

DataArts Studio

Preparations

Issue 01
Date 2024-05-31



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.

Huawei Cloud Computing Technologies Co., Ltd.

Address: Huawei Cloud Data Center Jiaoxinggong Road
Qianzhong Avenue
Gui'an New District
Gui Zhou 550029
People's Republic of China

Website: <https://www.huaweicloud.com/intl/en-us/>

Contents

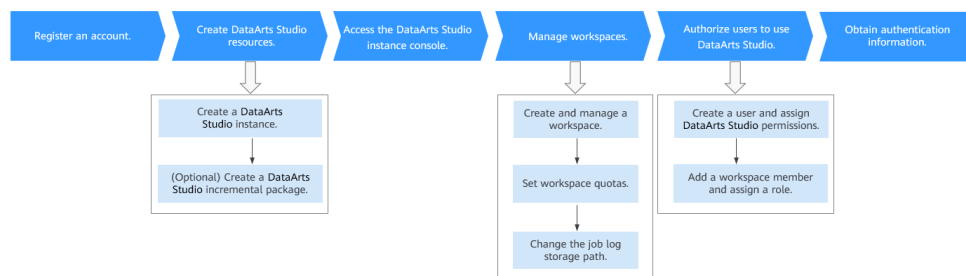
1 Preparations.....	1
2 Buying DataArts Studio Instances.....	2
2.1 Buying a DataArts Studio Basic Package.....	2
2.2 (Optional) Buying a DataArts Studio Incremental Package.....	9
3 Accessing the DataArts Studio Instance Console.....	16
4 Managing a Workspace.....	18
4.1 Creating and Managing a Workspace.....	18
4.2 Setting Workspace Quotas.....	22
4.3 (Optional) Changing the Job Log Storage Path.....	24
5 Authorizing Users to Use DataArts Studio.....	26
5.1 Creating an IAM User and Assigning DataArts Studio Permissions.....	26
5.2 Adding a Member and Assigning a Role.....	28
6 (Optional) Obtaining Authentication Information.....	32

1 Preparations

Before using DataArts Studio, you need to register a Huawei account, buy a DataArts Studio instance, access the DataArts Studio instance console, manage workspaces, and authorize users to use DataArts Studio.

The following figure shows the preparations and operations, which will be described in detail in later sections.

Figure 1-1 Preparations for using DataArts Studio



2 Buying DataArts Studio Instances

2.1 Buying a DataArts Studio Basic Package

DataArts Studio is billed based on the basic package and incremental package. The basic package refers to the DataArts Studio instance. For details about how to purchase a DataArts Studio instance, see [Buying a DataArts Studio Basic Package](#).

Background

- Only users who have the **DAYU Administrator** or **Tenant Administrator** permissions can buy DataArts Studio instances or DataArts Studio incremental packages. To allow users to buy instances or incremental packages, you must grant the required permissions to them.

NOTE


- Users with the **Tenant Administrator** permissions can perform all operations except IAM user management. For security purposes, you are not advised to grant the **Tenant Administrator** permissions to IAM users.
- Only users who have the **Security Administrator** permissions can create cloud service agencies. Cloud service agencies allow DataArts Studio to perform operations such as task scheduling and resource O&M on cloud services on your behalf.

Prerequisites

You have obtained a VPC, subnet, and security group. You can also apply for them when you buy a DataArts Studio instance.

For details, see [Virtual Private Cloud User Guide](#).

Logging In to DataArts Studio Console

1. Log in to the Huawei Cloud console.
2. In the upper left corner of the console, click  , and choose **DataArts Studio** to access the DataArts Studio console.

Buying a DataArts Studio Basic Package

- Step 1** Access the page for [purchasing a DataArts Studio instance](#).
- Step 2** On the displayed page, set parameters for the DataArts Studio instance. [Table 2-1](#) provides descriptions of the parameters.

Table 2-1 DataArts Studio instance parameters

Parameter	Example	Description
Region	None	<p>The region where the instance resides. Resources in different regions cannot communicate with each other.</p> <p>When selecting a region, consider the following factors:</p> <ul style="list-style-type: none">• Location You are advised to select a region closest to your target users. This reduces network latency and improves access rate. However, Chinese mainland regions provide the same infrastructure, BGP network quality, and operations and configurations on resources. Therefore, if your target users are in the Chinese mainland, you do not need to consider the network latency differences when selecting a region.• Relationship between cloud services Cloud services in different regions cannot communicate with each other through an internal network.• Resource price Resource pricing may vary in different regions. For details, see Product Pricing Details. <p>For details, see AZs.</p>

Parameter	Example	Description
Enterprise Project	default	<p>Enterprise project associated with the default workspace of the DataArts Studio instance. An enterprise project facilitates management of cloud resources. For details, see Enterprise Management User Guide.</p> <p>This parameter is available only when an enterprise project has been created. If you want to connect the DataArts Studio instance to a cloud service (such as DWS, MRS, and RDS), ensure that the enterprise project of the DataArts Studio workspace is the same as that of the cloud service instance.</p> <ul style="list-style-type: none">You can buy only one DataArts Studio instance for an enterprise project.If you want to enable communication between DataArts Studio and another cloud service, ensure that the enterprise project of DataArts Studio is the same as that of the cloud service.
Version	Basic	<p>Select a DataArts Studio version. For details about the version differences, see Versions.</p> <p>NOTE</p> <p>The DataArts Studio instance contains a CDM cluster by default. You are advised to use this cluster as a connection agent. For data migration jobs, purchase a DataArts Migration incremental package of a higher version. For details, see Buying a DataArts Migration Incremental Package.</p>
Billing Mode	Yearly/Monthly	The DataArts Studio basic package is billed on a monthly/yearly basis.
Instance Name	DataArts Studio-test	Name of the DataArts Studio instance The instance name cannot be changed.

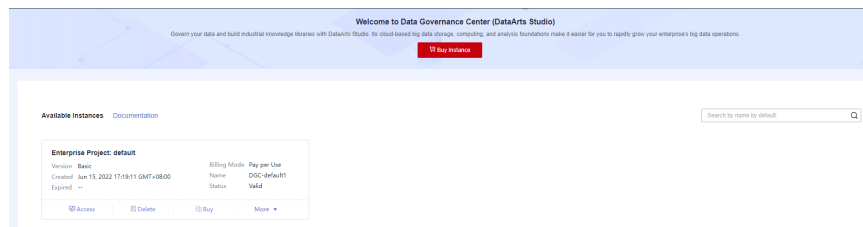
Parameter	Example	Description
AZ	AZ1	<p>AZ of the DataArts Studio instance, that is, the AZ where the CDM cluster is located. The DataArts Studio instance communicates with other services through the CDM cluster.</p> <p>When you buy a DataArts Studio instance or incremental package for the first time, there is no requirement on the AZ. When you buy a new DataArts Studio instance or incremental package, determine whether to select the same AZ as the existing one based on your DR and network latency demands.</p> <ul style="list-style-type: none">• If your application requires good DR capability, deploy resources in different AZs in the same region.• If your application requires a low network latency between instances, deploy resources in the same AZ. <p>For details, see AZs.</p>
VPC	vpc1	<p>VPC, subnet, and security group to which the CDM cluster in the DataArts Studio instance belongs. The DataArts Studio instance communicates with other services through the CDM cluster.</p> <p>If you want to connect the DataArts Studio instance or CDM cluster to a cloud service (such as DWS, MRS, and RDS), ensure that the CDM cluster can communicate with the cloud service. If the CDM cluster and the cloud service are in the same region, VPC, subnet, and security group, they can communicate with each other by default. If the CDM cluster and the cloud service are in the same VPC but in different subnets or security groups, you must configure routing rules and security group rules.</p> <p>For details about the operations on VPCs, subnets, and security groups, see Virtual Private Cloud User Guide.</p> <p>NOTE</p> <ul style="list-style-type: none">• After a DataArts Studio instance is created, the VPC, subnet, and security group of the default CDM cluster cannot be changed. Exercise caution when setting them during the instance creation.• You can select a VPC subnet shared by the VPC owner when you buy a DataArts Studio instance. Through VPC subnet sharing, you can easily configure and manage multiple accounts' resources at low costs. For details about how to share a VPC subnet, see VPC Sharing.
Subnet	subnet-1	
Security Group	sg-1	

Parameter	Example	Description
Required Duration	1 year	Select a duration based on your requirements.
Auto-renew	N/A	You can select Auto-renew to automatically renew the subscription by month or by year. If you choose the monthly billing mode, your subscription will be automatically renewed each month. If you choose the yearly billing mode, your subscription will be automatically renewed each year.
Tag	Tag key: key1 Tag value: asd	DataArts Studio instance tags can be used in the following scenarios: <ul style="list-style-type: none">• If there are a large number of cloud resources, you can add tags to them (including DataArts Studio instances) by user, maintainer, or usage. Then you can use Tag Management Service (TMS) to identify tags and manage cloud resources easily.• If there are multiple DataArts Studio instances, you can add tags to them by user, maintainer, or usage. Then you can search for and identify DataArts Studio instances by tag on the DataArts Studio instance list page. You can set tags when creating a DataArts Studio instance. <ol style="list-style-type: none">1. Enter a tag key in the text box or select a predefined tag key from the drop-down list. NOTE To add a predefined tag, you need to create one on TMS and then select it from the Tag key drop-down list. You can click View predefined tags to enter the Predefined Tag page of TMS. Then, click Create Tag to create a predefined tag. For details, see Creating Predefined Tags in <i>Tag Management Service User Guide</i>.2. Enter a tag value in the text box or select a predefined tag value from the drop-down list.3. Click Add to add the tag. NOTE<ul style="list-style-type: none">- A maximum of 20 tags can be added.- Only one tag value can be added to a tag key.- The key name must be unique in the same instance.

Step 3 Confirm the settings and click **buy Now**.

Step 4 Click **Next**. After you pay for your instance, wait until the instance is created. The created instance is displayed on the homepage.

Figure 2-1 Viewing a DataArts Studio instance



Step 5 When you return to the DataArts Studio management console, the **Authorize Access** dialog box is displayed, prompting you to authorize the listed services. DataArts Studio interacts with the cloud services and needs to collaborate with them. Therefore, you need to create a cloud service agency to delegate permissions to DataArts Studio so that DataArts Studio can use these cloud services and perform task scheduling and resource O&M on behalf of you.

NOTICE

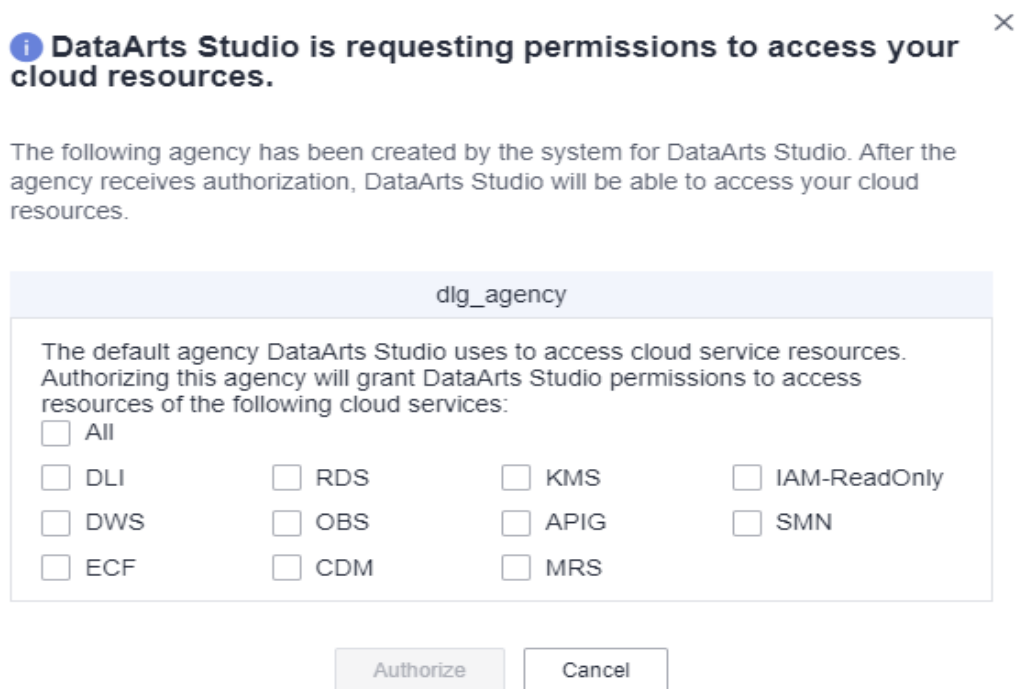
Only users who have the **Security Administrator** permissions can create cloud service agencies. Cloud service agencies allow DataArts Studio to perform operations such as task scheduling and resource O&M on cloud services on your behalf.

Cloud service agencies include permissions related to DWS, MRS, RDS, OBS, SMN, and KMS. You can access the IAM agency page to view the agency scope. You do not need to apply for an agency for users. The agency of the account is used.

Select all services and click **Authorize**. The platform automatically creates the default agency **dlg_agency**.

- After the authorization is complete, the **Authorize Access** dialog box will not be displayed when you access the DataArts Studio console homepage next time.
- If you select only some services for authorization, the system still displays the dialog box next time you access the DataArts Studio console homepage, prompting you to authorize access to unauthorized cloud services.

Figure 2-2 Authorize Access dialog box



Step 6 In the list of instances, locate your instance and click **Access** to access the DataArts Studio console.

----End

2.2 (Optional) Buying a DataArts Studio Incremental Package

DataArts Studio provides basic and incremental packages. If the basic package cannot meet your requirements, you can buy an incremental package. Before you buy an incremental package, ensure that you have a DataArts Studio instance.

DataArts Studio provides the following types of incremental packages:

- DataArts Migration incremental package
CDM clusters can migrate data to the cloud and integrate data into the data lake. It provides wizard-based configuration and management and can integrate data from a single table or an entire database incrementally or periodically. The DataArts Studio basic package contains a CDM cluster. If the cluster cannot meet your requirements, you can buy a DataArts Migration incremental package.
DataArts Migration incremental packages support the following billing modes:
 - **Buying a Pay-Per-Use CDM Cluster:** You can enable or disable the package as you like and you are charged by the usage duration.
 - **Buying a Yearly/Monthly CDM Cluster:** This mode is more cost-effective and a better choice if you want to use the package for a long time.
- DataArts DataService incremental package

DataArts DataService is a platform where you can develop, test, and deploy your data services. DataArts DataService ensures agile response to data service needs, easier data retrieval, better experience for data consumers, higher efficiency, and better monetization of data assets. To use DataArts DataService Exclusive, you need to purchase an exclusive cluster incremental package in addition to the DataArts Studio basic package.

For details, see [Buying a DataArts DataService Exclusive Cluster](#).

Background

- You can choose **More > View Packages** to view incremental packages.
- You can buy a CDM cluster by buying an incremental package on the DataArts Studio console or buy a CDM cluster directly on the CDM console. The differences are as follows:
 - a. Billing: Clusters purchased on the DataArts Studio console only support packages purchased on the DataArts Studio console, while clusters purchased on the CDM console only support the discount packages purchased on the CDM console.
 - b. Permission control: For clusters purchased on the DataArts Studio console, permissions are managed based on the DataArts Studio permission system. For clusters purchased on the CDM console, permissions are managed based on the CDM permission system.
 - c. Application scenarios: Clusters purchased on the DataArts Studio console are isolated by workspace and can be used only in associated workspaces. Clusters purchased on the CDM console are not isolated by workspace and can be used in all DataArts Studio workspaces.

The former is recommended, and this section uses it as an example.

Buying a Pay-Per-Use CDM Cluster

If you buy a pay-per-use incremental package, the system automatically creates a CDM cluster based on the specifications you select.

1. Locate an enabled instance and click **Buy**.
2. On the displayed page, set parameters based on [Table 2-2](#).

Table 2-2 Parameters for the CDM incremental package

Parameter	Description
Package	Select DataArts Migration .
Billing Mode	Select Pay per Use .

Parameter	Description
AZ	<p>When you buy a DataArts Studio instance or incremental package for the first time, there is no requirement on the AZ. When you buy a new DataArts Studio instance or incremental package, determine whether to select the same AZ as the existing one based on your DR and network latency demands.</p> <ul style="list-style-type: none">• If your application requires good DR capability, deploy resources in different AZs in the same region.• If your application requires a low network latency between instances, deploy resources in the same AZ. <p>For details, see AZs.</p>
Workspace	Select the workspace where the DataArts Migration incremental package will be used. The CDM cluster can only be used in this workspace.
Enterprise Project	If the CDM cluster is associated with multiple workspaces, you need to specify an enterprise project for the CDM cluster.
Cluster	Customize the cluster name.
Instance	<p>The following CDM cluster flavors are available:</p> <ul style="list-style-type: none">• cdm.large: 8 vCPUs and 16 GB of memory. The maximum and assured bandwidths are 3 Gbit/s and 0.8 Gbit/s. Up to 16 jobs can be executed concurrently.• cdm.xlarge: 16 vCPUs and 32 GB of memory. The maximum and assured bandwidths are 10 Gbit/s and 4 Gbit/s. Up to 32 jobs can be executed concurrently. This flavor is suitable for migrating terabytes of data that requires a bandwidth of 10GE.• cdm.4xlarge: 64 vCPUs and 128 GB of memory. The maximum and assured bandwidths are 40 Gbit/s and 36 Gbit/s. Up to 128 jobs can be executed concurrently.• The free ECS with 4 vCPUs and 8 GB memory provided by DataArts Studio can run only one job.

Parameter	Description
VPC	VPC, subnet, and security group to which the CDM cluster in the DataArts Studio instance belongs.
Subnet	If you want to connect the DataArts Studio instance or CDM cluster to a cloud service (such as DWS, MRS, and RDS), ensure that the CDM cluster can communicate with the cloud service. If the CDM cluster and the cloud service are in the same region, VPC, subnet, and security group, they can communicate with each other by default. If the CDM cluster and the cloud service are in the same VPC but in different subnets or security groups, you must configure routing rules and security group rules. For details about the operations on VPCs, subnets, and security groups, see Virtual Private Cloud User Guide .
Security Group	
	NOTE <ul style="list-style-type: none">• After a CDM instance is created, the VPC, subnet, and security group of the instance cannot be changed. Exercise caution when setting them during the instance creation.• You can select a VPC subnet shared by the VPC owner when you buy a CDM cluster. Through VPC subnet sharing, you can easily configure and manage multiple accounts' resources at low costs. For details about how to share a VPC subnet, see VPC Sharing.
IPv6 Dual Stack	If you enable this function, both private IPv4 and IPv6 addresses can be used to access the cluster. NOTE <ul style="list-style-type: none">• If you enable this function, you can only select subnets for which IPv6 CIDR blocks are enabled. If you want to select a subnet for which IPv6 CIDR blocks are disabled, enable IPv6 CIDR blocks for the subnet in the VPC service.• IPv6 dual stack can be enabled for private IP addresses, but not for public IP addresses.

NOTICE

You cannot modify the specifications of an existing cluster. If you need higher specifications, create another cluster.

3. Click **Buy Now**, confirm the specifications, and click **Next**.
4. View the CDM cluster in the corresponding workspace.

Buying a Yearly/Monthly CDM Cluster

If you want to enjoy the preferential price of the yearly/monthly incremental package, you can buy a yearly/monthly incremental package and then buy a pay-per-use incremental package which is in the same region and has the same specifications as the yearly/monthly incremental package.

1. Locate an enabled instance and click **Buy**.

2. On the displayed page, set the following parameters:
 - a. Select **DataArts Migration** for **Package**.
 - b. Select **Yearly/Monthly** for **Billing Mode**.
 - c. Specify a validity period in **Usage Duration** for the package.
 - d. Enter the number of subscribed packages in **Purchase Quantity**. For example, if you specify **1 month** for **Usage Duration** and **2** for **Purchase Quantity**, you will have a 1490-hour quota and the validity period will be 1 month.
3. Click **Buy Now**. If there are no changes made to the specifications, submit the order.
4. After you buy the yearly/monthly package, the system will not automatically create a CDM cluster. You need to buy a pay-per-use incremental package in the same region and with the same specifications as the yearly/monthly package by following the instructions in [Buying a Pay-Per-Use CDM Cluster](#). Then you can enjoy the preferential price of the yearly/monthly incremental package.

Buying a DataArts DataService Exclusive Cluster

Step 1 Locate an enabled instance and click **buy**.

Step 2 On the displayed page, set parameters based on [Table 2-3](#).

Table 2-3 Parameters for buying an exclusive DataArts DataService instance

Parameter	Description
Package	Select DataArts DataService .
Billing Mode	Currently, Yearly/Monthly is supported.
Workspace	The workspace for which you want to use the incremental package. For example, if you want to use DataArts DataService Exclusive in workspace A of the DataArts Studio instance, select workspace A. After you buy an exclusive DataArts DataService cluster, you can view it in workspace A.
AZ	When you buy a DataArts Studio instance or incremental package for the first time, you can select any available AZ. When you buy another DataArts Studio instance or incremental package, determine whether to deploy your resources in the same AZ based on your DR and network latency demands. <ul style="list-style-type: none">• If your application requires good DR capability, deploy resources in different AZs in the same region.• If your application requires a low network latency between instances, deploy resources in the same AZ. For details, see AZs .
Name	N/A
Description	A description of the exclusive DataArts DataService cluster.

Parameter	Description
Version	Cluster version of the exclusive DataArts DataService cluster.
Cluster Details	The number of concurrent API requests supported varies depending on the instance specifications.
Enabling public IP address	If you select Enabling public IP address , external services can call the APIs created in exclusive instances through the Internet address.
Bandwidth	Bandwidth range on the Internet.
VPC	<p>A VPC is a secure, isolated, and logical network environment. Cloud resources (such as ECSs) within the same VPC can call APIs using the private IP address of DataArts DataService Exclusive.</p> <p>Deploy the DataArts DataService Exclusive instance in the same VPC as your other services to facilitate network configuration and secure network access.</p> <p>NOTE After the DataArts DataService instance is created, the VPC cannot be changed.</p>
Subnet	<p>A subnet provides dedicated network resources that are logically isolated from other networks for network security.</p> <p>Deploy the DataArts DataService Exclusive instance in the same subnet of the same VPC as your other services to facilitate network configuration and secure network access.</p> <p>NOTE After the DataArts DataService instance is created, the subnet cannot be changed.</p>
Security Group	<p>A security group is used to set port access rules, define ports that can be accessed by external services, and determine the IP addresses and ports that can be accessed externally.</p> <p>For example, if the backend service is deployed on an external network, configure security group rules to allow access to the IP address and listening port of the backend service.</p> <p>NOTE</p> <ol style="list-style-type: none">1. If Enabling the public IP address is selected, the security group must allow access from ports 80 (HTTP) and 443 (HTTPS) in the inbound direction.2. After the DataArts DataService instance is created, the security group cannot be changed.
Managing Cluster Resources Using an Enterprise Project	Enterprise project associated with the exclusive DataArts DataService cluster. An enterprise project facilitates management of cloud resources. For details, see Enterprise Management User Guide .
Nodes	N/A

Parameter	Description
Required Duration	N/A

Step 3 Click **buy Now**, confirm the settings, and click **Next**.

----End

3 Accessing the DataArts Studio Instance Console

Prerequisites

A DataArts Studio instance has been obtained. For details, see [Buying a DataArts Studio Basic Package](#).

Procedure

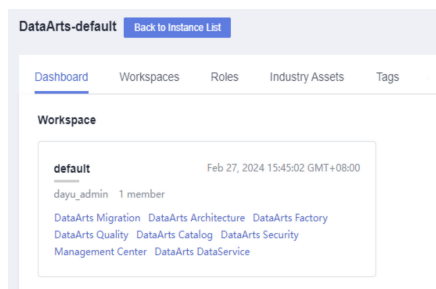
- Step 1** Log in to the Huawei Cloud console. Choose DataArts Studio from the service list in the upper left corner to access the DataArts Studio instance console.
- If there are multiple DataArts Studio instances in the current region, the instance list is displayed by default. Locate an instance and click **Access** to access the DataArts Studio console.

Figure 3-1 Instance list



- If there is only one DataArts Studio instance in the current region, the DataArts Studio console homepage is displayed by default.

Figure 3-2 Console homepage



----End

4 Managing a Workspace

4.1 Creating and Managing a Workspace

The system will create a default workspace and assign the admin role to you after you buy a DataArts Studio instance. You can use the default workspace or create another workspace.

A workspace in a DataArts Studio instance is the basic unit for member management and role and permission allocation. It provides all DataArts Studio functions. Workspaces are allocated by branch or subsidiary (such as the group, subsidiary, and department), business domain (such as the procurement, production, and sales), or implementation environment (such as the development, test, and production environment). There are no fixed rules.

Admins can manage user (member) permissions, resources, and compute engines in a workspace. To enable users to work together, admins can add users to a workspace and assign the preset roles of DataArts Studio (admin, developer, operator, and visitor) to the users. Users other than admins can access Management Center, DataArts Migration, DataArts Architecture, DataArts Catalog, DataArts Quality, DataArts DataService, DataArts Security, and DataArts Factory only after they are added to a workspace and assigned relevant roles.

Constraints

- There is no limit on the number of workspaces that can be created for a DataArts Studio instance. You can plan the quota as needed.
- The storage of job logs and dirty data depends on the OBS service.

Prerequisites

A DataArts Studio instance has been obtained. For details, see [Buying a DataArts Studio Basic Package](#).

Background

- The system will create a default workspace and assign the admin role to you after you buy a DataArts Studio instance.

- In a DataArts Studio instance created by a master account, if an IAM user of the account needs to create a workspace, the IAM user must be assigned the **DAYU Administrator** or **Tenant Administrator** policy. By default, the master account has all the permissions for a DataArts Studio instance created by a sub-user.
- Users assigned the permissions of **DAYU User** can access a workspace only after they are added as members of the workspace.

Creating a Workspace

Step 1 Log in to the DataArts Studio console as user **DAYU Administrator** or **Tenant Administrator**. For details, see [Accessing the DataArts Studio Instance Console](#).

Step 2 On the **Workspaces** page, click **Create**. In the displayed dialog box, set parameters based on [Table 4-1](#).

Figure 4-1 Creating a workspace

The screenshot shows a 'Create' dialog box with the following fields and controls:

- Name:** A text input field with a placeholder 'Enter a name.'
- Description:** A text input field with a placeholder 'Enter a description.' and a character count '0/4,096'.
- Mode:** A dropdown menu currently set to 'Simple'.
- Enterprise Project:** A dropdown menu with a 'C' icon.
- Job Log Path:** A text input field with a 'Select' button and a help icon.
- Dirty Data Path:** A text input field with a 'Select' button and a help icon.
- tags:** A section with a note: 'It is recommended that you use TMS's predefined tag function to add the same tag to different cloud resources. [View predefined tags](#) C'. Below this is a text area for adding tags, followed by input fields for 'Enter a tag key', 'Enter a tag value', and an 'Add' button. A note below says 'Tags you can still add: 20'.
- Buttons:** 'OK' (red) and 'Cancel' (grey) buttons at the bottom.

Table 4-1 Parameters for creating a workspace

Parameter	Description
Name	Workspace names can contain only letters, numbers, underscores (_), and hyphens (-). They cannot exceed 32 characters. In a DataArts Studio instance, workspace names must be unique.
Description	A description of the workspace.

Parameter	Description
Mode	<p>Workspace mode</p> <ul style="list-style-type: none">• Simple: This traditional mode is easy to use but does not allow you to strictly control the data development process and table permissions.• Enterprise: You can isolate the development environment from the production environment in the DataArts Factory and Management Center modules of DataArts Studio. This prevents developers' operations from affecting services in the production environment. For details about the enterprise mode, see DataArts Studio Enterprise Mode Overview.
Enterprise Project	<p>Enterprise project associated with the default workspace of the DataArts Studio instance. An enterprise project facilitates management of cloud resources. For details, see Enterprise Management User Guide.</p> <p>This parameter is available only when an enterprise project has been created. If you want to connect the DataArts Studio instance to a cloud service (such as DWS, MRS, and RDS), ensure that the enterprise project of the DataArts Studio workspace is the same as that of the cloud service instance.</p> <ul style="list-style-type: none">• You can buy only one DataArts Studio instance for an enterprise project.• If you want to enable communication between DataArts Studio and another cloud service, ensure that the enterprise project of DataArts Studio is the same as that of the cloud service.
OBS Bucket for Job Logs	<p>The OBS bucket for storing logs of DataArts Studio data development jobs. To use the DataArts Factory module of DataArts Studio, workspace members must have the read and write permissions on the OBS bucket for storing job logs. Otherwise, the system cannot read or write job logs generated during data development.</p> <ul style="list-style-type: none">• Click Select. You can select a created OBS bucket. The selected OBS bucket is globally configured in the current workspace.• If this parameter is not set, job logs generated during data development are stored in the OBS bucket named dlf-log-{projectId} by default. {projectId} indicates the project ID, which can be obtained by referring to Obtaining a Project ID and Account ID.


Parameter	Description
OBS Bucket for DLI Dirty Data	<p>The OBS bucket for storing dirty data generated during DLI SQL execution in DataArts Studio Data Development. To use DataArts Studio DataArts Factory to develop and execute DLI SQL statements, workspace members must have the read and write permissions on the OBS bucket where DLI dirty data is stored. Otherwise, the system cannot read or write the dirty data generated during DLI SQL execution.</p> <ul style="list-style-type: none">• Click Select. You can select a created OBS bucket. The selected OBS bucket is globally configured in the current workspace.• If this parameter is not set, dirty data generated during DLI SQL execution is stored in the OBS bucket named dlf-log-{projectId} by default.
Tag	<p>If you have multiple workspaces, you can add tags to classify them by user, maintainer, or usage. On the Workspaces page, you can search for and identify workspaces by tag.</p> <p>A tag consists of a tag key and a tag value. You can select tags predefined in Tag Management Service (TMS) or enter custom tags. Each tag key must be unique. Only one tag value can be added to a tag key.</p>

Step 3 Click **OK**.

----End

Related Operations


- **Disable a workspace:** After a workspace is created, it is enabled by default. You can disable or enable it as needed.

On the **Workspaces** page, locate the target workspace, and click  in the **Status** column to disable the workspace. In the displayed **Disable Workspace** dialog box, learn about the impact of disabling the workspace and click **Yes** to disable the workspace.

NOTE

If you disable a workspace, you cannot access the workspace, edit the workspace, view quotas of the workspace, or schedule jobs in the workspace.

CDM clusters in the workspace will still be billed.

- **Enable a workspace:** On the **Workspaces** page, locate the target workspace, and click  in the **Status** column to enable the workspace. In the displayed **Enable Workspace** dialog box, click **Yes**.
- **Edit a workspace:** On the **Workspaces** page, locate the target workspace, and click **Edit** in the **Operation** column. In the displayed **Workspace Information** dialog box, modify the workspace parameters listed in [Table 4-1](#) and click **OK**.

Different from creating a workspace, when editing a workspace, you cannot change tags (see [Add/Edit a tag](#)), but can add workspace members (see

Adding a Member and Assigning a Role) and set workspace quotas (see **Setting Workspace Quotas**).

- Add/Edit a tag: On the **Workspaces** page, locate a workspace, click **More** in the **Operation** column, and select **tags**. In the displayed **tags** dialog box, click **Add/Edit Tag** to associate tags with the workspace.
- View quota usage: On the **Workspaces** page, locate a workspace and click **Quota Usage** in the **Operation** column. In the displayed **Quota Usage** dialog box, you can view the quota usage of each service in the workspace.
- Pin a workspace to top: On the **Workspaces** tab page, locate the target workspace, click **More** in the **Operation** column, and select **Pin to Top**.
- Delete a workspace: On the **Workspaces** page, locate the target workspace, click **More** in the **Operation** column, and select **Delete**. In the displayed **Delete Workspace** dialog box, click **OK**.

 **NOTE**

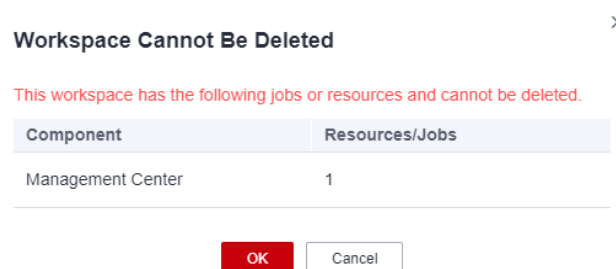
To prevent service loss caused by mis-deletion, the system allows only the **DAYU Administrator** or **Tenant Administrator** to delete workspaces. Before deleting a workspace, ensure that no resource is available in any of the following modules:

- Management Center: data connections
- DataArts Migration: CDM clusters
- DataArts Architecture: subjects, logical models, standards, physical models, dimensional models, and metrics
- DataArts Factory: jobs, job directories, scripts, script directories, and resources
- DataArts Quality: quality jobs and comparison jobs
- DataArts Catalog: tables and files in technical assets and metadata collection tasks
- DataArts DataService: clusters, APIs, and apps
- DataArts Security: sensitive data discovery tasks, masking policies, static masking tasks, and data watermarking tasks

If any component of a workspace has any service resource, deleting the workspace will fail.

If any module has resources, delete the resources as prompted and try again.

Figure 4-2 Message indicating that the workspace cannot be deleted



4.2 Setting Workspace Quotas

Before using DataArts Studio, you need to set quotas for the current workspace (only API quota for DataArts DataService Exclusive). If the used quota of the

current workspace exceeds the allocated quota or the total used quota exceeds the total allocated quota, some functions will be unavailable, for example, you cannot create APIs in DataArts DataService Exclusive.

- Used quota: quota that has been used in the current workspace. It is automatically collected by the system.
- Allocated quota: quota allocated to the current workspace by the administrator
- Total used quota: quota that has been used in the current instance. It is automatically collected by the system.
- Total allocated quota: quota allocated to all workspaces in the current instance. It is automatically collected by the system.
- Total quota: maximum quota for the current instance. The value is fixed and cannot be changed.

Prerequisites

You have changed your account to either of the following:

- **DAYU Administrator or Tenant Administrator**
- **DAYU User**, which is the workspace administrator

Procedure

- Step 1** Log in to the DataArts Studio console by following the instructions in [Accessing the DataArts Studio Instance Console](#).
- Step 2** On the **Workspaces** page, locate the target workspace and click **Edit** in the **Operation** column.

Figure 4-3 Workspace Information dialog box

Workspace Information ×

* Name: default

Description: Enter a description. 0/4,096

* Mode: Simple Upgrade

* Enterprise Project: default C

Job Log Path: Select

API Quota of DataArts: Used: 9

DataService Exclusive: Allocated: 10 Save

Total used: 9

Total allocated: 10

Total: 6,000

* Workspace Members

Account	Account ...	Role	Added	Operation
<input type="checkbox"/>	User	admin	Feb 20, 2024 16:07:24 GMT+0...	Edit
<input type="checkbox"/>	User	admin	Jan 27, 2024 16:33:00 GMT+0...	Edit
<input type="checkbox"/>	User	admin	Jan 25, 2024 19:41:42 GMT+0...	Edit
<input type="checkbox"/>	User	admin	Jan 18, 2024 14:47:06 GMT+0...	Edit

OK Cancel

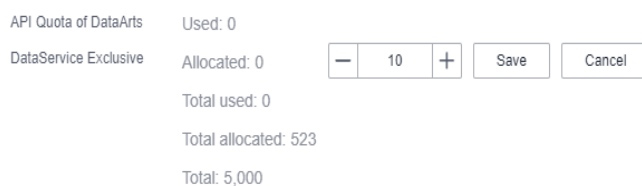
Step 3 In the displayed **Workspace Information** dialog box, locate **API Quota of DataArts DataService Exclusive** and click **Edit** in the **Operation** column. Set the quota and click **Save**.

Allocated indicates the quota that has been allocated to the current workspace. The allocated quota cannot be less than the used quota or greater than the unallocated quota (that is, the total quota minus the total allocated quota).

 **NOTE**

You can create 10 DataArts DataService Exclusive APIs for free in each DataArts Studio instance, and you will be charged for each extra API.

Figure 4-4 Setting the allocated quota



API Quota of DataArts	Used: 0
DataService Exclusive	Allocated: 0

Total used: 0
Total allocated: 523
Total: 5,000

Step 4 After setting the **Allocated** parameter, click **OK** in the **Workspace Information** dialog box.

----End

4.3 (Optional) Changing the Job Log Storage Path

By default, job logs and DLI dirty data are stored in an OBS bucket named **dlf-log-*{project ID}***. You can customize a log storage path. You can configure an OBS bucket globally based on the workspace.

Constraints

- This function depends on the OBS service.
- The OBS path is only supported for OBS buckets and not for parallel file systems.

Prerequisites

You have changed your account to either of the following:

- **DAYU Administrator** or **Tenant Administrator**
- **DAYU User**, which is the workspace administrator

Procedure

Step 1 Log in to the DataArts Studio console by following the instructions in [Accessing the DataArts Studio Instance Console](#).

Step 2 On the **Workspaces** page, locate the target workspace and click **Edit** in the **Operation** column.

Figure 4-5 Workspace Information dialog box

Workspace Information [Close]

* Name:

Description: 0/4,096

* Mode: Upgrade

* Enterprise Project: Clear

Job Log Path Select

API Quota of DataArts DataService Exclusive
 Used: 9
 Allocated: 10 Save
 Total used: 9
 Total allocated: 10
 Total: 6,000

* Workspace Members

Add Remove Search

<input type="checkbox"/>	Account	Account ...	Role	Added	Operation
<input type="checkbox"/>	User		admin	Feb 20, 2024 16:07:24 GMT+0...	Edit
<input type="checkbox"/>	User		admin	Jan 27, 2024 16:33:00 GMT+0...	Edit
<input type="checkbox"/>	User		admin	Jan 25, 2024 19:41:42 GMT+0...	Edit
<input type="checkbox"/>	User		admin	Jan 18, 2024 14:47:06 GMT+0...	Edit

OK Cancel

Step 3 In the **Workspace Information** dialog box, click **Select** next to **Job Log Path** and **Dirty Data Path** to select the path for storing logs and DLI dirty data. You can select a specific directory.

Figure 4-6 Changing the path for storing logs and DLI dirty data

Job Log Path Select

Dirty Data Path Select

Step 4 Click **OK**.

----End

5 Authorizing Users to Use DataArts Studio

5.1 Creating an IAM User and Assigning DataArts Studio Permissions

Identity and Access Management (IAM) can be used for fine-grained permissions management on your DataArts Studio resources. With IAM, you can:

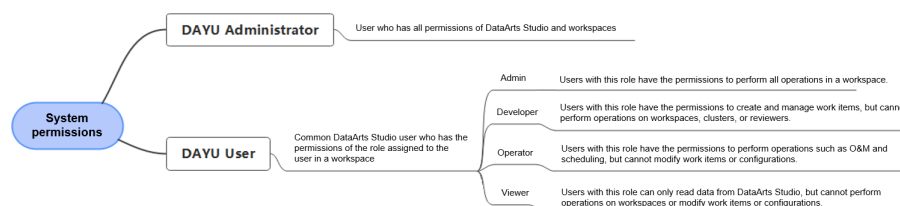
- Create IAM users for employees based on your enterprise's organizational structure. Each IAM user will have their own security credentials for accessing DataArts Studio resources.
- Assign users only the permissions required to perform a task.
- Entrust a Huawei account or cloud service to perform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

If you do not require individual IAM users for permissions management, skip this section.

Background

- Before assigning permissions to a user group, you need to understand the DataArts Studio permission system and select permissions as needed. For details about the permissions supported by DataArts Studio, see [DataArts Studio Permissions Management](#).

Figure 5-1 Permission system



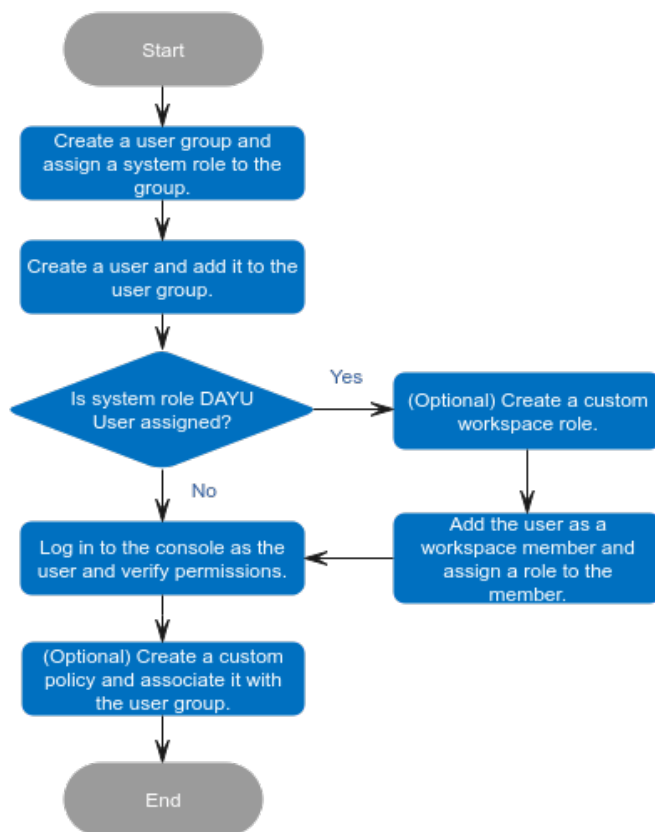
- For details on the system policies of other services, see [Permissions Policies](#).

Constraints

- The DAYU User role provides permissions related to instances, workspaces, and dependent services. Workspace roles provide service operation permissions in workspaces.
- IAM provides the following two authorization mechanisms: Note that DataArts Studio supports only the IAM role-based authorization and does not support the IAM policy-based authorization.
 - **IAM Roles:** IAM initially provides a coarse-grained authorization mechanism to define permissions based on users' job responsibilities. This mechanism provides only a limited number of service-level roles for authorization. However, traditional IAM roles are not an ideal choice for fine-grained authorization and secure access control.
 - **IAM Policies:** A type of fine-grained authorization mechanism that defines permissions required to perform operations on specific cloud resources under certain conditions. This type of authorization is more flexible and is ideal for least privilege access.

Procedure

Figure 5-2 Authorization process



Step 1 Create a user group and assign a system role to the group.

Log in to the IAM console using a Huawei account, create a user group and assign a DataArts Studio system role to the group. For example, the system role can be DAYU Administrator or DAYU User.

For details, see [Creating a User Group and Assigning Permissions](#).

 NOTE

- When configuring DataArts Studio permissions for a user group, enter **DAYU** in the search box to search for the permissions and select the permissions to be granted to the user group, for example, **DAYU User**.
- DataArts Studio is a project-level service deployed in specific physical regions. If you select **All resources** for **Scope**, the permission takes effect in all projects of all regions. If you select **Region-specific projects** for **Scope**, the permission takes effect only for a specified project. When accessing DataArts Studio, the IAM user must switch to the region where they have been assigned the required permissions.

Step 2 Create a user and add it to the user group.

Create users on the IAM console and add them to the group created in [step 1](#).

For details, see [Creating an IAM User and Adding It to a User Group](#).

 NOTE

An IAM user can pass the authentication and access DataArts Studio through an API or SDK only if **Programmatically access** is selected for **Access Type** during the creation of the IAM user.

Step 3 Create a custom workspace role for DAYU User, add it as a workspace member, and assign a role to the member.

DataArts Studio workspace roles determine the permissions of DAYU User in a workspace. There are five preset roles: admin, developer, deployer, operator, and viewer. For details about how to add a workspace member and assign a role to the member, see [Adding a Member and Assigning a Role](#).

For details about the permissions of the roles, see [DataArts Studio Permissions](#).

Step 4 Log in to the console and verify permissions.

Log in to the console using the created user and verify permissions of the user.

- Choose **Service List > DataArts Studio**. Locate a DataArts Studio instance and click **Access**. Check whether the workspace list is displayed.
- Access a service module (for example, Management Center) to which your current user has been added and check whether you can perform the operations allowed for the workspace role assigned to you.

----End

Follow-up Operations

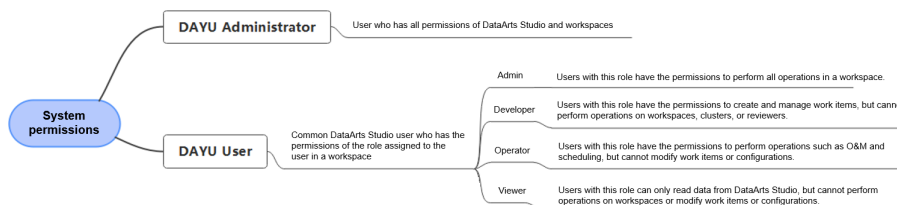
- The preset DAYU User has excessive permissions of dependent services, which may cause security risks. You can manually remove unnecessary permissions to comply with the principle of least privilege (PoLP). For details, see [Authorizing IAM Users to Use DataArts Studio by Complying with the Principle of Least Privilege](#).

5.2 Adding a Member and Assigning a Role

For IAM users with the **DAYU User** permission, a workspace role determines their permissions in the workspace. If you want to allow another IAM user with the

DAYU User permission to use your DataArts Studio instance, create an IAM user by referring to [Creating an IAM User and Assigning DataArts Studio Permissions](#), and add the user as the workspace member and configure a role for the user by following the instructions in this section.

Figure 5-3 Permission system



A workspace role determines the permissions of a user in the workspace. Currently, the following preset roles are available: admin, developer, deployer, operator, and viewer. For details about the permissions of each role, see [Permissions](#).

- **Admin:** This role has all operation permissions in a workspace. You are advised to assign the admin role to the project owner, development owner, and O&M administrator.
- **Developer:** This role has permissions to create and manage resources in a workspace. You are advised to assign this role to users who develop and process tasks.
- **Operator:** This role has the operation permissions of services such as O&M and scheduling in a workspace, but cannot modify resources or configurations. You are advised to assign this role to users responsible for O&M management and status monitoring.
- **Viewer:** This role can view data in a workspace but cannot perform any other operation. You are advised to assign this role to users who only need to view data in a workspace but do not need to perform operations.
- **Deployer:** This role is unique to the enterprise mode and has permissions to release task packages in a workspace. In enterprise mode, when a developer submits a script or job version, the system generates a release task. After the developer confirms the release and the deployer approves the release request, the modified job is synchronized to the production environment.

Background

If the permissions of **DAYU User** has been assigned to the created IAM user, you must add workspace members and roles. Otherwise, the IAM user cannot view existing DataArts Studio workspaces.

Constraints

Due to the limitation of the authentication cache mechanism, changes to the roles of workspace members do not take effect immediately. Instead, the changes take effect only after the workspace members stop accessing the DataArts Studio console and wait for 6 minutes.

Prerequisites

You have changed your account to either of the following:

- **DAYU Administrator** or **Tenant Administrator**
- **DAYU User**, which is the workspace administrator

Adding a Member and Assigning a Role

Step 1 Log in to the DataArts Studio console by following the instructions in [Accessing the DataArts Studio Instance Console](#).

Step 2 On the **Workspaces** page, locate the target workspace and click **Edit** in the **Operation** column.

Figure 5-4 Workspace Information dialog box

Workspace Information [Close]

* Name:

Description: (0/4,096)

* Mode:

* Enterprise Project:

Job Log Path:

API Quota of DataArts DataService Exclusive: Used: 9, Allocated: 10, Total used: 9, Total allocated: 10, Total: 6,000.

* Workspace Members:

<input type="checkbox"/>	Account	Account ...	Role	Added	Operation
<input type="checkbox"/>	User		admin	Feb 20, 2024 16:07:24 GMT+0...	Edit
<input type="checkbox"/>	User		admin	Jan 27, 2024 16:33:00 GMT+0...	Edit
<input type="checkbox"/>	User		admin	Jan 25, 2024 19:41:42 GMT+0...	Edit
<input type="checkbox"/>	User		admin	Jan 18, 2024 14:47:06 GMT+0...	Edit

Step 3 Click **Add** under **Workspace Members**. In the displayed **Add Member** dialog box, select **Add User** or **Add Group**, select a member account from the drop-down list, and select a role for it.

Figure 5-5 Adding a member

Add Member [Close]

* Account Type: User Group

* Member Account:

* Role: admin developer operator viewer

Step 4 Click **OK**. You can view or modify the members and roles in the member list, or remove members from the workspace.

----End

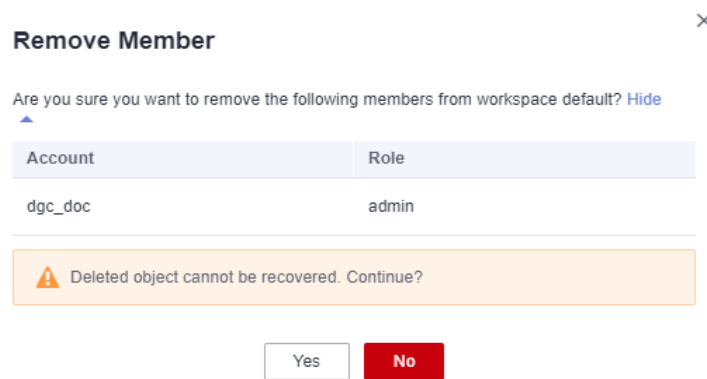
Related Operations

- Remove workspace members: In the **Workspace Information** dialog box, select the members to remove and click **Remove**. In the **Remove Member** dialog box, click **Yes**.

NOTE

The creator of a workspace cannot be removed.

Figure 5-6 Removing members



6 (Optional) Obtaining Authentication Information

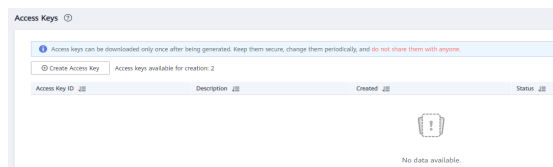
When using DataArts Studio, for example, when creating OBS links, calling APIs, or locating issues, you may need to obtain information such as access keys, project IDs, and endpoints. This section describes how to obtain such information.

Obtaining an Access Key

To obtain an access key, perform the following steps:

1. Log in to the management console, move the cursor to the username in the upper right corner, and select **My Credentials** from the drop-down list.
2. On the **My Credentials** page, choose **Access Keys**, and click **Create Access Key**. See [Figure 6-1](#).

Figure 6-1 Clicking Create Access Key



3. Click **OK** and save the access key file as prompted. The access key file will be saved to your browser's configured download location. Open the **credentials.csv** file to view **Access Key Id** and **Secret Access Key**.

NOTE

- Only two access keys can be added for each user.
- To ensure access key security, the access key is automatically downloaded only when it is generated for the first time and cannot be obtained from the management console later. Keep them properly.

Obtaining a Project ID and Account ID

A project is a group of tenant resources, and an account ID corresponds to the current account. The IAM ID corresponds to the current user. You can view the project IDs, account IDs, and user IDs in different regions on the corresponding pages.

1. Register with and log in to the management console.
2. Hover the cursor on the username in the upper right corner and select **My Credentials** from the drop-down list.
3. On the **API Credentials** page, obtain the account name, account ID, IAM username, and IAM user ID, and obtain the project ID from the project list.

Obtaining a DataArts Studio Instance ID and Workspace ID

To obtain a DataArts Studio instance ID and workspace ID, perform the following steps:

1. On the DataArts Studio console, locate a workspace and click any module, such as **Management Center**.
2. On the **Management Center** page, obtain the values of **instanceId** and **workspace** in the browser address bar, which are the instance ID and workspace ID, respectively.

As shown in [Figure 6-2](#), the instance ID is **6b88...2688**, and the workspace ID is **1dd3bc...d93f0**.

Figure 6-2 Obtaining the instance ID and workspace ID



Obtaining an Endpoint

An endpoint is the **request address** for calling an API. Endpoints vary depending on services and regions.

DataArts Studio endpoints are in the following formats. Select an endpoint based on your service requirements.

- DataArts Migration (or CDM): `cdm.{region_id}.myhuaweicloud.com`. See [Table 6-1](#) for details.
- DataArts Factory: `dayu-dlf.{region_id}.myhuaweicloud.com`. See [Table 6-2](#) for details.

Table 6-1 CDM endpoints

Region Name	Region ID	Endpoint	Protocol
CN-Hong Kong	ap-southeast-1	cdm.ap-southeast-1.myhuaweicloud.com	HTTPS
AP-Bangkok	ap-southeast-2	cdm.ap-southeast-2.myhuaweicloud.com	HTTPS
AP-Singapore	ap-southeast-3	cdm.ap-southeast-3.myhuaweicloud.com	HTTPS
LA-Sao Paulo1	sa-brazil-1	cdm.sa-brazil-1.myhuaweicloud.com	HTTPS
LA-Santiago	la-south-2	cdm.la-south-2.myhuaweicloud.com	HTTPS

Region Name	Region ID	Endpoint	Protocol
LA-Mexico City1	na-mexico-1	cdm.na-mexico-1.myhuaweicloud.com	HTTPS

Table 6-2 DataArts Factory endpoints

Region Name	Region ID	Endpoint	Protocol
CN-Hong Kong	ap-southeast-1	dayu-dlf.ap-southeast-1.myhuaweicloud.com	HTTPS
AP-Bangkok	ap-southeast-2	dayu-dlf.ap-southeast-2.myhuaweicloud.com	HTTPS
AP-Singapore	ap-southeast-3	dayu-dlf.ap-southeast-3.myhuaweicloud.com	HTTPS
LA-Sao Paulo1	sa-brazil-1	dayu-dlf.sa-brazil-1.myhuaweicloud.com	HTTPS
LA-Santiago	la-south-2	dayu-dlf.la-south-2.myhuaweicloud.com	HTTPS
LA-Mexico City1	na-mexico-1	dayu-dlf.na-mexico-1.myhuaweicloud.com	HTTPS

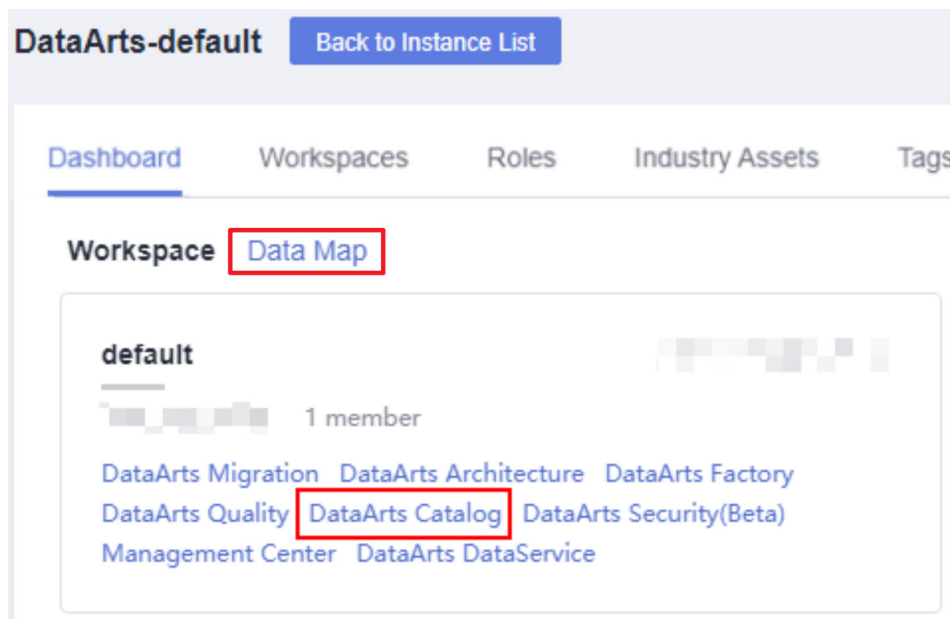
Obtaining the GUID of DataArts Catalog

Each logical, technical, or metric asset has a GUID, which is the unique identifier of the asset. Some URLs in the calls to DataArts Catalog or Data Map APIs require the GUID.

To obtain the GUID of a data asset in DataArts Catalog or Data Map, perform the following operations:

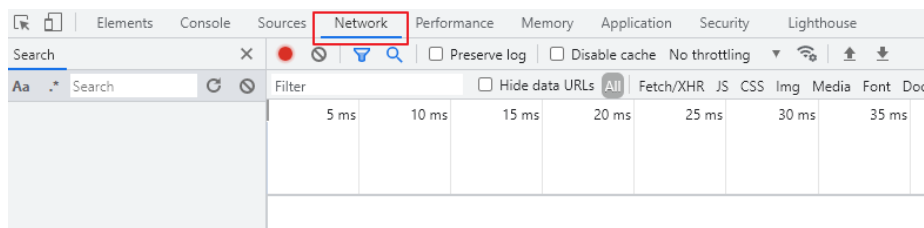
1. On the DataArts Studio console, locate an instance and click **Access**. On the displayed page, click **Data Map** or locate a workspace and click **DataArts Catalog**.

Figure 6-3 Clicking Data Map or DataArts Catalog



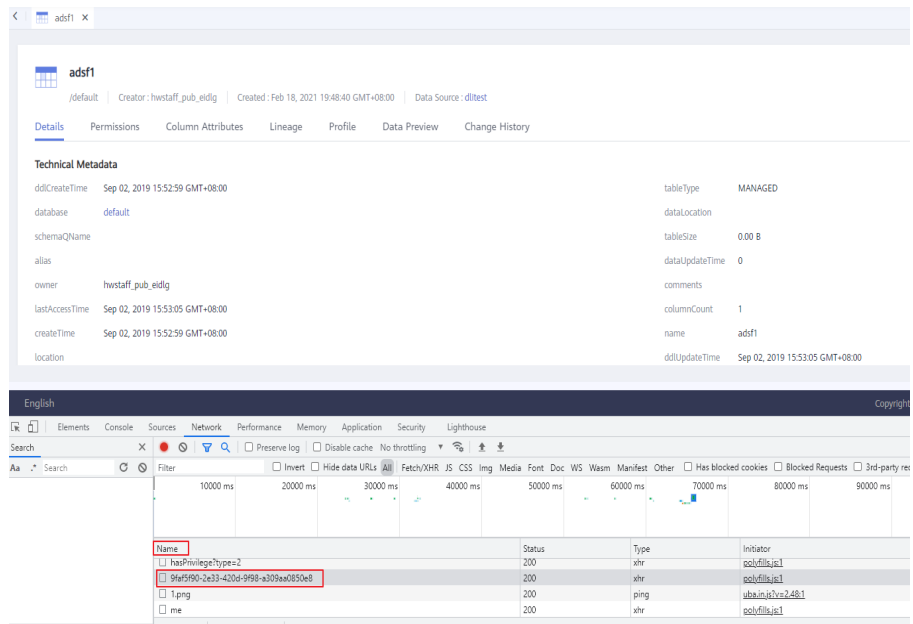
2. Press **F12** to open the developer debugging tool and click the **Network** tab.

Figure 6-4 Network



3. On the **Home** or **Data Search** page of the Data Map console, or on the **Dashboard** or **Data Catalog** page of the DataArts Catalog console, locate an asset and click its name to access its details page.
 - a. In the network request, search for a long string whose name is similar to **09318f28-939f-4ab6-a374-9e621096652c**.

Figure 6-5 Searching for a long string



b. Click the string and obtain the GUID from **Request URL**.

Figure 6-6 Obtaining the GUID

