud Account

Register a Huawei Cloud account to access all services on Huawei Cloud and pay only for the services you use.

If you already have a Huawei Cloud account, skip this part.

Procedure

- Step 1** Visit the [Huawei Cloud website](#).
- Step 2** Click **Register** and complete the registration as instructed. For details, see [Registering a Huawei ID and Enabling Huawei Cloud Services](#).
The system redirects you to the personal information page.
- Step 3** Complete real-name authentication for individual or enterprise accounts. For details, see [Real-Name Authentication](#).

----End

1.2 Configuring Cloud Service Authorization

Before using LakeFormation, you need to authorize it to access related cloud services.

- **IAM ReadOnlyAccess:** allows your instance to obtain user group and user information when running.
- **OBS OperateAccess:** provides the storage function for your instance.
- **OBS AccessLabel:** allows your instance to control permissions by using tagging.
- **OBS Bucket Lifecycle:** This permission is required to manage the lifecycle of instances.
- **VPC Endpoint Administrator:** allows you to operate on VPC endpoint for instance access management.

- **DNS FullAccess:** allows you to modify DNS private domain names for instance access management.

Procedure

Step 1 Log in to the LakeFormation console.

Step 2 Click the service list and choose **Analytics > LakeFormation**. On the displayed page, click **Service Authorization** the left navigation pane.

Figure 1-1 Authorization

Service Authorization

To use LakeFormation, grant the following permissions:

- IAM ReadOnlyAccess**
Allows you to obtain the user group and user information during instance running.
- OBS OperateAccess**
Allows you to store instances in OBS.
- OBS AccessLabel**
Allows you to control instance permissions through labelling.
- VPC Endpoint Administrator**
Allows you to operate on VPC endpoint for instance access management.
- DNS FullAccess**
Allows you to modify DNS private domain names for instance access management.

After agreeing to the authorization, LakeFormation will create an agency named `lakeformation_admin_trust` on IAM. To ensure the normal running of LakeFormation, do not delete this agency.

I have read and agree with the [LakeFormation Service Statement](#).

Authorize

Step 3 Select "I have read and agree with the LakeFormation Service Statement." and click **Authorized**.

NOTE

After cloud service authorization, LakeFormation will create an agency named **lakeformation_admin_trust** in Identity and Access Management (IAM). Do not delete the agency when using LakeFormation.

If the agency fails to be automatically created, log in to the IAM console, delete the agency or contact the administrator to increase the quota, and check whether the current user has the permission to create an agency.

----End

1.3 Managing Permissions

1.3.1 Creating a User and Assigning Permissions

This chapter describes how to use [What Is IAM?](#) to implement fine-grained permissions control for your LakeFormation instances. With IAM, you can:

- Create IAM users for employees based on your enterprise's organizational structure. Each IAM user will have their own security credentials for accessing the resources.
- Assign users only the permissions required to perform a given task based on their job responsibilities.
- Entrust an account or a cloud service to perform professional and efficient O&M on your resources.

If your cloud account does not need individual IAM users, skip this section.

This section describes the procedure for granting permissions. See [Figure 1-2](#) for details.

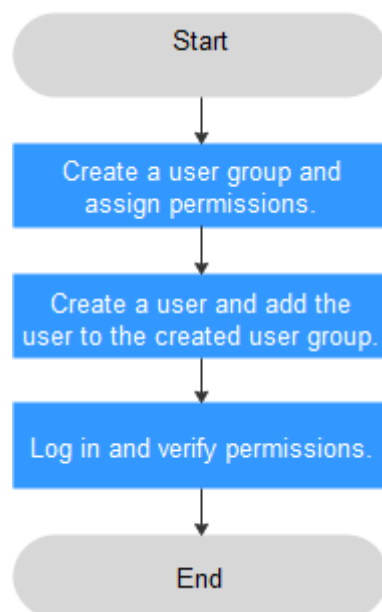
Prerequisites

Before assigning permissions to a user group, learn about LakeFormation permissions in [LakeFormation Permissions](#) and select them as needed.

If you need to assign permissions to other services, check all the policies supported by IAM in [System-defined Permissions](#).

Procedure

Figure 1-2 Granting LakeFormation permissions



1. **Creating a User Group and Assigning Permissions**
Create a user group on the IAM console, and assign LakeFormation permissions to the group.
2. **Create an IAM user.**
Create a user on the IAM console and add the user to the group created in 1.
3. **Log in** and verify permissions.
Log in to the console as the user you created, and verify that the user has the assigned permissions.
You can perform the following operation to verify a permission:
Click the service list and choose LakeFormation. On the **Overview** page, click **Buy Instance** in the upper right corner. If the instance creation page is displayed, the **lakeformation:role:create permission** has already taken effect.

LakeFormation Permissions

By default, new IAM users lack permissions assigned. Add a user to one or more groups, and attach permissions policies or roles to these groups. Users inherit permissions from the groups to which they are added and can perform specified operations on cloud services.

You can grant permissions to a role or by creating a policy.

- **Roles:** A coarse-grained IAM authorization strategy to assign permissions based on user responsibilities. Only a limited number of service-level roles are available. Some roles depend other roles to take effect. When you assign such roles to users, remember to assign the roles they depend on. Roles are not an ideal choice for fine-grained authorization and secure access control.

NOTE

- **Read-only permission authorization using IAM:** To assign the read-only permission of LakeFormation in an IAM project to a sub-user, create a user group for this user as a tenant administrator and add the LakeFormation ReadOnlyAccess system-defined policy to this user group.
- **Enterprise project authorization:** To assign all permission of LakeFormation in an enterprise project to a sub-user, create a user group for this user as a tenant administrator and add the LakeFormation CommonAccess permission to the user group and make this permission apply globally. Then, grant this permission to this enterprise project.
- **Policies:** A type of fine-grained authorization that defines permissions required to perform operations on specific cloud resources under certain conditions. This type of authorization is more flexible and ideal for secure access control. Most fine-grained policies split permissions by API. For details about how to customize IAM policies for LakeFormation, see [Creating a Custom Policy](#).

Table 1-1 LakeFormation system permissions

Role/Policy Name	Description	Type	Dependency
LakeFormation FullAccess	Administrator permissions for LakeFormation. Users granted these permissions can use all LakeFormation functions.	System policy	N/A
LakeFormation ReadOnlyAccess	Read-only permissions for LakeFormation. Users granted these permissions can query LakeFormation data.	System policy	N/A
LakeFormation CommonAccess	Basic permissions for LakeFormation, including viewing, authorizing, and canceling the LakeFormation service agreement and basic permissions for dependent services such as OBS and TMS.	System policy	N/A

Table 1-2 LakeFormation fine-grained permissions

Operation Type	Item	Description
Read-only	lakeformation:instance:describe	Permission to query LakeFormation instances.
	lakeformation:catalog:describe	Permission to query the data catalogs of LakeFormation metadata.
	lakeformation:database:describe	Permission to query the databases of LakeFormation metadata.
	lakeformation:table:describe	Permission to query the data tables of LakeFormation metadata.
	lakeformation:function:describe	Permission to query the functions of LakeFormation metadata.
	lakeformation:policy:describe	Permission to query LakeFormation permission policies.
	lakeformation:policy:export	Permission to query LakeFormation permission policies in batches.

Operation Type	Item	Description
	lakeformation:agency:describe	Permission to query LakeFormation agencies.
	lakeformation:credential:describe	Permission to obtain the authentication information for accessing LakeFormation.
	lakeformation:group:describe	Permission to obtain the relationship between a LakeFormation user group and its associated roles.
	lakeformation:user:describe	Permission to obtain the relationship between a LakeFormation user and its associated roles.
	lakeformation:role:describe	Permission to query LakeFormation roles.
	lakeformation:configuration:describe	Permission to query user configurations.
	lakeformation:access:describe	Permission to query the client access permission.
	lakeformation:job:describe	Permission to query LakeFormation tasks.
Write	lakeformation:instance:create	Permission to create LakeFormation instances.
	lakeformation:role:create	Permission to create LakeFormation roles.
	lakeformation:policy:create	Permission to create LakeFormation permission policies.
	lakeformation:function:create	Permission to create the functions of LakeFormation metadata.
	lakeformation:catalog:create	Permission to create the data catalogs of LakeFormation metadata.
	lakeformation:database:create	Permission to create the databases of LakeFormation metadata.
	lakeformation:table:create	Permission to create the tables of LakeFormation metadata.

Operation Type	Item	Description
	lakeformation:access:create	Permission to create the client access permission.
	lakeformation:agency:create	Permission to create LakeFormation agencies.
	lakeformation:job:create	Permission to create LakeFormation tasks.
	lakeformation:instance:alter	Permission to modify LakeFormation instances.
	lakeformation:catalog:alter	Permission to modify the data catalogs of LakeFormation metadata.
	lakeformation:database:alter	Permission to modify the databases of LakeFormation metadata.
	lakeformation:table:alter	Permission to modify the tables of LakeFormation metadata.
	lakeformation:function:alter	Permission to modify the functions of LakeFormation metadata.
	lakeformation:role:alter	Permission to modify the relationship between a LakeFormation role and its associated user groups.
	lakeformation:group:alter	Permission to modify the relationship between a LakeFormation user group and its associated roles.
	lakeformation:user:alter	Permission to modify the relationship between a LakeFormation user and its associated roles.
	lakeformation:job:alter	Permission to modify LakeFormation tasks.
	lakeformation:instance:drop	Permission to delete LakeFormation instances.
	lakeformation:role:drop	Permission to delete LakeFormation roles.
	lakeformation:policy:drop	Permission to delete LakeFormation permission policies.

Operation Type	Item	Description
	lakeformation:function:drop	Permission to delete the functions of LakeFormation metadata.
	lakeformation:catalog:drop	Permission to delete the data catalogs of LakeFormation metadata.
	lakeformation:database:drop	Permission to delete the databases of LakeFormation metadata.
	lakeformation:table:drop	Permission to delete the tables of LakeFormation metadata.
	lakeformation:access:delete	Permission to delete the client access permission.
	lakeformation:agency:drop	Permission to delete LakeFormation agencies.
	lakeformation:job:drop	Permission to delete LakeFormation tasks.
	lakeformation:transaction:operate	Permission to operate LakeFormation transactions.
	lakeformation:instance:access	Permission to query a LakeFormation instance or apply for the access to it.
	lakeformation:job:exec	Permission to execute LakeFormation tasks.

1.3.2 Creating a Custom Policy

Custom policies can be created for LakeFormation to supplement system-defined policies.

To create a custom policy, choose either visual editor or JSON.

- Visual editor: Select cloud services, actions, resources, and request conditions. This does not require knowledge of policy syntax.
- JSON: Create a policy in the JSON format from scratch or based on an existing policy.

For details, see [Creating a Custom Policy](#).

The following section contains examples of common LakeFormation custom policies.

Example Custom Policies

- Example 1: Grant the read-only permission on LakeFormation in batches to a user.

```
{
  "Version": "1.1",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "lakeformation:instance:describe",
        "lakeformation:role:describe",
        "lakeformation:policy:export",
        "lakeformation:group:describe",
        "lakeformation:function:describe",
        "lakeformation:catalog:describe",
        "lakeformation:policy:describe",
        "lakeformation:table:describe",
        "lakeformation:database:describe"
      ]
    }
  ]
}
```

- Example 2: Deny data deletion.

A policy with only "Deny" permissions must be used together with other policies. If the permissions granted to an IAM user contain both "Allow" and "Deny", the "Deny" permissions take precedence over the "Allow" permissions.

The following method can be used if you need to assign permissions of the **Admin** policy to a user but you want to prevent the user from deleting LakeFormation catalogs, databases, and tables. Create a custom policy for denying deletion, and attach both policies to the group to which the user belongs. Then, the user can perform all operations on catalogs, databases, and tables except deleting them.

Example policy denying MRS cluster deletion:

```
{
  "Version": "1.1",
  "Statement": [
    {
      "Effect": "Deny",
      "Action": [
        "lakeformation:database:drop",
        "lakeformation:table:drop",
        "lakeformation:catalog:drop"
      ]
    }
  ]
}
```

- Example 3: Create a custom policy that contains the actions of multiple services.

```
{
  "Version": "1.1",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "obs:bucket:CreateBucket"
      ]
    },
    {
      "Effect": "Allow",
      "Action": [
        "lakeformation:table:create",
        "lakeformation:database:create",

```

```
    "lakeformation:catalog:create"  
  }  
]  
}
```


2 Instance Management

2.1 Creating an Instance

Before using LakeFormation, create an instance. Then, you can perform operations on the instance, such as managing metadata and setting metadata permissions.

Procedure

Step 1 Log in to the LakeFormation console.

Step 2 In the upper left corner, click  and choose **Analytics > LakeFormation** to access the LakeFormation console.

Step 3 Click **Buy Now** or **Buy Instance** in the upper right corner.

If a LakeFormation instance exists on the page, **Buy Instance** is displayed. Otherwise, **Buy Now** is displayed.

Step 4 Configure the parameters listed in the following table.

Table 2-1 Instance creation parameters

Parameter	Description
Type	Select the instance type. <ul style="list-style-type: none">• Shared: Resources are reused between shared instances to maximize the usage of resources such as clusters or GaussDB(for MySQL) instances.• Exclusive: You are charged based on the upper limit of query per second (QPS) and metadata usage.
Region	Select a region . LakeFormation instances in different regions cannot communicate with each other over an intranet. For lower network latency and quick resource access, select the nearest region.

Parameter	Description
Billing Mode	Billing mode of instances. <ul style="list-style-type: none"> • Pay-per-use: You are billed by the usage duration of LakeFormation instances.
Project	Select the project to which the instance belongs.
Name	User-defined LakeFormation instance name.
QPS	Maximum number of requests per second. You do not need to set this parameter when Type is set to Shared . You will be billed by LakeFormation based on your usage of metadata objects.
Enterprise Project	Enterprise project to which the cluster belongs. If no enterprise project is available, click Create to create one. An enterprise project facilitates project-level management and grouping of cloud resources and users.
Description	Description of the instance.
Label	Enter content in the boxes and click Add to add a tag. You are advised to click View Predefined Tags to use predefined tags and add the same tag to different cloud resources.

Step 5 Click **Buy Now**, confirm the configured information, and pay.

Step 6 Click **Back to Console**. You can view information about the newly created LakeFormation instance on the console.

 **NOTE**

When creating an instance, pay attention to quota notification. If a resource quota is insufficient, increase the resource quota as prompted and create an instance.

Wait until the instance status changes to  **Running**, indicating that the instance is created.

After an instance is created, you can view its basic information and data overview.

----End

Failed to Create an Instance

If an instance fails to be created, the **Failed** page is displayed.


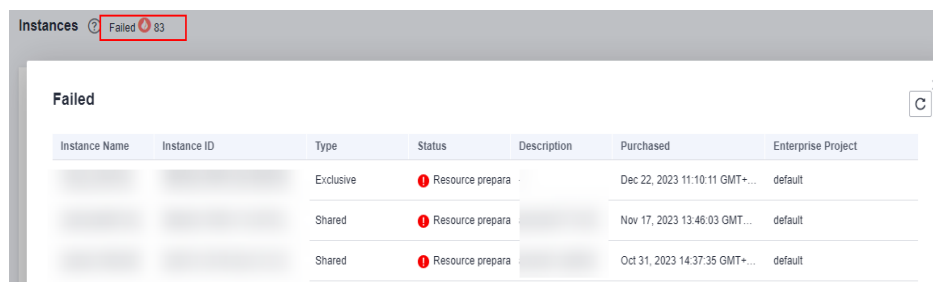
Choose **Analytics > LakeFormation**. In the upper left corner of the **Overview** page, click **Failed** . In the displayed window, view information about the instance that fails to be created.

Figure 2-1 Failed instances




Instance Name	Instance ID	Type	Status	Description	Purchased	Enterprise Project
		Exclusive	Resource preparation failed		Dec 22, 2023 11:10:11 GMT+...	default
		Shared	Resource preparation failed		Nov 17, 2023 13:46:03 GMT...	default
		Shared	Resource preparation failed		Oct 31, 2023 14:37:35 GMT+...	default

2.2 Configuring an Instance


After a LakeFormation instance is created, you can perform the following operations on the instance on the **Overview** page.

- **Modifying Instance Specifications:** Change the QPS value of the instance. Only exclusive instances support this operation.
- **Setting an Instance as Default:** Set the instance as the default one. If no instance ID is specified when other services interconnect with the LakeFormation instance, this operation will modify the instance interconnected with the service.

Modifying Instance Specifications

- Step 1** Log in to the LakeFormation console.
 - Step 2** In the upper left corner, click  and choose **Analytics > LakeFormation** to access the LakeFormation console.
 - Step 3** Select the target LakeFormation instance from the drop-down list box on the left.
 - Step 4** Click **Scale Instance** in the upper right corner of the page. The **Scale Instance** page is displayed. Shared instances do not support this operation.
 - Step 5** Select the target QPS value and click **Next**.
 - Step 6** Confirm the instance information and its modified specification, and click **Submit**.
- End

Setting an Instance as Default

- Step 1** Log in to the LakeFormation console.
- Step 2** In the upper left corner, click  and choose **Analytics > LakeFormation** to access the LakeFormation console.
- Step 3** Select the target LakeFormation instance from the drop-down list box on the left.
- Step 4** Check the value of **Default Instance** in the **Basic Information** area.
 - **true** indicates that the instance is the default instance.
 - **false** indicates that the current instance is not the default instance.

If the current instance is the default instance, the **Set as Default** button is unavailable.

Step 5 To set the instance as the default one, click **Set as Default** in the upper right corner of the page, confirm the operation impact, and click **OK**.


NOTICE

If other services need to connect to LakeFormation and no LakeFormation instance ID is specified, the system automatically accesses the default instance. This default instance change may affect the services that use LakeFormation. Exercise caution when performing this operation.

----End

2.3 Deleting an Instance

Step 1 Log in to the LakeFormation console.

Step 2 In the upper left corner, click  and choose **Analytics > LakeFormation** to access the LakeFormation console.

Step 3 Select the target LakeFormation instance from the drop-down list box on the left.

Step 4 Click **Delete Instance** in the upper right corner.

Step 5 In the displayed dialog box, confirm the deletion impact, and click **OK** to delete the instance.

 **NOTE**

- Deleted instances are moved to the recycle bin and continue to be billed until they are deleted from the recycle bin.
- Instances that have been stored in the recycle bin for more than one day will be automatically deleted and cannot be restored.
- Instances can be forcibly deleted only 15 minutes after they are moved to the recycle bin to prevent service interruption.

----End

3 Metadata Management

3.1 Creating a Metadata Storage Path

The data files and directories mapped to LakeFormation metadata are stored in the OBS parallel file system. Before creating LakeFormation metadata, you need to create an OBS parallel file system for data storage.

If an OBS parallel file system is available, skip this section.

Procedure

- Step 1** Create an OBS parallel file system. For details, see [Creating a Parallel File System](#). For example, the file system name is **lakeformation-test**.
- Step 2** On the **Parallel File Systems** page, click the name of the created file system, that is **lakeformation-test**.
- Step 3** Click **Files** in the navigation pane, click **Create Folder**, enter a folder name, and click **OK**. Click the folder name and click **Create Folder** to create a subfolder.

Repeat this step to create paths for storing metadata in sequence. The following paths are examples:

- Catalog storage path: **lakeformation-test/catalog1**
- Database storage path: **lakeformation-test/catalog1/database1**
- Table storage path: **lakeformation-test/catalog1/database1/table1** and **lakeformation-test/catalog1/database1/table**
- Function storage path: **lakeformation-test/catalog1/database1/udf1**

----End

3.2 Managing Catalogs

A data catalog is a metadata management object that can contain multiple databases.


Multiple catalogs can be created and managed in LakeFormation to isolate metadata of different external clusters.

Prerequisites

- A LakeFormation instance has been created and is running properly.
- Catalog data is stored in OBS and the permission to perform operations on OBS is obtained.
- You have created an OBS parallel file system for storing catalog data by referring to [Creating a Metadata Storage Path](#).

Creating a Catalog

Step 1 Log in to the LakeFormation console.

Step 2 In the upper left corner, click  and choose **Analytics > LakeFormation** to access the LakeFormation console.

Step 3 Select the LakeFormation instance to be operated from the drop-down list on the left and choose **Metadata > Catalog** in the navigation pane.

Step 4 Click **Create** and set the parameters.

1. In the **Basic Information** area, set the related parameters.

Table 3-1 Parameters for creating a catalog

Parameter	Description
Catalog Name	Enter a catalog name. The value should contain 1 to 256 characters. Only letters, numbers, and underscores (_) are allowed.
Catalog Type	The options are as follows: <ul style="list-style-type: none"> - DEFAULT - CLICKHOUSE

Parameter	Description
Select Location	Select a location where catalog data is stored in the OBS parallel file system. Click + , select a location, and click OK . <ul style="list-style-type: none">- The selected location must start with obs:// and must contain one storage object. For example, select obs://lakeformation-test/catalog1. If no suitable parallel file system is available, click go to OBS and create one.- You are advised to select a folder that is not selected by other catalogs.
Description	Enter a description of the created catalog. The content length must be between 0 and 4000 bytes (3 bytes per Chinese character).

2. (Optional) Click **Add** under **Database Storage Locations**. On the **Add Database Storage Location** page, click **+** to manually select a database storage path as required. Click **OK**.

Click **+** to add more storage paths. Click **🗑** to delete a storage path.

NOTE

Selecting database storage path is an optional operation. If a database storage path is added, the databases in this catalog must be stored in the subpath of the database storage path or that of the selected catalog location path.

3. Click **Submit**.

Step 5 After the catalog is created, you can view the catalog information on the **Catalog** page.

Click **Modify** in the **Operation** column to modify the configurations of a catalog.

Click **Database** in the **Operation** column to view the databases in a catalog.

Click **More** to authorize or delete a catalog, or view the permissions of a catalog.

NOTE

If files are deleted when metadata is deleted, the system processes the files based on the deletion policy that you have configured. For how to configure a deletion policy, see [Configuring Metadata Lifecycle](#). If no deletion policy is configured, data is moved to the recycle bin (OBS path **lake-formation-trash-dir/table_id**) by default.

----End

3.3 Managing Databases


Multiple databases can be created under a LakeFormation catalog. Centralized metadata management can maximize the value of data assets.

Prerequisites

- A LakeFormation instance has been created and is running properly.
- A catalog has been created.
- You have created an OBS parallel file system for storing databases by referring to [Creating a Metadata Storage Path](#).

Create a Database

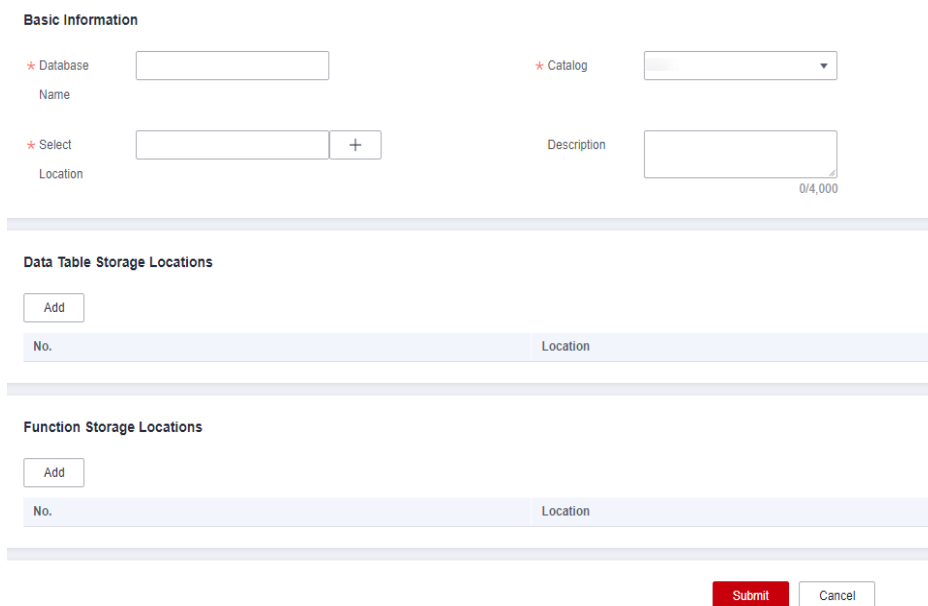
Step 1 Log in to the LakeFormation console.

Step 2 In the upper left corner, click  and choose **Analytics > LakeFormation** to access the LakeFormation console.

Step 3 Select the LakeFormation instance to be operated from the drop-down list on the left and choose **Metadata > Database** in the navigation pane.

Step 4 Select the name of the desired catalog from the **Catalog** drop-down list box in the upper right corner. You can view the databases contained in this catalog.

Step 5 Click **Create**.





1. In the **Basic Information** area, set the related parameters.

Table 3-2 Parameters for creating a database

Parameter	Description
Database Name	Enter a database name. The value should contain 1 to 128 characters. Only letters, numbers, and underscores (_) are allowed.
Catalog	Select the name of the catalog to which the database to be created belongs.

Parameter	Description
Select Location	<p>Select a location where the database information is stored in the OBS parallel file system.</p> <p>Click +, select a location, and click OK.</p> <ul style="list-style-type: none"> - The selected location must start with obs:// and must contain one storage object. For example, select obs://lakeformation-test/catalog1/database1. If no suitable parallel file system is available, click go to OBS and create one. - The path must be different from the storage path of its catalog (the one specified by Select Location on the catalog creation page). - If a database storage path is added for the catalog to which the database belongs, set this parameter to the subpath of the database storage path or that of the path specified by Select Location during catalog creation.
Description	<p>Description of the created database.</p> <p>The content length must be between 0 and 4000 bytes (3 bytes per Chinese character).</p>



2. (Optional) Click **Add** under **Data Table Storage Locations**. On the **Add Table Storage Location** page, click **+** to manually select a database storage path as required. Click **OK**.

Click  to add more storage paths. Click  to delete a storage path.

 **NOTE**

- Selecting data table storage path is an optional operation.
- The data table storage path can be set to the catalog path and its subpaths, or database storage location path and its subpaths.
- If a data table storage path is added, the tables in this database must be the subpath of the data table storage path or that of the selected database location path.

3. (Optional) Click **Add** under **Function Storage Locations**. On the **Add Function Storage Location** page, click **+** to manually select a function storage path as required. Click **OK**.

Click  to add more storage paths. Click  to delete a storage path.

 **NOTE**

- Selecting function storage path is an optional operation.
- The function storage path can be set to the catalog path and its subpaths, or database storage location path and its subpaths.
- If a function storage path is added, the functions in this database must be the subpath of the function storage path or that of the selected database location path.

4. Click **Submit**.

Step 6 After the database is created, you can view the database name/ID, catalog, owner, and storage location on the **Database** page.

Click **Modify** in the **Operation** column to modify the configurations of a database.

Click **Table** in the **Operation** column to view the tables in a database.

Click **More** to authorize or delete a database, or view the permissions of a database.

 **NOTE**

If files are deleted when metadata is deleted, the system processes the files based on the deletion policy that you have configured. For how to configure a deletion policy, see [Configuring Metadata Lifecycle](#). If no deletion policy is configured, data is moved to the recycle bin (OBS path **lake-formation-trash-dir/table_id**) by default.

----End

3.4 Managing Tables


You can manage metadata databases and metadata tables in the data catalog and create data tables based on the service plan.

Prerequisites

- A LakeFormation instance has been created and is running properly.
- A catalog and a database have been created.
- You have created an OBS parallel file system for storing tables by referring to [Creating a Metadata Storage Path](#).

Creating a Table

Step 1 Log in to the LakeFormation console.

Step 2 In the upper left corner, click  and choose **Analytics > LakeFormation** to access the LakeFormation console.

Step 3 Select a LakeFormation instance from the drop-down list box on the left, choose **Metadata > Table**, and select names of the target catalog and database from the **Catalog** and **Database** drop-down lists in the upper right corner. You can view the tables contained in the selected database.

Step 4 Click **Create** and set related parameters.

The screenshot shows the 'Create Table' interface with the following sections:

- Basic Information:** Includes fields for Table Name, Catalog, Database, Table Type, Data Storage Location, Data Source Format, and Description (with a 0M,000 limit).
- Table Field:** A table with columns: Name, Type, Length/Setting, Description, and Operation. It includes an 'Add' button and search filters.
- Partition Key:** A table with columns: Name, Type, Length/Setting, Description, and Operation. It includes an 'Add' button and search filters.
- Table Attributes:** A table with columns: Attribute Name, Attribute Value, and Operation. It includes an 'Add' button and a search filter.

At the bottom, there are 'Submit' and 'Cancel' buttons.







1. In the **Basic Information** area, set the related parameters.

Table 3-3 Parameters for creating a table

Parameter	Description
Table Name	Enter a metadata table name. The value should contain 1 to 256 characters. Only letters, numbers, and underscores (_) are allowed.
Catalog	Select the catalog to which the table to be created belongs.
Database	Select the database to which the table to be created belongs.
Table Type	Type of the table to be created. Currently, the following types are supported: <ul style="list-style-type: none"> – MANAGED_TABLE: management table. If a management table or partition is deleted, the data and metadata associated with the table or partition are deleted. – EXTERNAL_TABLE: external table. An external table is used when the file already exists or is stored in a remote location. – VIRTUAL_VIEW: virtual view. It does not store actual data or occupy physical space. – MATERIALIZED_VIEW: materialized view. It stores actual data and occupies physical space.

Parameter	Description
Data Storage Location	<p>File directory of the OBS parallel file system to which the table is mapped.</p> <p>Click +, select a location for storing the table in the OBS parallel file system, and click OK.</p> <ul style="list-style-type: none"> - The selected location must start with obs:// and must contain one storage object. For example, select obs://lakeformation-test/catalog1/database1/table1. If no suitable parallel file system is available, click go to OBS and create one. - The path must be different from the storage path of the catalog and database to which the belongs. - If a table storage path is added for the database to which the table belongs, set this parameter to the subpath of the table storage path or that of the path specified by Select Location during database creation.
Compress Data	<p>Whether to compress a data table.</p> <p>By using table compression, you can store data in a compressed format and improve performance and storage space.</p>
Data Source Format	<p>Data source format of the table to be created. Currently, the following types are supported:</p> <ul style="list-style-type: none"> - Avro - Json - Xml - Parquet - Csv - Orc - Text - Rc - Sequence - Custom <p>Parameters Input Format, Output Format, Serde name, and SerializationLib are displayed if Data Source Format is set to Custom. Set these parameters based on the site requirements.</p>

Parameter	Description
Separator	This parameter is displayed if Data Source Format is set to Csv . The values include: <ul style="list-style-type: none">- Comma(,)- Vertical bar()- Semicolon(,)- Tab(\u0009)- Ctrl-A(\u0001)
Description	Description of the created table. The content length must be between 0 and 4000 bytes (3 bytes per Chinese character).

- (Optional) Click **Add** in the **Table Field** area. Manually add metadata table fields as required. Click **OK**.
Click  to add more table fields. Click  to delete a table field.
A table field is an independent piece of information that forms a record in a table.
- (Optional) Click **Add** in the **Partition Key** area. Manually add the partition key of the metadata as required. Click **OK**.
Click  to add more partition keys. Click  to delete a partition key.
A partition key is an ordered set of one or more table columns. The values in the table partition keys are used to determine the data partition that a row belongs to.
- (Optional) Click **Add** in the **Table Attributes** area. Add metadata table attributes as required. Click **OK**.
Click  to add more table attributes. Click  to delete a table attribute.
A table attribute enables you to tag table definitions with your own metadata key-value pairs.
- Click **Submit**.

Step 5 After the table is created, you can view the table name/ID, catalog, database, type, and storage location on the **Table** page.

Click **Modify** in the **Operation** column to modify the configurations of a table.

Click **More** and then click **Authorize** in the **Operation** column to authorize a table.

Click **More** to authorize or delete a database, or view the permissions of a table.

NOTE

If files are deleted when metadata is deleted, the system processes the files based on the deletion policy that you have configured. For how to configure a deletion policy, see [Configuring Metadata Lifecycle](#). If no deletion policy is configured, data is moved to the recycle bin (OBS path `lake-formation-trash-dir/table_id`) by default.

Step 6 Click a table name to view its detailed metadata information.

- The format and serialization information includes the storage format, input format, and output format.
- The field information includes the table field name, type, and description, and the field name, type, and description of the partition key.
- The table attribute information includes the name and value of each attribute in the table.

Locate the row that contains the target table and click **Edit** to modify the fields in it.

----End

3.5 Managing Functions


You can manage metadata in the data catalog and create functions based on the service plan.

Prerequisites

- A LakeFormation instance has been created and is running properly.
- A catalog and a database have been created.
- You have created an OBS parallel file system for storing functions by referring to [Creating a Metadata Storage Path](#) if a function location is added.

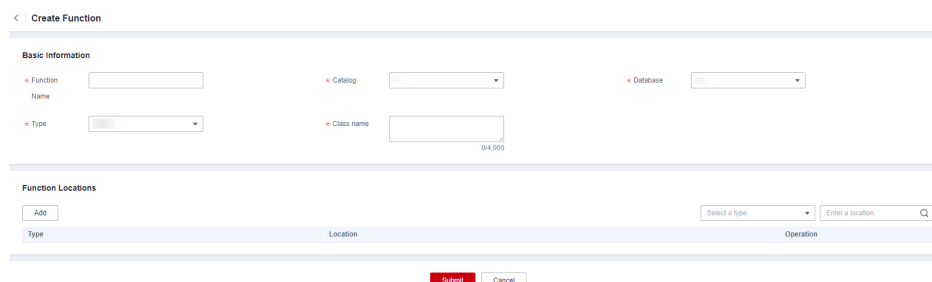
Creating a Function

Step 1 Log in to the LakeFormation console.

Step 2 In the upper left corner, click  and choose **Analytics > LakeFormation** to access the LakeFormation console.

Step 3 Select the target LakeFormation instance from the drop-down list box on the left and choose **Metadata > Function**. In the upper right corner, select the names of the target catalog and database from the **Catalog** and **Database** drop-down lists. You can view the functions contained in the selected database.

Step 4 Click **Create** and set related parameters.





The screenshot shows the 'Create Function' page in the LakeFormation console. It is divided into two main sections: 'Basic Information' and 'Function Locations'.
Basic Information: This section contains several input fields and dropdown menus. There are three dropdown menus for 'Catalog' and 'Database'. There are also text input fields for 'Function Name', 'Type', and 'Class name'.
Function Locations: This section features an 'Add' button, a table with columns for 'Type', 'Location', and 'Operation', and a search bar with the placeholder text 'Enter a location'.
Bottom: At the bottom of the form, there are two buttons: 'Submit' (in red) and 'Cancel'.

1. In the **Basic Information** area, set the related parameters.

Table 3-4 Parameter for creating a function

Parameter	Description
Function Name	Enter a metadata function name. The value should contain 1 to 256 characters. Only letters, numbers, and underscores (_) are allowed.
Catalog	Catalog to which the function to be created belongs.
Database	Database to which the function to be created belongs.
Type	Type of the function to be created. Currently, JAVA is supported.
Class Name	Enter a function class name.

- (Optional) Click **Add** under **Function Locations** to add the function package type and location as required. Click **OK**.

Click  to add more function package types and locations. Click  to delete a function package type and location.

 **NOTE**

- Selecting function package type and location is an optional operation.
 - If a function storage path is added for the database to which the function belongs, the storage location must be the path specified by **Select Location** during database creation or its subpaths, or the function storage path or its subpaths.
- Click **Submit**.

Step 5 After the function is created, you can view the function name/ID, catalog, database, type, and class name on the **Function** page.

Click **Modify** in the **Operation** column to modify the configurations of a function.

Click **More** and then click **Authorize** in the **Operation** column to authorize a function.

Click **More** to authorize or delete a function, or view the permissions of a function.

 **NOTE**

If files are deleted when metadata is deleted, the system processes the files based on the deletion policy that you have configured. For how to configure a deletion policy, see [Configuring Metadata Lifecycle](#). If no deletion policy is configured, data is moved to the recycle bin (OBS path **lake-formation-trash-dir/table_id**) by default.


----End

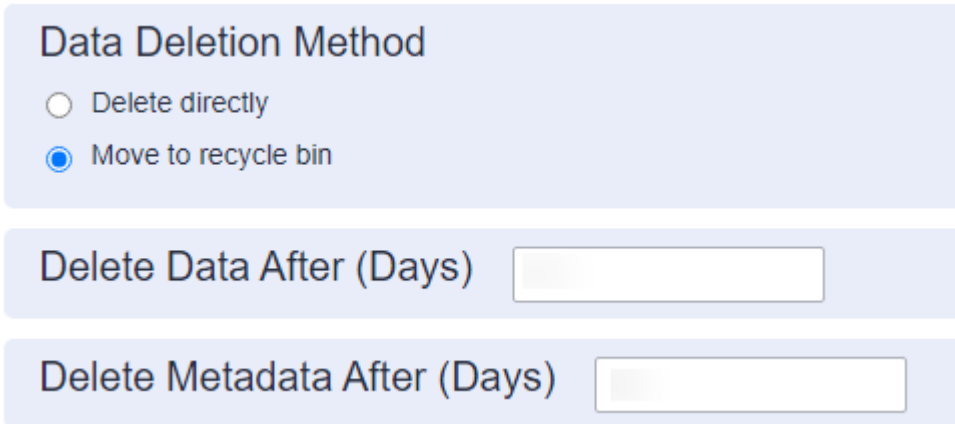
3.6 Configuring Metadata Lifecycle

- You can specify data deletion policies on the LakeFormation console to improve system flexibility and delete invalid data in a timely manner, saving space and cutting costs.

- This configuration applies to all instances that use the same OBS parallel file system. Changing the configuration for one instance may result in overwriting the configurations of other instances.

Procedure

- Step 1** Log in to the LakeFormation console.
- Step 2** In the upper left corner, click  and choose **Analytics > LakeFormation** to access the LakeFormation console.
- Step 3** Select the LakeFormation instance to be operated from the drop-down list on the left and choose **Metadata > Lifecycle** in the navigation pane.
- Step 4** Configure the metadata and data lifecycle based on the site requirements.



Data Deletion Method

Delete directly

Move to recycle bin

Delete Data After (Days)

Delete Metadata After (Days)

Table 3-5 Parameters of metadata and data lifecycle configuration

Parameter	Description
Data Deletion Method	<p>Select a data deletion method.</p> <ul style="list-style-type: none"> • Delete directly: The system marks the metadata for deletion when you delete the data stored in OBS. This will result in the direct physical deletion of the data, which cannot be recovered. • Move to recycle bin: The system marks the metadata for deletion when you delete the data. This will result in the data being moved to the recycle bin directory (lake-formation-trash-dir/table_id).

Parameter	Description
Delete Data After (Days)	<p>Deleted data expiration date. The value ranges from 1 to 30 or -1.</p> <p>This parameter cannot be configured if Delete directly is selected.</p> <ul style="list-style-type: none"> • After the data is deleted, the system will automatically delete expired data based on the number of days you set. • If you do not want the system to automatically delete expired data, set this parameter to -1. • If the value of this parameter is changed after the data is deleted, the change will take effect after a certain period of time. After the system periodically synchronizes rules, data that was deleted before the change will also be subject to the new expiration period. • The value of Delete Data After (Days) should be no bigger than that of Delete Metadata After (Days).
Delete Metadata After (Days)	<p>The value ranges from 1 to 30.</p> <ul style="list-style-type: none"> • After the metadata is deleted, the system will automatically delete expired metadata based on the number of days you set. • If you change the number of days, metadata that was deleted before the change will also be subject to the new expiration period. • The value of Delete Data After (Days) should be no bigger than that of Delete Metadata After (Days).

Step 5 Click Submit.

After the configuration is complete, the system will create a lifecycle rule named "lakeformation_trash_lifecycle" for the OBS bucket you are using in the OBS service. If you have already created a lifecycle rule for the entire file system, the system will not create this rule. Data in the recycle bin directory will be deleted based on the lifecycle rule you created for the entire file system in OBS.

----End

4 Data Permission Management

4.1 Data Permission Overview

Data lake permissions can be configured in three dimensions: database, table, and function.

Authorization and fine-grained permission control are not supported for catalog objects and their sub-metadata objects created by users.

Cloud service administrators can configure permissions of different user groups for different managed objects to centrally manage data lake resources.

NOTE

You can centrally manage permissions on resources in the data lake on the LakeFormation console. IAM users and user groups can also be associated with fine-grained permission policies of LakeFormation for authorization. For details, see [Creating a Custom Policy](#). If there are a large number of data resources in the data lake, you are advised to use the LakeFormation console to centrally manage permissions on resources in the data lake.

The following table lists the main elements in LakeFormation permission configuration.

Table 4-1 Permission configuration elements

Element	Description
Authorization Entity	You can specify any user, user group, or role to be the authorization entity. The name of an authorization entity cannot contain hyphens (-). Otherwise, the operation may fail.

Element	Description
Granted To	<ul style="list-style-type: none"> ● Resources <ul style="list-style-type: none"> - Data catalogs - Databases - Tables - Columns - Functions ● Paths
Authorization Permission	Access permission on the authorization object that the authorization entity has. Different authorization objects support different access methods. For details, see Table 4-2 .

Table 4-2 Operation rights of different authorization objects

Object	Operation Type	Description
Catalog	ALL	Perform all operations on catalogs.
	ALTER	Modify catalogs.
	CREATE_DATABASE	Create databases.
	DROP	Delete catalogs.
	DESCRIBE	Check the metadata of catalogs or switch catalogs.
	LIST_DATABASE	View the resource list in a catalog.
Database	ALL	Perform all operations on databases.
	ALTER	Modify databases.
	DROP	Delete databases.
	DESCRIBE	Check the metadata of databases or switch databases.
	LIST_TABLE	View the resource list in a database.
	LIST_FUNC	View functions in a database.
	CREATE_TABLE	Create a table in a database.
	CREATE_FUNC	Create a function in a database.
Table	ALL	Perform all operations on tables.
	ALTER	Modify tables.

Object	Operation Type	Description
	DROP	Delete tables.
	DESCRIBE	Check the metadata of tables.
	UPDATE	Update table data.
	INSERT	Insert table data.
	SELECT	Query data in a table.
	DELETE	Delete data from a table.
Column	SELECT	Query column data in a table.
Function	ALL	Perform all operations on functions.
	ALTER	Modify functions.
	DROP	Delete functions.
	DESCRIBE	Check the metadata of functions.
	EXEC	Execute functions.
Path	READ	Read files stored in a path.
	WRITE	Write data into the files stored in a path.

4.2 Granting permissions


You can centrally manage permissions on resources in the data lake and grant permissions to different authorization entities on the LakeFormation console.

Before authorization, ensure that the entity to be authorized exists. For example, the IAM user group has been created.

NOTE

You can centrally manage permissions on resources in the data lake on the LakeFormation console. IAM users and user groups can also be associated with fine-grained permission policies of LakeFormation for authorization. For details, see [Creating a Custom Policy](#). If there are a large number of data resources in the data lake, you are advised to use the LakeFormation console to centrally manage permissions on resources in the data lake.

Adding an Authorization Policy


- Step 1** Log in to the LakeFormation console.
- Step 2** In the upper left corner, click  and choose **Analytics > LakeFormation** to access the LakeFormation console.
- Step 3** Select the target LakeFormation instance from the drop-down list box on the left and choose **Data Permissions > Data Authorization**.

Step 4 Click **Authorize**. In the displayed dialog box, set parameters by referring to the following table and click **OK**.

 **NOTE**

The name of an authorization entity (including users, user groups, and roles) cannot contain hyphens (-). Otherwise, the operation may fail.

Table 4-3 Data authorization parameters

Parameter		Description
Entity Type	User group	Select User Group(s) : Select the user group to be authorized, for example, IAM user group. You can create one on the IAM console in advance.
	Role	Role : Select the role to be authorized. You can create roles in advance by following the instructions provided in Creating a Role and Binding a User with It .
	User	Select User(s) : Select the user to be authorized.
Granted To		<ul style="list-style-type: none"> • Resources: authorizes the resources in LakeFormation instances. • Paths: authorizes the paths in the OBS service. This authorization type is used to grant permissions to foreign tables or functions.
Resource Type	Catalog	This parameter is displayed when Granted To is set to Resources . Select the type of the resource to be authorized. NOTE When granting the SELECT permission to a table, you need to select columns at the same time. For example, set Column to * to select all columns.
	Database	
	Table	
	Column	
	Function	
Row Filter Criterion		Whether to set row filtering criteria for the permission policy. This parameter is displayed when Resource Type is set to Table or Column . After the row filter criteria are set, the operation type can only be SELECT .
Paths		This parameter is displayed when Granted To is set to Paths . Click  to select a path in the authorized OBS file system. You can select a maximum of 10 paths.
Operation Type		Select the operation type to be authorized. Different authorization types have different operation types. For details, see Table 4-2 .

Parameter	Description
Grant Authorization Permission	Whether to grant the authorization permission. After authorization permission is granted, an authorization entity has the permission to authorize an object to other authorization entities.

Step 5 Click **Cancel Authorization** in the **Operation** column and click **OK** to cancel the authorization of a user group or role.

NOTICE


Revoked authorizations cannot be recovered.

----End

Adding Permissions for a Specified Resource

You can add permissions for a specified resource (database or table).

Step 1 Log in to the LakeFormation console.

Step 2 In the upper left corner, click  and choose **Analytics > LakeFormation** to access the LakeFormation console.

Step 3 Select the target LakeFormation instance from the drop-down list box on the left.

Step 4 Access the page for granting permissions to a specified resource.

- **Catalog:** In the navigation pane, choose **Metadata > Catalog**. In the **Operation** column of the catalog to be authorized, choose **More > Authorize**.
- **Database:** In the navigation pane on the left, choose **Metadata > Database**. In the upper right corner, select the name of the catalog to which the target database belongs from the **Catalog** drop-down list, and choose **More > Authorize** in the **Operation** column of the database.
- **Data Table:** In the navigation pane, choose **Metadata > Table**. In the upper right corner, select the catalog to which the target data table belongs and the database name from the **Catalog** and **Database** drop-down lists. Click **Authorize** in the **Operation** column of the data table.
- **Function:** In the navigation pane on the left, choose **Metadata > Function**. In the upper right corner, select the catalog to which the target function belongs and the name of the database from the **Catalog** and **Database** drop-down lists. Click **Authorize** in the **Operation** column of the function.

Step 5 Configure related information by referring to [Table 4-3](#) and click **OK**.

Step 6 Choose **Data Permissions > Data Authorization** to view the authorization information.


After the authorization is complete, users in the selected user group or users or user groups with the selected role can perform operations on the current database.

----End

4.3 Canceling Authorization

This section describes how to cancel an authorized permission.

Procedure

- Step 1** Log in to the LakeFormation console.
- Step 2** In the upper left corner, click  and choose **Analytics > LakeFormation** to access the LakeFormation console.
- Step 3** Select the target LakeFormation instance from the drop-down list box on the left and choose **Data Permissions > Data Authorization**.
- Step 4** Search for the authorization policy to be canceled and click **Operation** next to the authorization policy and click **Revoke Authorization**.
- Step 5** Click **OK** to revoke the authorization.

NOTICE


Revoked authorizations cannot be recovered.

----End

4.4 Querying Authorization

This section describes how to query an authorized permission.

Procedure

- Step 1** Log in to the LakeFormation console.
- Step 2** In the upper left corner, click  and choose **Analytics > LakeFormation** to access the LakeFormation console.
- Step 3** Select the target LakeFormation instance from the drop-down list box on the left and choose **Data Permissions > Data Authorization**.

You can select the target permission information using **Authorization Entity**, **Entity Type**, and **Entity Source** in the upper right corner as filters.

- Step 4** View the data authorization information in the displayed list.

The following table lists the information items:

Table 4-4 Authorization information

Item	Description
Policy Type	The values include: <ul style="list-style-type: none">• DEFAULT: default permission policy.• ROW_FILTER: row filtering permission policy, including row filtering criteria.
Authorization Entity	Name of the authorized entity.
Entity Type	Type of the authorized entity. GROUP indicates the user group, ROLE indicates role, and USER indicates user.
Authorization Object	Name or path of the authorized resource. If the authorization type is set to Resources , the format is <i>Catalog.[Database].[Table]</i> .
Resource Type	The values include: <ul style="list-style-type: none">• CATALOG: catalog• DATABASE: database• TABLE: table• COLUMN: column• FUNC: function• URI: path
Permission	Name of the authorized permission. For details about the permission description, see Table 4-2 .
Authorized Permission	Authorized permission.

----End

4.5 Performing Role-based Authorization

If a role has some permissions on resources (such as databases), users or user groups with this role also have the corresponding resource operation permissions.


NOTE

If the service interconnected with the LakeFormation instance requires role authorization, the agency for interconnecting with LakeFormation must contain the permissions of the role.

For example, if the query permission of a role is required after LakeFormation is interconnected with an MRS cluster, select **lakeformation:role:describe** when creating a LakeFormation agency.

Creating a Role and Binding a User with It

Step 1 Log in to the LakeFormation console.

- Step 2** In the upper left corner, click  and choose **Analytics > LakeFormation** to access the LakeFormation console.
- Step 3** Select the target LakeFormation instance from the drop-down list box on the left and choose **Data Permission > Role**.
- Step 4** Click **Create**, set **Role Name** and **Description**, and click **OK**.
- Step 5** On the **Roles** page, click **Add** in the **Operation** column, select the target role and user, and click **OK**.

 **NOTE**

- You can also choose **Data Permissions > Users** in the navigation pane, locate the row that contains the user to be bound with the role, click **Add** in the **Operation** column, select the target role, and click **OK**.
- After the role is authorized, the users bound with the role have its permissions.

----End

Granting Permissions to the Created Role

Grant permissions to the created role. For details, see [Granting permissions](#).

5 Data Migration Management

5.1 Granting the Job Management Permission

Scenario


LakeFormation supports full or incremental migration of metadata and permissions from external services to the current LakeFormation instance for unified management.

Before managing jobs, you need to delegate LakeFormation access permissions to the current user so that related data can be written during metadata and permission migration.

Prerequisites

You have created a user by following the instructions provided in [Creating a User and Assigning Permissions](#) and added the user to the **admin** user group.

Procedure

- Step 1** Log in to the LakeFormation console as a user in the admin user group.
- Step 2** In the upper left corner, click  and choose **Analytics > LakeFormation** to access the LakeFormation console.
- Step 3** Select the LakeFormation instance to be operated from the drop-down list on the left and choose **Tasks > Job Authorization** in the navigation pane.
- Step 4** Click **Authorized** to grant the LakeFormation job management permission to the current user.

To cancel the permission, click **Cancel Authorization**.

NOTE

After the authorization is approved, LakeFormation automatically creates an agency named **lakeformation_job_trust**. Do not delete the agency during job running.

----End

5.2 Migrating Metadata

Scenario


Migrate external metadata to LakeFormation and store the data in OBS for unified management.

Prerequisites

- The current instance has been interconnected with the service whose metadata is to be migrated.
- A catalog for storing migration metadata has been created for the current instance.
- The target user has the permission to perform operations on OBS and the catalog for storing migration metadata.
- You have created an OBS parallel file system for storing migrated data.
- The name of the table owner can contain 1 to 49 characters, including only letters, digits, and underscores (_). The value cannot contain other characters such as hyphens (-).

Procedure

Step 1 Log in to the LakeFormation console.

Step 2 In the upper left corner, click  and choose **Analytics > LakeFormation** to access the LakeFormation console.

Step 3 Select the LakeFormation instance to be operated from the drop-down list on the left and choose **Tasks > Metadata Migration** in the navigation pane.

Step 4 Click **Create Migration Task**, set related parameters, and click **Submit**.

< Create Migration Task

Basic Information

* Task Name Description 0/4,000

Data Source Configurations

* Data Source * JDBC URL

* Username Password

Migration Policy

* Source Catalog * Target Catalog

* Conflict Resolution Default Owner ⓘ

* Log Path + Force Table Creation ⓘ

* Metadata Objects to Migrate DATABASE FUNCTION TABLE PARTITION

Original Path	New Path	Operation
---------------	----------	-----------

Network Connection Scheme

* Network Connection EP VPC Peering

Table 5-1 Creating a metadata migration task

Parameter	Description
Task Name	Name of the metadata migration task.
Description	Description of the created migration task.
Data Source	Type of the data to be migrated. The options are as follows: <ul style="list-style-type: none"> ● DLF ● MRS RDS for MySQL ● OpenSource HiveMetastore for MySQL ● MRS RDS FOR PostgreSQL ● MRS LOCAL GaussDB

Parameter	Description
JDBC URL	<p>URL of the JDBC link of the metadata to be migrated, followed by ?useSSL=false&permitMysqlScheme.</p> <p>NOTE Some examples are as follows:</p> <ul style="list-style-type: none"> JDBC URL of the MySQL data source type: jdbc:mysql://IP address.Port number/Database name?useSSL=false&permitMysqlScheme JDBC URL of the PostgreSQL data source type: jdbc:postgresql://IP address.Port number/Database name?useSSL=false&permitMysqlScheme <p>Set this parameter when Data Source is not set to DLF. Set the following parameters when Data Source is not set to DLF:</p> <ul style="list-style-type: none"> Username: username for accessing the data source. Password: password for accessing the data source. If the user has a password, this parameter is mandatory. Otherwise, leave this parameter blank.
Access Point	<p>Access point of the metadata service to be migrated. This parameter is displayed when Data Source is set to DLF. In addition, you need to set the following parameters:</p> <ul style="list-style-type: none"> Access Key: Obtain the AK from DLF O&M personnel. Secret Key: Obtain the SK from DLF O&M personnel.
Source Catalog	Name of the catalog to which the metadata to be migrated belongs.
Target Catalog	Name of the catalog to which metadata is migrated in LakeFormation.
Conflict Resolution	Policy for resolving conflicts during migration. Currently, only Update old metadata is supported.
Default Owner	<p>Default owner of metadata after migration.</p> <ul style="list-style-type: none"> If the configured default owner does not have the corresponding metadata operation permission, the migrated metadata cannot be added, deleted, modified, or queried. In this case, you can grant permissions to the owner or migrate permissions. If all metadata can be used properly before the migration, you do not need to set this parameter.
Log Path	<p>Storage location of logs generated during migration. Click + to select a path.</p> <p>The path must exist in OBS. If the path is customized, the migration task will fail.</p>

Parameter	Description
Force Table Creation	Selecting this option will bypass OBS path restrictions when creating an internal table.
Metadata Objects to Migrate	<p>Select the metadata objects to be migrated.</p> <ul style="list-style-type: none"> ● All: Migrate databases, functions, data tables, and partitions. ● Database: Migrate databases. ● Function: Migrate functions. ● Table: Migrate tables. ● Partition: Migrate partitions. <p>NOTE</p> <ul style="list-style-type: none"> - Select All to migrate all metadata for the first migration task. - Ensure that the upper-level directory of the selected metadata exists if All is not selected. For example, you need to ensure that the target catalog contains the database (for example, DB_1) where the tables are located if you plan to set this parameter to Table. Otherwise, the table migration will fail. - Ensure that the function class name exists if you plan to set this parameter to Function to guarantee a successful function migration task.
Add Location Rule	<ul style="list-style-type: none"> ● If the prefix of the metadata storage path is not obs://, click Add Location Rule to replace the prefix with obs:// and ensure that the corresponding OBS storage path exists. For example, if the current metadata storage path is file:/a/b, set Original Path to file:/ and New Path to obs://. Ensure that the obs://a/b path exists in the OBS parallel file system, the new metadata storage path is obs://a/b. ● You can create multiple rules at the same time. If a rule conflict occurs, the rule on the top of the page prevails.

Step 5 Click **Start** in the **Operation** column to run the migration task.

 **NOTE**

- Before running a migration task, you need to authorize the task by referring to [Granting the Job Management Permission](#).
- After the migration task starts, if new metadata is added to the source database, the new metadata will not be migrated. You need to run the migration task again to migrate the new metadata.
- If the task fails to be executed, you can click **Start** in the **Operation** column to retry after rectifying the fault.

You can click **Metadata** on the navigation pane and click the name of target metadata object to view the metadata object after the migration. For example, choose **Metadata > Database** to view the migrated database.

Click **Edit** or **Delete** in the **Operation** column to modify or delete a task.

Step 6 Click **View Log** in the **Operation** column to view the logs generated during task running.

By default, the latest 50 lines of logs are displayed. You can click the hyperlink at the bottom of the log to view the complete log.

The following table lists some error messages in logs and their causes.

Error Message	Cause
field 'storageDescriptor.location' must match '^(obs har)://.+.\$'	Incorrect location rule is configured. (The metadata storage path should start with obs://.)
Invalid input parameter	The input parameter of the metadata is invalid or LakeFormation does not support such metadata.
Incorrect type of column xxx.	The column type is invalid or LakeFormation is incompatible with the column type.
No permission to perform this operation on resources.	The default owner is incorrectly configured or the owner does not have the metadata operation permission.

----End

5.3 Migrating Permissions

Scenario

After metadata migration is complete, you can migrate metadata permissions to LakeFormation. After the migration is successful, the default owner bound to the metadata will have the metadata operation permissions.

Prerequisites


- Metadata has been migrated.
- You have obtained the permission to perform operations on OBS and have created an OBS parallel file system for storing data.
- The permission policy files to be migrated have been exported and uploaded to the OBS parallel file system. For details about how to export permissions, contact the corresponding service support personnel.
- Authorization entities (except roles) in the permission policy have been created in advance and their names are consistent. Metadata objects contained in the permission policy already exist and their names are consistent.

If the migration type is set to **DLF**, the mapping and migration policies are as follows:

- RAM user: IAM user (If the corresponding IAM user does not exist, the permission policy will not be migrated.)
- RAM role: IAM user group (If the corresponding IAM user group does not exist, the permission policy will not be migrated.)
- DLF role: LakeFormation role (If this role does not exist, it will be automatically created.)

Procedure

Step 1 Log in to the LakeFormation console.

Step 2 In the upper left corner, click  and choose **Analytics > LakeFormation** to access the LakeFormation console.

Step 3 Select the LakeFormation instance to be operated from the drop-down list on the left and choose **Tasks > Permission Migration** in the navigation pane on the left.

Step 4 Click **Create Migration Task**, set related parameters, and click **Submit**.

Table 5-2 Creating a permission migration task

Parameter	Description
Task Name	Name of the permission migration task to be created.
Description	Description of the created migration task.
Policy Type	Type of the permission policy to be migrated. <ul style="list-style-type: none"> ● DLF ● RANGER
Policy File Path	Storage location of the permission policy file to be migrated in OBS.
Policy File	Name of the file whose permission policies are to be migrated.

Parameter	Description
Log Path	Storage location of logs generated during migration.
Catalog ID	Catalog name of the permission source.
Ranger Authorization Conversion Objects	<p>Parameter for specifying the conversion relationship between authorization objects of the Ranger permission policy. The prefix and suffix are added to the name of the authorization object.</p> <p>This parameter is displayed when Policy Type is set to RANGER.</p> <p>In addition, you need to select conversion objects and enter prefixes and suffixes in the User, User Group, and Role areas.</p>

Step 5 Click **Run** in the **Operation** column to run the migration task.

Click **View Log** to view the logs generated during task running.

 **NOTE**

- Before running a migration task, you need to authorize the task by referring to [Granting the Job Management Permission](#).
- If the task fails to be executed, you can click **Start** in the **Operation** column to retry after rectifying the fault.

Click **Edit** or **Delete** in the **Operation** column to modify or delete a task.


----End

6 Managing Clients

On this page, you can create and manage clients, view client IP addresses, and use the IP addresses to connect multiple services to LakeFormation instances.

Procedure

Step 1 Log in to the LakeFormation console.

Step 2 In the upper left corner, click  and choose **Analytics > LakeFormation** to access the LakeFormation console.

Step 3 Select the LakeFormation instance to be operated from the drop-down list on the left and click **Clients** in the navigation pane.

Step 4 Click **Create**. In the displayed dialog box, set the following parameters and click **OK**.

If no suitable VPC or subnet is available, click **create one** to access the VPC console to create one.

Table 6-1 Parameters for creating an access management client

Parameter	Description
Client	Enter a client name.
VPC	Select the VPC to which the client to be connected belongs. The value must be the same as that of the service to be connected.
Subnet	Select the subnet to which the client to be connected belongs. The value must be the same as that of the service to be connected.

Step 5 Click **View Details** in the **Operation** column to open the details page.

In the **Basic Information** area, you can view the ID, client name, status, access mode, VPC, and subnet.

In the **Accessed Connections** area, you can view the VPC endpoint ID and access IP address.

Basic Information					
ID	[Redacted]	Client	[Redacted]	Status	Running
Access Mode	[Redacted]	VPC	[Redacted]	Subnet	[Redacted]
Created	[Redacted]				

Accessed Connections		
Endpoint ID	Owner	IP Address
[Redacted]	[Redacted]	[Redacted]

10 Total Records: 1 < 1 >

----End

7 Viewing Audit Logs

Cloud Trace Service (CTS) records operations on LakeFormation instances for query, audit, and backtrack.

Operations Recorded by CTS

Table 7-1 Operations recorded by CTS

Operation	Resource Type	Trace Name
Creating a catalog	Catalog	createCatalog
Deleting a catalog	Catalog	dropCatalog
Modifying a catalog	Catalog	alterCatalog
Creating a database	Database	createDatabase
Deleting a database	Database	dropDatabase
Modifying a database	Database	alterDatabase
Creating a data table	Table	createTable
Deleting a data table.	Table	dropTable
Modifying a data table.	Table	alterTable
Clearing the data in a table	Table	truncateTable
Creating a function	Function	createFunction
Modifying the attributes of a function	Function	alterFunction
Deleting a function	Function	dropFunction
Creating an instance	instance	createInstance
Modifying an instance	instance	updateInstance

Operation	Resource Type	Trace Name
Deleting an instance	instance	deleteInstance
Granting permissions	Access	grantAccess
Revoking authorization	Access	revokeAccess
Updating statistics about a specified column in a table	TableColumnStatistic	setTableColumnStatistics
Deleting statistics about a specified column in a table	TableColumnStatistic	deleteTableColumnStatistics
Adding column restrictions for creating tables in batches	TableConstraint	addConstraints
Deleting a column restriction	TableConstraint	deleteConstraints
Adding partition information in batches	Partition	addPartitions
Modifying partition information in batches	Partition	alterPartitions
Deleting partition information in batches	Partition	dropPartitions
Clearing list information in batches	Partition	truncatePartitions
Setting partition column statistics in batches	PartitionColumnStatistic	setPartitionColumnStatistics
Deleting partition column statistics	PartitionColumnStatistic	deletePartitionColumnStatistics

Viewing an Audit Log

Query LakeFormation traces on the CTS console.

For details, visit [Querying Real-Time Traces](#).

8 Change History

Released On	Description
2024-02-01	This issue is the first official release.