

Alibaba Cloud

DataWorks
Preparation

Document Version: 20220627

Legal disclaimer

Alibaba Cloud reminds you to carefully read and fully understand the terms and conditions of this legal disclaimer before you read or use this document. If you have read or used this document, it shall be deemed as your total acceptance of this legal disclaimer.

1. You shall download and obtain this document from the Alibaba Cloud website or other Alibaba Cloud-authorized channels, and use this document for your own legal business activities only. The content of this document is considered confidential information of Alibaba Cloud. You shall strictly abide by the confidentiality obligations. No part of this document shall be disclosed or provided to any third party for use without the prior written consent of Alibaba Cloud.
2. No part of this document shall be excerpted, translated, reproduced, transmitted, or disseminated by any organization, company or individual in any form or by any means without the prior written consent of Alibaba Cloud.
3. The content of this document may be changed because of product version upgrade, adjustment, or other reasons. Alibaba Cloud reserves the right to modify the content of this document without notice and an updated version of this document will be released through Alibaba Cloud-authorized channels from time to time. You should pay attention to the version changes of this document as they occur and download and obtain the most up-to-date version of this document from Alibaba Cloud-authorized channels.
4. This document serves only as a reference guide for your use of Alibaba Cloud products and services. Alibaba Cloud provides this document based on the "status quo", "being defective", and "existing functions" of its products and services. Alibaba Cloud makes every effort to provide relevant operational guidance based on existing technologies. However, Alibaba Cloud hereby makes a clear statement that it in no way guarantees the accuracy, integrity, applicability, and reliability of the content of this document, either explicitly or implicitly. Alibaba Cloud shall not take legal responsibility for any errors or lost profits incurred by any organization, company, or individual arising from download, use, or trust in this document. Alibaba Cloud shall not, under any circumstances, take responsibility for any indirect, consequential, punitive, contingent, special, or punitive damages, including lost profits arising from the use or trust in this document (even if Alibaba Cloud has been notified of the possibility of such a loss).
5. By law, all the contents in Alibaba Cloud documents, including but not limited to pictures, architecture design, page layout, and text description, are intellectual property of Alibaba Cloud and/or its affiliates. This intellectual property includes, but is not limited to, trademark rights, patent rights, copyrights, and trade secrets. No part of this document shall be used, modified, reproduced, publicly transmitted, changed, disseminated, distributed, or published without the prior written consent of Alibaba Cloud and/or its affiliates. The names owned by Alibaba Cloud shall not be used, published, or reproduced for marketing, advertising, promotion, or other purposes without the prior written consent of Alibaba Cloud. The names owned by Alibaba Cloud include, but are not limited to, "Alibaba Cloud", "Aliyun", "HiChina", and other brands of Alibaba Cloud and/or its affiliates, which appear separately or in combination, as well as the auxiliary signs and patterns of the preceding brands, or anything similar to the company names, trade names, trademarks, product or service names, domain names, patterns, logos, marks, signs, or special descriptions that third parties identify as Alibaba Cloud and/or its affiliates.
6. Please directly contact Alibaba Cloud for any errors of this document.

Document conventions

Style	Description	Example
 Danger	A danger notice indicates a situation that will cause major system changes, faults, physical injuries, and other adverse results.	 Danger: Resetting will result in the loss of user configuration data.
 Warning	A warning notice indicates a situation that may cause major system changes, faults, physical injuries, and other adverse results.	 Warning: Restarting will cause business interruption. About 10 minutes are required to restart an instance.
 Notice	A caution notice indicates warning information, supplementary instructions, and other content that the user must understand.	 Notice: If the weight is set to 0, the server no longer receives new requests.
 Note	A note indicates supplemental instructions, best practices, tips, and other content.	 Note: You can use Ctrl + A to select all files.
>	Closing angle brackets are used to indicate a multi-level menu cascade.	Click Settings > Network > Set network type .
Bold	Bold formatting is used for buttons, menus, page names, and other UI elements.	Click OK .
Courier font	Courier font is used for commands	Run the <code>cd /d C:/window</code> command to enter the Windows system folder.
<i>Italic</i>	Italic formatting is used for parameters and variables.	<code>bae log list --instanceid</code> <i>Instance_ID</i>
[] or [a b]	This format is used for an optional value, where only one item can be selected.	<code>ipconfig [-all -t]</code>
{ } or {a b}	This format is used for a required value, where only one item can be selected.	<code>switch {active stand}</code>

Table of Contents

1. Activate the DataWorks service and log on to the DataWorks console	05
1.1. Activate DataWorks	05
1.2. Prepare a RAM user	07
1.3. Use a RAM user to log on to the DataWorks console and ...	10
1.4. (Advanced) Use a RAM role to log on to the DataWorks console	12
2. Create a workspace	20
3. Add workspace members and assign roles to them	30
4. Work as a workspace administrator	35

1. Activate the DataWorks service and log on to the DataWorks console

1.1. Activate DataWorks

DataWorks provides an all-in-one professional, efficient, secure, and reliable platform for big data development and governance. Before you use DataWorks, you must activate the service by using your Alibaba Cloud account. After the service is activated, you can authorize Resource Access Management (RAM) users to use DataWorks. This topic describes how to activate DataWorks.

Context

You can log on to DataWorks by using your Alibaba Cloud account or as a [RAM](#) user. The logon method is the same as that for other Alibaba Cloud services.

- An Alibaba Cloud account owns Alibaba Cloud resources. Resource usage is measured and billed to the Alibaba Cloud account. You can use the Alibaba Cloud account to create RAM users for your enterprise. You can also use the Alibaba Cloud account to manage and grant permissions to RAM users.
- RAM users are created and managed by an Alibaba Cloud account in the RAM console. RAM users do not own resources, so the resource usage of a RAM user is not measured or billed to the RAM user. The Alibaba Cloud account centrally manages all RAM users and pays for the resources used by RAM users.

Therefore, before you use DataWorks, you must have an Alibaba Cloud account and manage your RAM users in the RAM console.

Step 1: Create an Alibaba Cloud account

Only an Alibaba Cloud account or a RAM user that has the `AliyunBSSOrderAccess` or `AliyunDataWorksFullAccess` permission can activate DataWorks. If you do not have an Alibaba Cloud account, perform the following operations to create an Alibaba Cloud account and complete security configurations such as real-name verification for your account. These operations are required to activate DataWorks.

If you have an Alibaba Cloud account that meets the requirements, skip this step and perform subsequent steps.

- For information about how to create an Alibaba Cloud account, see [Sign up with Alibaba Cloud](#).
- For information about how to complete real-name verification for an Alibaba Cloud account, see [Real-name registration FAQs](#).

 **Note**

- To ensure smooth subsequent operations, you must complete real-name verification for your Alibaba Cloud account. If you are an enterprise user, we recommend that you complete enterprise-level verification to enjoy more benefits. We also recommend that you complete the security configurations mentioned in the referenced topics to ensure account security.
- The Alibaba Cloud account that you create is a privileged account to which resource usage is billed. To keep your account safe, we recommend that you regularly change your password and do not share your account with others.

Step 2: Activate DataWorks

1. Log on to the [Alibaba Cloud international site \(alibabacloud.com\)](https://alibabacloud.com) by using your Alibaba Cloud account. Go to the [DataWorks homepage](#) and click **Buy Now**.

DataWorks provides Basic Edition and four advanced editions: Standard Edition, Professional Edition, Enterprise Edition, and Ultimate Edition. The editions differ in terms of supported features and prices. You can view details of each edition in the [Differences among DataWorks editions](#) topic and select the edition that meets your needs.

2. Select the region in which you want to activate DataWorks.

DataWorks is available in multiple regions in and outside China. You can select a region based on your needs.

3. Learn more about the pay-as-you-go billable items and billing details.

The buy page displays the billing standards of DataWorks services that are billed in pay-as-you-go mode. You can click the corresponding links for more details.

4. Click **Buy Now**.

5. Follow the instructions to confirm your settings. After you read and agree to the service agreement of DataWorks, click **Activate Now** and complete the payment.

Step 3: Log on to DataWorks and use the services

After you activate DataWorks, you can use the DataWorks services such as DataStudio, Data Integration, and Data Governance Center. Generally, the development and governance of big data require collaborative efforts from multiple roles. For example, developers perform operations such as data development and data integration, data governance engineers manage data, and administrators manage permissions and configurations.

- Log on to the DataWorks console.

- Log on to the DataWorks console as a RAM user.

To facilitate collaboration and meet the requirements of account security and permission isolation, you can use an Alibaba Cloud account to create multiple RAM users and grant different permissions to the users. The credentials of RAM users can be allocated to different engineers so that these engineers can independently log on to the DataWorks console. For more information about how to use DataWorks as a RAM user, see [Use a RAM user to log on to the DataWorks console and use DataWorks](#).

- Create or join a DataWorks workspace.

Workspaces are basic units for managing tasks, members, roles, and permissions in DataWorks. A workspace administrator can add members to a workspace and assign the Workspace Manager, Development, O&M, Deploy, Safety Manager, or Visitor role to each member. This way, the workspace members with different roles can collaborate with each other. Therefore, before you use DataWorks services such as DataStudio, you must create a workspace or apply to join a workspace.

DataWorks provides a variety of workspace modes and permission models.

- You can create a workspace in **standard mode** or **basic mode**. A workspace in standard mode isolates the development environment from the production environment, whereas a workspace in basic mode does not. For more information about the differences between the two modes, see [Basic mode and standard mode](#).
- You can manage permissions by using the permission models provided by DataWorks. The permission models allow you to grant different permissions to different roles. For more information, see [Overview](#).

For more information about how to create a workspace and add members to a workspace, see [Create a workspace](#) and [Add workspace members and assign roles to them](#).

- Use DataWorks services.

DataWorks provides an all-in-one data development and management console that integrates a variety of services, including Data Integration, DataStudio, Operation Center, DataAnalysis, Data Modeling, and Data Governance Center. For more information about the services, see related topics in the DataWorks documentation.

1.2. Prepare a RAM user

This topic describes how to create a RAM user, create an AccessKey pair for the RAM user, grant permissions to the RAM user, and allow a user to use the RAM user to access DataWorks.

Context

- If you want to access DataWorks alone, prepare your Alibaba Cloud account by following the instructions that are described in [Activate DataWorks](#) and skip the operations that are described in this topic.
- If you want to work with other users to access DataWorks, prepare RAM users by following the instructions that are described in this topic.

Create a RAM user

You need to use your Alibaba Cloud account to log on to the RAM console and create a RAM user.

1. Log on to the [RAM console](#) by using your Alibaba Cloud account.
2. In the left-side navigation pane, choose **Identities > Users**.
3. On the **Users** page, click **Create User**.
4. In the **User Account Information** section of the **Create User** page, configure the **Logon Name** and **Display Name** parameters.

 **Note** You can click **Add User** to create multiple RAM users at a time.

5. In the **Access Mode** section, select an access mode.
 - **Console Access**: If you select this option, you must complete the logon security settings. These

settings specify whether to use a system-generated or custom logon password, whether the password must be reset upon the next logon, and whether to enable multi-factor authentication (MFA).

 **Note** If you select Custom Logon Password in the Console Password section, you must specify a password. The password must meet the complexity requirements. For more information about the complexity requirements, see [Configure a password policy for RAM users](#).

- **OpenAPI Access:** If you select this option, an AccessKey pair is automatically created for the RAM user. The RAM user can call API operations or use other development tools to access Alibaba Cloud resources.

 **Note** To ensure the security of the Alibaba Cloud account, we recommend that you select only one access mode for the RAM user. This prevents the RAM user from using an AccessKey pair to access Alibaba Cloud resources after the RAM user leaves the organization.

6. Click **OK**.

 **Notice** After you create a RAM user, keep the username and password of the RAM user secure and send the logon information to the user that is allowed to access DataWorks by using the RAM user.

(Optional) Create an AccessKey pair for a RAM user

An AccessKey pair is not required if you use a RAM user only to run nodes in DataWorks. If you have special business requirements, you can create an AccessKey pair for the RAM user in the **RAM console**.

If the RAM user is granted permissions to manage AccessKey pairs, you can use the RAM user to create an AccessKey pair in the RAM console.

To create an AccessKey pair for a RAM user, perform the following steps:

1. Log on to the [RAM console](#) by using your Alibaba Cloud account.
2. In the left-side navigation pane, choose **Identities > Users**.
3. On the **Users** page, click the username of a specific RAM user.
4. In the **User AccessKeys** section, click **Create AccessKey Pair**.
5. In the **Create AccessKey Pair** dialog box, view the AccessKey ID and AccessKey secret.

You can click **Download CSV File** to download the AccessKey pair or click **Copy** to copy the AccessKey pair.

6. Click **Close**.

 **Note**

- The AccessKey secret is displayed only when you create an AccessKey pair, and is unavailable for subsequent queries. We recommend that you save the AccessKey secret for subsequent use.
- If the AccessKey pair is disclosed or lost, you must create another AccessKey pair. You can create a maximum of two AccessKey pairs.

Grant a RAM user the permissions to perform operations in the DataWorks console

If you want to grant a RAM user the permissions to perform operations in the DataWorks console, you must log on to the [RAM console](#) and grant the permissions to the RAM user.

Perform the following steps to grant the permissions to the RAM user:

Allow a user to use DataWorks as a RAM user

If you need to work with other users for collaborative development in the DataWorks console, you need to prepare RAM users and provide the users with the information that can be used to log on to the Alibaba Cloud Management Console as the RAM users. If an AccessKey pair is required for a RAM user that you prepared, you must provide the AccessKey pair that is available to a user that uses DataWorks as the RAM user.

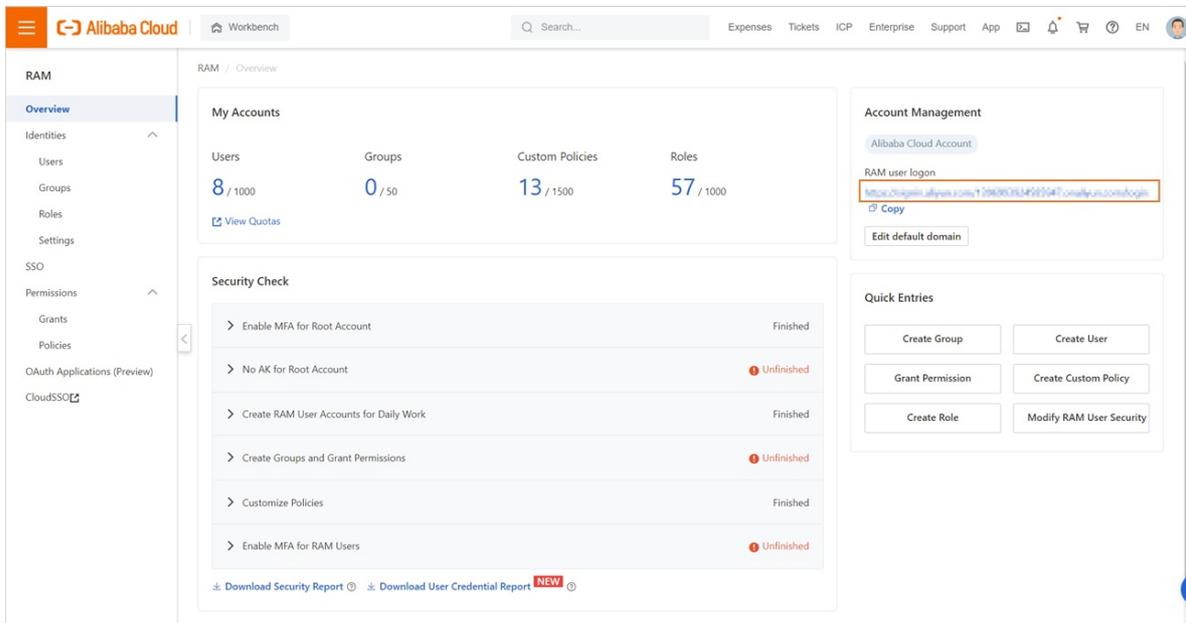
Note

- A RAM user belongs to an Alibaba Cloud account and does not own resources. Resource usage of a RAM user is not measured and billed.
- All charges generated when you use Alibaba Cloud services as a RAM user are paid by your Alibaba Cloud account.
- You must obtain the link that can be used to log on to the Alibaba Cloud Management Console as a RAM user and the default domain name or domain alias of your Alibaba Cloud account. Then, you must send the information to a user that is allowed to access DataWorks as a RAM user.

You must provide the following information to a user that is allowed to access DataWorks as a RAM user:

- Link that can be used to log on to the Alibaba Cloud Management Console as a RAM user

Log on to the **RAM console** by using your Alibaba Cloud account. In the left-side navigation pane, click **Overview**. In the upper-right corner of the Overview page, click **Copy** on the right side of the logon link specified by the **RAM user logon** parameter in the Account Management section and provide the link for a user that is allowed to access DataWorks as the RAM user. For more information about how to log on to the Alibaba Cloud Management Console as a RAM user, see [Log on to the Alibaba Cloud Management Console as a RAM user](#).



- Domain alias or default domain name of your Alibaba Cloud account

Log on to the **RAM console** by using your Alibaba Cloud account. In the left-side navigation pane, choose **Identities > Settings**. On the Settings page, click the **Advanced** tab to view the **default domain name** and **domain alias**.

- Username and password of the RAM user
- AccessKey ID and AccessKey secret of the RAM user

In addition to providing the preceding information, take note of the following points:

- The permissions to log on to the Alibaba Cloud Management Console are granted to the RAM user.
- The permissions to manage AccessKey pairs are granted to the RAM user. For more information, see [Configure security policies for RAM users](#).

What's next

After you prepare a RAM user, you can use the RAM user to log on to the DataWorks console, create a workspace, and perform operations such as data development in the workspace. For more information, see [Create a workspace](#).

1.3. Use a RAM user to log on to the DataWorks console and use DataWorks

This topic describes how to use a RAM user to log on to the DataWorks console, update the personal information of an alert contact, and create a workspace.

Prerequisites

The first time that you use a RAM user to log on to the DTplus console and use DataWorks, make sure that the following information is confirmed:

- The enterprise alias of the Alibaba Cloud account to which the RAM user belongs.
- The username and password of the RAM user.
- An AccessKey ID and an AccessKey Secret are created for the RAM user by using your Alibaba Cloud account to meet your business requirements. DataWorks no longer requires you to provide the AccessKey pair information before you use DataWorks.
- In the RAM console, you set Console Password Logon to Enabled for the RAM user in the Modify Logon Settings panel of the Users page.

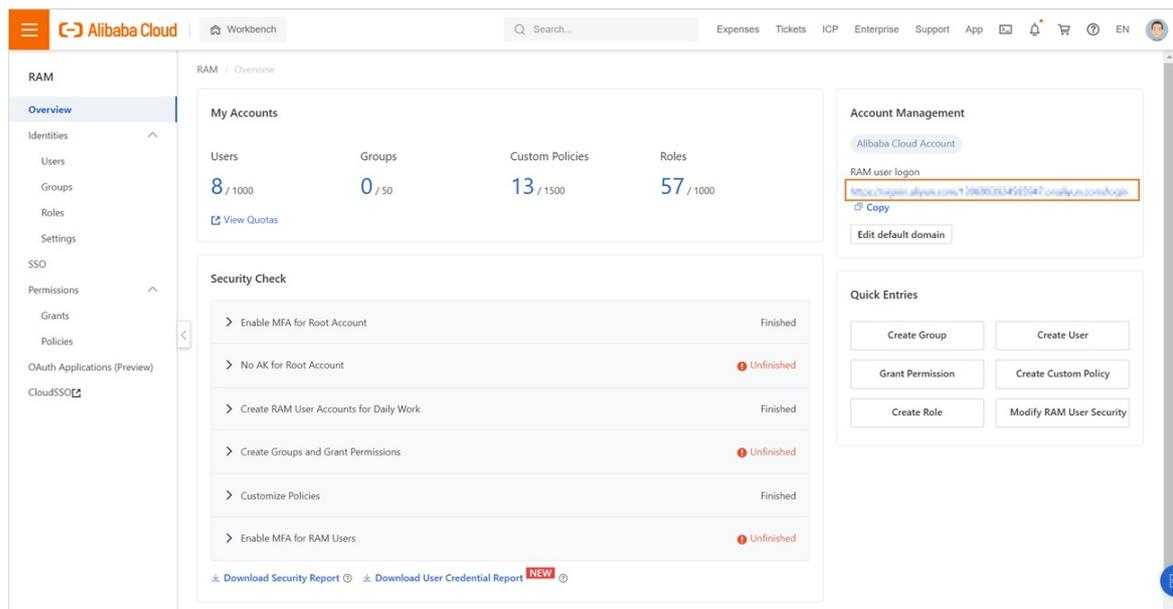
If the preceding information cannot be confirmed, make sure that the Alibaba Cloud account to which the RAM user belongs meets the following requirements:

- All required configurations and information are prepared for the RAM user. For more information, see [Activate DataWorks](#) and [Prepare a RAM user](#).
- A workspace is created. The RAM user is added to the workspace as a member and is assigned the required role. For more information, see [Create a workspace](#) and [Add workspace members and assign roles to them](#).

Log on to the DataWorks console

1. Log on to the Alibaba Cloud Management Console as a RAM user by using the link that is provided by your Alibaba Cloud account.

Log on to the **RAM console** by using your Alibaba Cloud account. In the left-side navigation pane, click Overview. In the upper-right corner of the Overview page, click Copy to the right of the logon link specified by the **RAM user logon** parameter in the Account Management section and provide the link for a user that is allowed to access DataWorks as the RAM user. For more information about how to log on to the Alibaba Cloud Management Console as a RAM user, see [Log on to the Alibaba Cloud Management Console as a RAM user](#).



2. On the homepage of the Alibaba Cloud Management Console, search for DataWorks and go to the DataWorks console.

3. In the left-side navigation pane, click **Workspaces**.

 **Note** After you log on to the DataWorks console as a RAM user, you can view only the workspaces to which you are added as a member on the **Workspaces** page.

If no workspace is displayed on the Workspaces page, contact the administrator of the workspace to which you want to be added. The administrator can add you to the workspace on the **User Management** page of the **Workspace Management** page. For more information, see [Manage workspace-level roles and members](#).

If you want to use the RAM user to create a workspace, attach the AliyunDataWorksFullAccess policy to the RAM user by using your Alibaba Cloud account. For more information, see [How do I grant a RAM user the permissions to create a DataWorks workspace?](#).

4. Find the workspace that is created and click **Data Development** in the Actions column to go to the DataStudio page.

Update personal information

You can configure the personal information of a RAM user on the **Alert Contacts** page of DataWorks. The personal information is used to receive alert notifications that are triggered during data development and O&M in DataWorks. When an alert is triggered based on the alert rules that you configure for a node, the related alert notification is sent to the alert contact that you configure on the Alert Contacts page. This allows the alert contact to identify and handle the exception at the earliest opportunity. This section describes how to view and configure the personal information of an alert contact.

1. Log on to the **DataWorks console**.
2. In the left-side navigation pane, click **Alerts**. Click **Alert Contacts**. For more information about how to configure the personal information of an alert contact, see [Configure and view alert contacts](#).

Create a workspace

To use a RAM user to create a workspace, you must attach the AliyunDataWorksFullAccess policy to the RAM user by using your Alibaba Cloud account. For more information, see [Prepare a RAM user](#) and [Create a workspace](#).

What's next

You can learn more about DataWorks by using the DataWorks tutorial. For more information, see [Workshop introduction](#).

1.4. (Advanced) Use a RAM role to log on to the DataWorks console and use DataWorks

This topic describes how to use role-based single sign-on (SSO) to log on to the DataWorks console and use DataWorks.

Context

Alibaba Cloud allows enterprise users to manage and use cloud resources in the Alibaba Cloud

Management Console. In most cases, enterprise users use their accounts and passwords to log on to the Alibaba Cloud Management Console. The accounts can be Alibaba Cloud accounts or RAM users. However, as the regulatory requirements for enterprise security become increasingly strict, some enterprises prefer to use role-based SSO to log on to the Alibaba Cloud Management Console. For more information, see [Overview](#).

Note

If you use role-based SSO to log on to the DataWorks console, you can use only exclusive resource groups for scheduling to run nodes. For more information about how to create and use an exclusive resource group for scheduling, see [Create and use an exclusive resource group for scheduling](#).

You can use role-based SSO to access only DataWorks workspaces that use the MaxCompute compute engine.

Logon methods supported by DataWorks

You can use one of the following methods to log on to the DataWorks console:

- Use an Alibaba Cloud account or a RAM user to log on to the DataWorks console.

You can use your Alibaba Cloud account or RAM user and the password to log on to the DataWorks console and then use DataWorks.

In this case, the Alibaba Cloud account or RAM user becomes a member of a DataWorks workspace and has the permissions to use the features of DataWorks.

- Use role-based SSO to log on to the DataWorks console.

You can use role-based SSO to log on to the DataWorks console and then use DataWorks. For more information, see [Overview](#).

In this case, the RAM role becomes a member of a DataWorks workspace. The user who assumes the RAM role has the same permissions as an Alibaba Cloud account or a RAM user. For more information, see [RAM role overview](#).

Log on to the DataWorks console by using role-based SSO

1. Create a role to be assumed.

 **Note** To create a role by using a RAM user or a RAM role, you must attach the AliyunRAMFullAccess policy to the RAM user or RAM role.

We recommend that you create the following roles:

- RAM role that is used as the access identity of the MaxCompute compute engine
 - a. Log on to the RAM console by using your Alibaba Cloud account and password. On the page that appears, choose **Identities > Roles**. On the Roles page, click **Create Role**. In the **Select Role Type** step of the **Create Role** panel, configure the **Select Trusted Entity** parameter and click **Next**.
 - If you want a RAM user to assume the RAM role by switching the identity in the Alibaba Cloud Management Console, select **Alibaba Cloud Account** for **Select Trusted Entity**. For more information, see [Create a RAM role for a trusted Alibaba Cloud account](#).
 - If you want an on-premises identity provider (IdP) to assume the RAM role, select **IdP** for **Select Trusted Entity**. For more information, see [Create a RAM role for a trusted IdP](#).

- b. In the **Configure Role** step, configure the **RAM Role Name** parameter and other parameters based on your business requirements. Then, click **OK**.
- c. On the **Roles** page, click the name of the role that you created. On the page that appears, click **Edit Trust Policy** on the **Trust Policy Management** tab. In the **Edit Trust Policy** panel, replace the policy document with the following script:

```
{
  "Statement": [
    {
      "Action": "sts:AssumeRole",
      "Effect": "Allow",
      "Principal": {
        "Service": [
          "dataworks.aliyuncs.com"
        ]
      }
    }
  ],
  "Version": "1"
}
```

- o RAM role that is assumed by a RAM user or an on-premises IdP

- Create a RAM role that is assumed by a RAM user.

If you want a RAM user to assume the RAM role, go to the **Roles** page and click **Create Role**. In the Create Role panel, select **Alibaba Cloud Account** for Select Trusted Entity and click **Next**. In the Configure Role step, configure the parameters and click **OK**. After the RAM role is created, replace the document of the trust policy with the following script:

```
{
  "Statement": [
    {
      "Action": "sts:AssumeRole",
      "Effect": "Allow",
      "Principal": {
        "RAM": [
          "acs:ram::UID:root"
        ]
      }
    },
    {
      "Action": "sts:AssumeRole",
      "Effect": "Allow",
      "Principal": {
        "Service": [
          "dataworks.aliyuncs.com"
        ]
      }
    }
  ],
  "Version": "1"
}
```

 **Note** Replace UID in the preceding script with the actual UID of your Alibaba Cloud account.

- Create a RAM role that is assumed by an on-premises IdP.

If you want an on-premises IdP to assume the RAM role, go to the **Roles** page and click **Create Role**. In the Create Role panel, select **IdP** for Select Trusted Entity and click **Next**. In the Configure Role step, configure the parameters and click **OK**. After the RAM role is created, replace the document of the trust policy with the following script:

```
{
  "Statement": [
    {
      "Action": "sts:AssumeRole",
      "Condition": {
        "StringEquals": {
          "saml:recipient": "https://signin.aliyun.com/saml-role/sso"
        }
      },
      "Effect": "Allow",
      "Principal": {
        "Federated": [
          "acs:ram::UID:saml-provider/IDP"
        ]
      }
    },
    {
      "Action": "sts:AssumeRole",
      "Effect": "Allow",
      "Principal": {
        "Service": [
          "dataworks.aliyuncs.com"
        ]
      }
    }
  ],
  "Version": "1"
}
```

Note Replace UID in the preceding script with the UID of your Alibaba Cloud account and IDP with the name of IdP that you selected.

For example, you create a RAM role named user-role, which is assumed by a RAM user.

2. Create a DataWorks workspace.

Note To create a DataWorks workspace by using a RAM user or a RAM role, you must attach the AliyunDataWorksFullAccess policy to the RAM user or RAM role.

- Log on to the **DataWorks console**.
- In the left-side navigation pane, click **Workspaces**. In the top navigation bar, select the region in which you want to create a workspace. Then, click **Create Workspace**.
- In the **Create Workspace** panel, configure the parameters in the Basic Settings step and click **Next**.

Section	Parameter	Description
Basic Information	Workspace Name	The name of the workspace. The name must be 3 to 23 characters in length and can contain letters, underscores (_), and digits. The name must start with a letter.
	Display Name	The display name of the workspace. The display name can be a maximum of 23 characters in length. It can contain letters, underscores (_), and digits and must start with a letter.
	Mode	<p>The mode of the workspace. Valid values: Basic Mode (Production Environment Only) and Standard Mode (Development and Production Environments).</p> <ul style="list-style-type: none"> ▪ Basic Mode (Production Environment Only): A workspace in basic mode is associated with only one MaxCompute project. Workspaces in basic mode do not isolate the development environment from the production environment. In these workspaces, you can perform only basic data development and cannot strictly control the data development process and the permissions on tables. ▪ Standard Mode (Development and Production Environments): A workspace in standard mode is associated with two MaxCompute projects. One serves as the development environment, and the other serves as the production environment. Workspaces in standard mode allow you to develop code in a standard way and strictly control the permissions on tables. These workspaces impose limits on table operations in the production environment for data security. <p>For more information, see Basic mode and standard mode.</p>
	Description	The description of the workspace.

Section	Parameter	Description
Advanced Settings	Download SELECT Query Result	Specifies whether the query results that are returned by SELECT statements in DataStudio can be downloaded. If you turn off this switch, the query results cannot be downloaded. You can change the setting of this parameter for the workspace in the Workspace Settings panel after the workspace is created. For more information, see Configure security settings .

- iv. In the **Select Engines and Services** step, select **MaxCompute** in the **Compute Engines** section and click **Next**.
- v. In the **Engine Details** step, configure the parameters based on your business requirements. The following table describes the parameters.

Parameter	Description
Instance Display Name	The display name of the compute engine instance. The display name must start with a letter and can contain only letters, underscores (_), and digits.
Resource Group	The quotas of the computing resources and disk space of the compute engine instance.
MaxCompute Data Type Edition	The MaxCompute data type edition. This configuration takes effect within 5 minutes. For more information, see Data type editions . If you do not know which edition to select, we recommend that you contact the workspace administrator.
MaxCompute Project Name	The name of the MaxCompute project. By default, the MaxCompute project that serves as the production environment is named after the DataWorks workspace. The MaxCompute project that serves as the development environment is named in the format of DataWorks workspace name_dev.
Account for Accessing MaxCompute	The identity that you used to access the MaxCompute project. In the development environment, the value is fixed to Node Owner . If you want to access MaxCompute projects in the production environment, select Alibaba Cloud Account for this parameter.
RAM role selection	The RAM role that you created. Select DwRole from the RAM role selection drop-down list.

- vi. Click **Create Workspace**.

After the workspace is created, you can go to the **Workspace Management** page to view the identity that is used to run nodes in MaxCompute. The identity is the RAM role that you created.

 **Note** The RAM role that is displayed for MaxCompute Visitor Identity can be used to run only nodes that are in the production environment. You can run nodes only on exclusive resource groups for scheduling.

3. Add another RAM role as a member of the DataWorks workspace.

 **Note** To add a RAM role as a member of the DataWorks workspace by using a RAM user or a RAM role, you must specify the RAM user or RAM role as the administrator of the workspace.

- i. Find the workspace that you created and click **DataStudio** in the Actions column. On the **DataStudio** page, click the  icon in the upper-right corner. The **Workspace Management** page appears.
 - ii. In the left-side navigation pane, click **User Management**.
 - iii. On the **Manage Members** tab, click **Add Member** in the upper-right corner.
 - iv. In the **Add Member** dialog box, click **Refresh**. The created RAM roles are displayed.
 - v. In the Available Accounts section, select the account that you want to add, such as **user-role**, and click **>** to add the account to Added Accounts.
 - vi. Select the roles that you want to assign to the added member and click **Confirm**.
4. (Optional) Attach the **AliyunSTSAssumeRoleAccess** policy to the RAM user that assumes the RAM role.

If you want a RAM user instead of an IdP to assume the RAM role, attach the **AliyunSTSAssumeRoleAccess** policy to the RAM user first.

5. Use the RAM user or IdP to log on to the DataWorks console.

The following steps show how to use the RAM user that assumes the user-role RAM role to log on to the DataWorks console and use DataWorks.

- i. Log on to the DataWorks console.

The user assumes the RAM role user-role by using role-based SSO and logs on to the DataWorks Console. For more information, see [Overview](#).
- ii. Go to the **Workspaces** page in the DataWorks console, find the desired workspace, and then click **DataStudio** in the Actions column.
- iii. Create an ODPS SQL node. For more information, see [Create an ODPS SQL node](#). The owner of the node is the RAM role user-role.

2. Create a workspace

A workspace in DataWorks is similar to a project in MaxCompute. This topic describes how to create a workspace in the DataWorks console.

Prerequisites

An Alibaba Cloud account is available. For more information, see [Activate DataWorks](#). The Alibaba Cloud account is required to create a workspace, as described in this topic.

 **Note**

If a Resource Access Management (RAM) user is required to perform operations on behalf of an Alibaba Cloud account, the RAM user must be granted required permissions.

For more information about how to grant permissions to a RAM user, see [How do I use an Alibaba Cloud account to attach the AliyunDataWorksFullAccess policy to a RAM user?](#) and [User permission management](#).

Procedure

1. Log on to the [DataWorks console](#) by using your Alibaba Cloud account.
2. On the **Overview** page, click **create Workspace** in the Frequently Used Workspaces section on the right.
You can also click **Workspaces** in the left-side navigation pane and click **Create Workspace** on the page that appears.
3. In the **Create Workspace** panel, set the parameters in the Basic Settings step and click **Next**.

Section	Parameter	Description
	Workspace Name	The name of the workspace. The name must be 3 to 23 characters in length and can contain letters, underscores (_), and digits. The name must start with a letter.
	Display Name	The display name of the workspace. The display name can be a maximum of 23 characters in length. It can contain letters, underscores (_), and digits and must start with a letter.

Section	Parameter	Description
Basic Information	Mode	<p>The mode of the workspace. Valid values: Basic Mode (Production Environment Only) and Standard Mode (Development and Production Environments).</p> <ul style="list-style-type: none"> ◦ Basic Mode (Production Environment Only): A workspace in basic mode is associated with only one MaxCompute project. Workspaces in basic mode do not isolate the development environment from the production environment. In these workspaces, you can perform only basic data development and cannot strictly control the data development process and the permissions on tables. ◦ Standard Mode (Development and Production Environments): A workspace in standard mode is associated with two MaxCompute projects. One serves as the development environment, and the other serves as the production environment. Workspaces in standard mode allow you to develop code in a standard way and strictly control the permissions on tables. These workspaces impose limits on table operations in the production environment for data security. <p>For more information, see Basic mode and standard mode.</p>
	Description	The description of the workspace.
Advanced Settings	Download SELECT Query Result	<p>Specifies whether the query results that are returned by SELECT statements in DataStudio can be downloaded. If you turn off this switch, the query results cannot be downloaded. You can change the setting of this parameter for the workspace in the Workspace Settings panel after the workspace is created. For more information, see Configure security settings.</p>

4. In the **Select Engines and Services** step, select the compute engines and services based on your needs and click **Next**.

Note

- If you need to bind a compute engine, you must first activate the corresponding service, such as Realtime Compute, E-MapReduce, Hologres, Graph Compute, AnalyticDB for MySQL, or AnalyticDB for PostgreSQL. You cannot select the checkbox for a compute engine that is not activated.
- If you do not select a compute engine when you create a workspace, you cannot perform operations that are related to the compute engine on the **DataStudio** page. For example, you cannot create tables or cleanse data based on engine nodes.

DataWorks is available as a commercial service. If you have not activated DataWorks in a region, activate it before you create a workspace in the region.

Section	Parameter	Description
DataWorks Services <div style="border: 1px solid #ccc; background-color: #e0f2f1; padding: 5px; margin-top: 10px;"> <p>Note The services that are enabled for the workspace. By default, the check box in this section is selected.</p> </div>	Data Integration	Provides a stable, efficient, and scalable data synchronization platform. Data Integration is designed to efficiently transmit and synchronize data between heterogeneous data sources in complex network environments. For more information, see Data Integration .
	Data Analytics	Allows you to design a data computing process that consists of multiple mutually dependent nodes based on your business requirements. The nodes are run by the scheduling system of DataWorks. For more information, see DataStudio .
	Operation Center	Allows you to view all your nodes and node instances and perform operations on them. For more information, see Operation Center .
	Data Quality	Provides an end-to-end data quality solution that relies on DataWorks. This solution allows you to explore data, compare data, monitor data quality, scan SQL statements, and use intelligent alerting. For more information, see Data Quality .
	MaxCompute	Provides a rapid, fully managed data warehouse solution that can process terabytes or petabytes of data. MaxCompute supports fast computing on large amounts of data, effectively reduces costs for enterprises, and ensures data security. For more information, see the MaxCompute documentation . <div style="border: 1px solid #ccc; background-color: #e0f2f1; padding: 5px; margin-top: 10px;"> <p>Note After you create workspaces in DataWorks, you must associate them with MaxCompute projects. Otherwise, the error <code>project not found</code> is returned when you run commands in the workspaces.</p> </div>
	Realtime Compute	Allows you to develop streaming computing nodes in DataWorks.

Section	Parameter	Description
Compute Engines	E-MapReduce	<p>Allows you to use E-MapReduce (EMR) to develop big data processing nodes in DataWorks. For more information, see the EMR documentation.</p> <div style="border: 1px solid #add8e6; padding: 5px; margin-top: 10px;"> <p> Notice</p> <p>The operation of associating an EMR compute engine instance with a workspace is suspended in DataWorks. If you need to perform such an operation, submit a ticket.</p> </div>
	Hologres	Allows you to use HoloStudio in DataWorks to manage internal and foreign tables and develop Hologres SQL nodes.
	Graph Compute	Allows you to use Graph Studio in DataWorks to manage Graph Compute instances.
	AnalyticDB for PostgreSQL	<p>Allows you to develop AnalyticDB for PostgreSQL nodes in DataWorks. For more information, see Overview.</p> <div style="border: 1px solid #add8e6; padding: 5px; margin-top: 10px;"> <p> Note You can use the AnalyticDB for PostgreSQL compute engine only in DataWorks Standard Edition or a more advanced edition.</p> </div>
AnalyticDB for MySQL	<p>Allows you to develop AnalyticDB for MySQL nodes in DataWorks. For more information about AnalyticDB for MySQL, see Product introduction.</p> <div style="border: 1px solid #add8e6; padding: 5px; margin-top: 10px;"> <p> Note You can use the AnalyticDB for MySQL compute engine only in DataWorks Standard Edition or a more advanced edition.</p> </div>	
Machine Learning Services	PAI Studio	Uses statistical algorithms to learn large amounts of historical data and generate an empirical model to provide business strategies.

5. In the **Engine Details** step, set the parameters for the selected compute engines.

- o Associate a MaxCompute compute engine instance with the workspace

Parameter	Description
Method	Specifies whether to create a MaxCompute project or use an existing MaxCompute project. Valid values: Create Project and Associate Existing Project .

Parameter	Description
Instance Display Name	The display name of the MaxCompute compute engine instance. The display name must be 3 to 28 characters in length and can contain letters, underscores (_), and digits. The display name must start with a letter.
Region	The region of the workspace.
Payment mode	<p>The billing method of the MaxCompute compute engine instance. Valid values: The pay-as-you-go billing method, Monthly package, and Developer version.</p> <div style="background-color: #e6f2ff; padding: 5px;"> <p> Note A MaxCompute compute engine instance of the developer version cannot be associated with a workspace in standard mode.</p> </div>
Quota group	The quotas of computing resources and disk space for the MaxCompute compute engine instance.
MaxCompute data type	The data type edition of the MaxCompute compute engine instance. Valid values: 2.0 data type (recommended) , 1.0 data type (for users who already use 1.0 data type) , and Hive compatible types (for Hive migration users) . For more information, see Data type editions .
Whether to encrypt	Specifies whether to encrypt the MaxCompute compute engine instance.
Production Environment	<p>Configure the MaxCompute Project Name and Access Identity parameters for the production environment.</p> <ul style="list-style-type: none"> ▪ MaxCompute Project Name: the name of the MaxCompute project that you want to associate with the workspace as the compute engine instance in the production environment. ▪ Access Identity: the identity that is used to access the MaxCompute project. Valid values: Alibaba Cloud primary account, Alibaba Cloud sub-account, and Alibaba Cloud RAM role.
Development Environment	<p>Configure the MaxCompute Project Name and Access Identity parameters for the development environment.</p> <ul style="list-style-type: none"> ▪ MaxCompute Project Name: the name of the MaxCompute project that you want to associate with the workspace as the compute engine instance in the development environment. <div style="background-color: #e6f2ff; padding: 5px;"> <p> Note This MaxCompute project provides computing and storage resources.</p> </div> <ul style="list-style-type: none"> ▪ Access Identity: The default value is Task owner and cannot be changed.

- o Associate an E-MapReduce compute engine instance with the workspace

Parameter	Description
Instance Display Name	The display name of the EMR compute engine instance.
Region	The region of the workspace.
Access Mode	<ul style="list-style-type: none"> ■ In shortcut mode, if you run or schedule EMR nodes in DataWorks by using an Alibaba Cloud account or a RAM user, the code of the nodes is committed to the EMR compute engine instance and run by a Hadoop user in the EMR compute engine instance. ■ In security mode, if you run or schedule EMR nodes in DataWorks by using an Alibaba Cloud account or a RAM user, the code of the nodes is committed to the EMR compute engine instance and run by a user that has the same name as the Alibaba Cloud account or RAM user in the EMR compute engine instance. You can use EMR Ranger to manage the permissions of each user in the EMR compute engine instance. This ensures that Alibaba Cloud accounts, node owners, or RAM users have different data permissions when they run EMR nodes in DataWorks. This way, higher data security is implemented.
Scheduling access identity	<ul style="list-style-type: none"> ■ If you set the Access Mode parameter to Shortcut mode, you can commit the code of an EMR node to the EMR compute engine instance by using an Alibaba Cloud account or a RAM user after the node is committed and deployed to the scheduling system of DataWorks. ■ If you set the Access Mode parameter to Security mode, you can commit the code of an EMR node to the EMR compute engine instance by using an Alibaba Cloud account or a RAM user or as a node owner after the node is committed and deployed to the production environment. A Hadoop user that corresponds to the identity in the EMR compute engine instance is used to run the code. <div style="background-color: #e6f2ff; padding: 10px; margin-top: 10px;"> <p> Note</p> <ul style="list-style-type: none"> ■ This parameter is available only for the production environment. ■ Before you associate an EMR compute engine instance with a workspace, you must attach the AliyunEMRDevelopAccess policy to workspace roles such as developers and administrators. This way, the roles can be used to create and run EMR nodes in DataStudio. <ul style="list-style-type: none"> ■ The AliyunEMRDevelopAccess policy is attached to Alibaba Cloud accounts by default. ■ If you want to use a RAM user to run EMR nodes, you must attach the AliyunEMRDevelopAccess policy to the RAM user. For more information, see Grant permissions to RAM users. </div>

Parameter	Description
Access identity	<p>The identity that is used to commit the code of an EMR node in the development environment to the EMR compute engine instance. Default value: Task owner.</p> <div style="background-color: #e6f2ff; padding: 10px; border: 1px solid #d9e1f2;"> <p> Note</p> <ul style="list-style-type: none"> ■ This parameter is available only for the development environment of a workspace in standard mode. ■ Task owner can be an Alibaba Cloud account or a RAM user. <p>Before you associate an EMR compute engine instance with a workspace, you must attach the AliyunEMRDevelopAccess policy to workspace roles such as developers and administrators. This way, the roles can be used to create and run EMR nodes in DataStudio.</p> <ul style="list-style-type: none"> ■ The AliyunEMRDevelopAccess policy is attached to Alibaba Cloud accounts by default. ■ If you want to use a RAM user to run EMR nodes, you must attach the AliyunEMRDevelopAccess policy to the RAM user. </div>
Cluster ID	<p>The ID of the EMR cluster that you want to associate with the workspace as the compute engine instance. Select an ID from the drop-down list. The EMR cluster is used as the runtime environment of EMR nodes.</p>
Project ID	<p>The ID of the EMR project that you want to associate with the workspace. Select an ID from the drop-down list. The EMR project is used as the runtime environment of EMR nodes.</p> <div style="background-color: #e6f2ff; padding: 10px; border: 1px solid #d9e1f2;"> <p> Note If you set Access Mode to Security mode, no EMR project IDs are displayed and can be selected.</p> </div>
YARN resource queue	<p>The name of the YARN resource queue in the EMR cluster. Unless otherwise specified, set this parameter to <i>default</i>.</p>

Parameter	Description
Override DataStudio YARN resource queue	<ul style="list-style-type: none"> Override DataStudio YARN resource queue is selected: All EMR nodes are run based on the specified YARN resource queue. Override DataStudio YARN resource queue is not selected: <ul style="list-style-type: none"> If you configure the queue parameter for an EMR node on the Advanced Settings tab, the EMR node is run based on the configured YARN resource queue. If you do not configure the queue parameter for an EMR node or you delete the setting of the queue parameter for an EMR node on the Advanced Settings tab, the EMR node is run based on the specified YARN resource queue. <p>Note If the Override DataStudio YARN resource queue check box is not displayed, submit a ticket to upgrade the DataWorks edition.</p>
Endpoint	The endpoint of the EMR cluster. The value of this parameter cannot be changed.
Resource Group	<p>Select an exclusive resource group for scheduling that connects to the DataWorks workspace. If no exclusive resource group for scheduling is available, create one. For more information about how to create an exclusive resource group for scheduling and configure network connectivity, see Create and use an exclusive resource group for scheduling.</p> <p>After you select an exclusive resource group for scheduling, click Test Connectivity to test the connectivity between the exclusive resource group for scheduling and the EMR cluster. After the connectivity test is passed, the system initializes the exclusive resource group for scheduling.</p> <p>Note The exclusive resource group for scheduling must be reinitialized if the configuration of the EMR cluster is modified.</p>

- o Associate a Hologres compute engine instance with the workspace

Parameter	Description
Instance Display Name	The display name of the Hologres compute engine instance.

Parameter	Description
Access identity	<ul style="list-style-type: none"> The identity that is used to run the code of committed Hologres nodes. Valid values: Alibaba Cloud primary account and Alibaba Cloud sub-account. <div style="background-color: #e0f2f1; padding: 5px; margin: 5px 0;"> <p> Note This parameter is available only for the production environment.</p> </div> <ul style="list-style-type: none"> The default value of this parameter for the development environment is Task owner.
Hologres instance name	The name of the Hologres instance that you want to associate with the workspace as the compute engine instance.
Database name	The name of the database that is created in SQL Console , such as testdb.

- Associate a Graph Compute compute engine instance with the workspace
- Associate an AnalyticDB for PostgreSQL compute engine instance with the workspace

Parameter	Description
Instance Display Name	The display name of the AnalyticDB for PostgreSQL compute engine instance. The display name must be unique.
InstanceName	The name of the AnalyticDB for PostgreSQL instance that you want to associate with the workspace as the compute engine instance.
DatabaseName	The name of the AnalyticDB for PostgreSQL database that you want to associate with the workspace.
Username	The username that you can use to connect to the database. You can obtain the information from the Account Management page in the AnalyticDB for PostgreSQL console. For more information, see Create a database account .
Password	The password that you can use to connect to the database. You can obtain the information from the Account Management page in the AnalyticDB for PostgreSQL console. For more information, see Create a database account .
Connectivity Test	<p>AnalyticDB for PostgreSQL nodes must be run on exclusive resource groups for scheduling. Therefore, you must select an exclusive resource group for scheduling. For more information, see Exclusive resource group mode.</p> <p>Click Test Connectivity to test the connectivity between the specified exclusive resource group for scheduling and AnalyticDB for PostgreSQL instance. If no exclusive resource group for scheduling is available, click Create Exclusive Resource Group to create one.</p>

- o Associate an AnalyticDB for MySQL compute engine instance with the workspace

Parameter	Description
Instance Display Name	The display name of the AnalyticDB for MySQL compute engine instance. The display name must be unique.
InstanceName	The name of the AnalyticDB for MySQL cluster that you want to associate with the workspace as the compute engine instance.
DatabaseName	The name of the AnalyticDB for MySQL database that you want to associate with the workspace.
Username	The username that you can use to connect to the database. You can obtain the information from the Accounts page in the AnalyticDB for MySQL console. For more information, see Database accounts and permissions .
Password	The password that you can use to connect to the database. You can obtain the information from the Accounts page in the AnalyticDB for MySQL console. For more information, see Database accounts and permissions .
Connectivity Test	<p>AnalyticDB for MySQL nodes must be run on exclusive resource groups for scheduling. Therefore, you must select an exclusive resource group for scheduling. For more information, see Exclusive resource group mode.</p> <p>Click Test Connectivity to test the connectivity between the specified exclusive resource group for scheduling and AnalyticDB for MySQL cluster. If no exclusive resource group for scheduling is available, click Create Exclusive Resource Group to create one.</p>

6. Click **Create Workspace**.

After the workspace is created, you can view the information about the workspace on the **Workspaces** page.

Note

- o If you are the owner of a workspace, all data in the workspace belongs to you. Other users can access the workspace only after you grant permissions to them. If you create a workspace as a RAM user, the workspace belongs to the RAM user and the Alibaba Cloud account that manages the RAM user.
- o You can add a RAM user to a workspace so that the RAM user can use the workspace. This way, the RAM user does not need to create a workspace.

What's next

You have learned how to create a workspace. You can now proceed with the next tutorial to [add workspace members](#) or directly perform operations that are described in [Quick Start](#). Quick Start guides you through a complete process of data development and O&M.

3. Add workspace members and assign roles to them

This topic describes how to add workspace members and assign roles to them. This topic also describes the permissions of each role.

Prerequisites

If you need to use DataWorks together with other users, create Resource Access Management (RAM) users, create a workspace, and then follow the instructions in this topic to add members to your workspace. For more information about how to create RAM users and workspaces, see [Prepare a RAM user](#) and [Create a workspace](#).

Context

 **Note** If you plan to perform all operations in your workspace by using only the Alibaba Cloud account, you can skip this topic and directly create tables and import data. For more information, see [Create tables and import data](#).

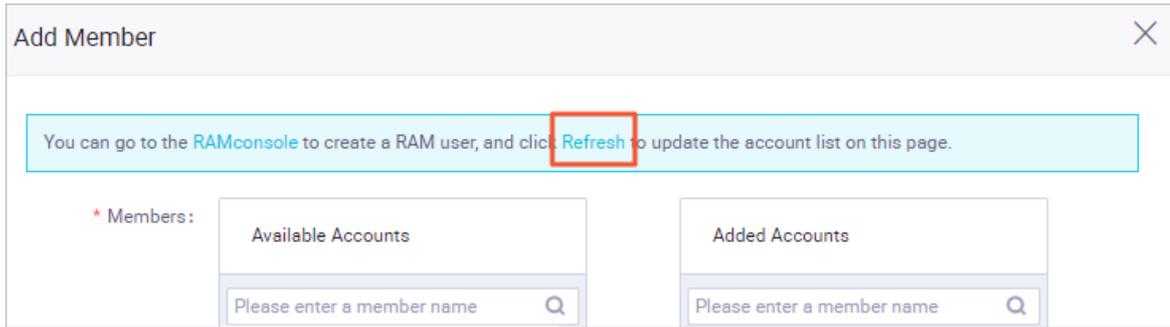
You can log on to the DataWorks console by using your Alibaba Cloud account or as a **RAM** user.

Alibaba Cloud adopts the following billing rules for RAM users:

- An Alibaba Cloud account owns Alibaba Cloud resources. Resource usage is measured and billed to the Alibaba Cloud account. You can use the Alibaba Cloud account to create RAM users for your enterprise. You can also use the Alibaba Cloud account to manage and grant permissions to RAM users.
- RAM users are created and managed by the Alibaba Cloud account in the RAM console. RAM users do not own resources, and therefore the resource usage of a RAM user is not measured or billed to the RAM user. The Alibaba Cloud account centrally manages all RAM users and pays for the resources used by RAM users.

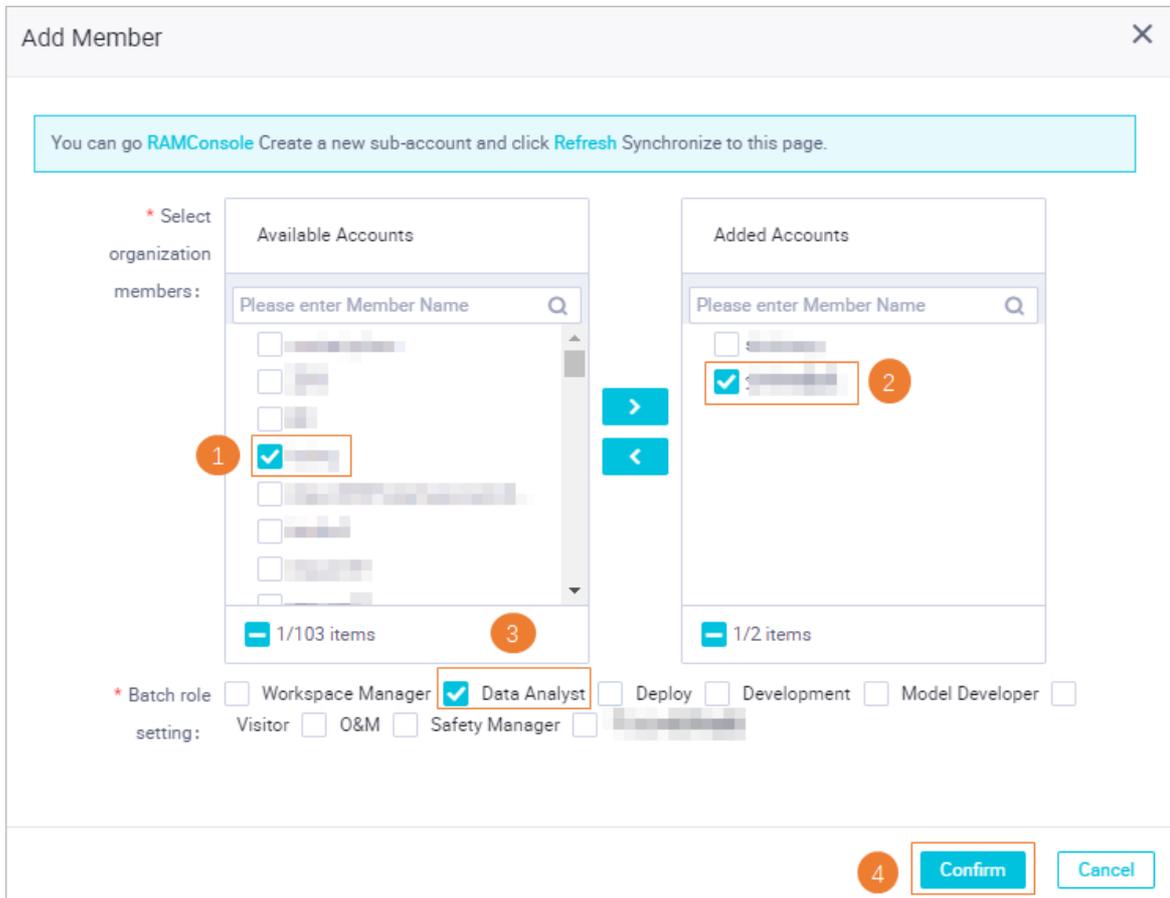
Procedure

1. Go to the **Workspace Management** page of the workspace to which you want to add members.
 - i. Log on to the [DataWorks console](#).
 - ii. In the left-side navigation pane, click **Workspaces**.
 - iii. On the Workspaces page, find the workspace to which you want to add members, move the pointer over the more icon in the Actions column, and then select **Workspace Settings**.
 - iv. In the **Workspace Settings** panel, click **More**. The **Workspace Management** page appears.
You can also click DataStudio in the Actions column of the workspace to which you want to add members. On the **DataStudio** page, click the  icon in the upper-right corner to go to the **Workspace Management** page.
2. In the left-side navigation pane, click **User Management**.
3. On the **User Management** page, click **Add Member** in the upper-right corner.
4. In the **Add Member** dialog box, click **Refresh**. All the RAM users created by your Alibaba Cloud account appear in the Available Accounts section.



Note If you need to create more RAM users, click **RAMConsole** in the Add Member dialog box to go to the RAM console and create RAM users as required. For more information about how to create a RAM user and allocate the RAM user to a person, see **Prepare a RAM user**.

- 5. Select RAM users in the **Available Accounts** section and click the > icon to move them to the **Added Accounts** section.
- 6. Select the roles that you want to assign to the RAM users and click **Confirm**.



Notice You must move RAM users from the Available Accounts section to the Added Accounts section before you assign roles to them.

- 7. Go to the User Management page and view or modify the roles of each added member. You can click Remove in the Actions column of a member to remove the member.

You can assign the following roles to a workspace member: **Project Owner, Workspace**

Manager, Data Analyst, Development, O&M, Deploy, Visitor, and Safety Manager. By default, the creator of a workspace is assigned the Workspace Manager role.

Role	Description
Project Owner	This role has full permissions on a workspace.
Workspace Manager	This role has all permissions of the Development and O&M roles. This role can also perform the following operations. For example, add a user to a workspace as a member, remove a member from a workspace, or create a custom resource group.
Data Analyst	By default, this role has permissions only on DataAnalysis.
Development	This role has permissions to perform design and maintenance operations on the DataStudio page of a workspace.
O&M	This role has permissions to manage the execution of and perform the required operations on all nodes in a workspace in Operation Center.
Deploy	This role has permissions to review the code of a node and determine whether to commit the node to Operation Center in a workspace in standard mode.
Visitor	This role has read-only permissions on workflows and code on the DataStudio page of a workspace.
Safety Manager	This role has permissions only on Data Security Guard. For more information about Data Security Guard, see Overview .

The DataWorks built-in roles and the roles in a MaxCompute project have a permission mapping. By default, a role in a MaxCompute project in the development environment has specific permissions on the project. However, the role does not have permissions on a MaxCompute project in the production environment. If the role wants to perform operations on data in a project in the production environment, the role can apply for the required permissions in [Security Center](#).

MaxCompute role	Permission on data in a MaxCompute project in the development environment	DataWorks built-in role	Permission on a DataWorks workspace
Project Owner	This role has all permissions on the project.	N/A	N/A
Super_Administrator	This role has management permissions on the project and all permissions on all types of resources in the project.	N/A	N/A

MaxCompute role	Permission on data in a MaxCompute project in the development environment	DataWorks built-in role	Permission on a DataWorks workspace
Admin	<p>When you create a project, the system automatically creates an Admin role for this project and grants the following permissions to the role: The Admin role has the permissions to access all objects in the project, manage users or roles, and grant permissions to users or roles.</p> <p>Unlike the Project Owner role, the Admin role does not have permissions to perform the following operations: assign the Admin role to users, configure security policies for the project, modify the authentication model for the project, and modify the permissions of the Admin role.</p> <p>The Project Owner role can assign the Admin role to a user and authorize the user to manage security configurations.</p>	N/A	N/A
Role_Project_Data_Analyst	<p>By default, this role does not have permissions to perform operations on data in a MaxCompute project in the development or production environment. If the role wants to perform operations on data in a project in the development or production environment, the role can apply for the required permissions in Security Center.</p>	Data Analyst	This role has permissions only on DataAnalysis by default.

MaxCompute role	Permission on data in a MaxCompute project in the development environment	DataWorks built-in role	Permission on a DataWorks workspace
Role_Project_Admin	This role has all permissions on tables, functions, resources, instances, jobs, and packages of a project.	Workspace Manager	The administrator of the workspace. This role has permissions to manage the basic properties, data sources, compute engine configurations, and members of the workspace and assign the Workspace Manager, Development, O&M, Deploy, or Visitor role to workspace members.
Role_Project_Dev	This role has all permissions on tables, functions, resources, instances, jobs, and packages of a project.	Development	This role has permissions to create workflows, script files, resources, user-defined functions (UDFs), tables, and deployment tasks, and delete tables, but no permissions to perform deployment operations.
Role_Project_Pe	This role has all permissions on functions, resources, instances, and jobs of a project. It also has read permissions on packages and both read and describe permissions on tables of the project.	O&M	This role has deployment and online O&M permissions that are granted by the Workspace Manager role but no permissions to develop data.
Role_Project_Deploy	No permissions by default.	Deploy	This role has the same permissions as the O&M role, except for online O&M permissions.
Role_Project_Guest	No permissions by default.	Visitor	This role has permissions to view data but no permissions to edit workflows or code.
Role_Project_Security	No permissions by default.	Safety Manager	This role has permissions to configure sensitivity rules and audit data risks in Data Security Guard.

For more information about, see [Users, roles, and permissions](#).

4. Work as a workspace administrator

This topic describes the responsibilities and permissions of a workspace administrator. By default, the Alibaba Cloud account that creates a workspace is the owner and administrator of the workspace and has full permissions on the workspace.

The owner can also specify a Resource Access Management (RAM) user as a workspace administrator.

Create a workspace

After you create a workspace by using your Alibaba Cloud account or as a RAM user that is managed by your account, the workspace belongs to the Alibaba Cloud account. For more information about how to create a workspace, see [Create a workspace](#).

Note

- A RAM user can create a workspace only if it has the AliyunDataWorksFullAccess permission. For more information, see [Grant a RAM user the permissions to perform operations in the DataWorks console](#).
- A workspace administrator needs to maintain stable execution of the workspace in the production environment, grant least permissions to workspace members, and control operation permissions on tables in the workspace.

Add workspace members

A workspace administrator can add RAM users as members of the workspace and assign roles to the members as required. For more information about the permissions of each role, see [Permissions of built-in workspace-level roles](#).

 **Note** We recommend that you do not assign the Development and O&M roles to the same member.

Manage permissions

DataWorks roles are divided into preset roles and workspace-level custom roles. Each role has different permissions. You can assign a role to a user when you add the user to a workspace. This way, the user obtains the permissions that are configured for the role. For more information, see [Manage workspace-level roles and members](#). Each role has different operation permissions on the DataWorks console. For more information, see [Permissions of built-in workspace-level roles](#).

If you use a MaxCompute engine, mappings are established between the preset roles of DataWorks and the MaxCompute roles for projects in Dev mode. In other words, after a preset role of DataWorks is assigned to a user, the user can manage the resources of the corresponding project in Dev mode in MaxCompute. For more information, see [Users, roles, and permissions](#).

To ensure stability and security of the production environment, DataWorks does not allow RAM users to perform operations on tables in the production environment. For example, RAM users cannot modify or delete tables in the production environment. In addition, workspace members must be granted related permissions before they can commit nodes.

Other permissions:

- **Data Integration:** Only the owner and administrator of a workspace can perform operations such as adding data sources and migrating tables to the cloud in the workspace.
- **MaxCompute Management:** A workspace administrator can bind a resource group to a workspace. Then, O&M engineers can view the system status, allocate resource groups, and monitor nodes in MaxCompute Management.
- **Operation Center:** Only a member that is assigned the O&M role or an administrator of a workspace can perform advanced operations in Operation Center.
- **DataStudio:** Only a developer or an administrator of a workspace can perform development operations in DataStudio.

When you create a workspace, you must first specify whether to use your **Alibaba Cloud account** or the credentials of a **RAM user** to run nodes in the workspace. An invalid setting will damage the permission system of DataWorks.

- **Alibaba Cloud account:** The AccessKey ID and AccessKey secret of your Alibaba Cloud account are required to execute SQL statements. The SQL statements can be executed on tables in all workspaces in the specified region. Exercise caution when you select this option.
- **RAM user:** The AccessKey ID and AccessKey secret of a RAM user are required to execute SQL statements. The permissions of RAM users are strictly controlled. Only authorized RAM users can perform operations on tables in the production environment.

 **Note**

To ensure data security, we recommend that you assign roles with least permissions to RAM users.