

Alibaba Cloud Data Online Migration

Migrate data from HTTP/HTTPS sources to
OSS

Issue: 20200601









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 due to product version upgrades, adjustments, or other reasons. Alibaba Cloud reserves the right to modify the content of this document without notice and the updated versions of this document will be occasionally released through Alibaba Cloud-authorized channels. You shall 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 the document in the context that Alibaba Cloud products and services are provided on an "as is", "with all faults" and "as available" basis. 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 bear any liability for any errors or financial losses incurred by any organizations, companies, or individuals arising from their download, use, or trust in this document. Alibaba Cloud shall not, under any circumstances, bear responsibility for any indirect, consequential, exemplary, incidental, 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 contact Alibaba Cloud directly if you discover any errors in this document.

Document conventions

Style	Description	Example
	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.
	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.
	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.
	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.
Italic	Italic formatting is used for parameters and variables.	<code>bae log list --instanceid Instance_ID</code>
[] or [a b]	This format is used for an optional value, where only one item can be selected.	<code>ipconfig [-all -t]</code>

Style	Description	Example
{ } or {a b}	This format is used for a required value, where only one item can be selected.	switch {active stand}

Contents

Legal disclaimer.....	1
Document conventions.....	1
1 Background information.....	1
2 Preparations.....	2
4 Manage migration jobs.....	5

1 Background information

This topic describes how to migrate network resources to Object Storage Service (OSS).

Alibaba Cloud Data Transport is used as a data channel between various data stores. With Data Transport, you can migrate data from third-party data stores to OSS or between OSS buckets.

With Data Transport, you only need to log on to the console, specify a source data address and a destination OSS endpoint, and then create a migration job. After starting a migration job, you can perform management tasks for the job such as viewing the progress and status of the job. Additionally, you can generate the migration report to view the list of migrated files and the list of files that failed to migrate.

This guide includes the following topics:

- [Prerequisites](#)
- [Create a migration job](#)
- [Manage migration jobs](#)

2 Preparations

This topic describes what you need to prepare before migration.

Network resources

- Estimate the amount of data to be migrated.

Estimate the size and number of files to be migrated.

**Note:**

To ensure a successful migration, you must enter an appropriate size and number of files when [creating a migration job](#).

- Load network resources

1. Create an on-premises list file.

The list file has two columns.

- The first column includes a list of HTTP/HTTPS URLs. The migration service uses the HTTP GET method to download a file from an HTTP/HTTPS URL and the HTTP HEAD method to obtain metadata of a file.
- The second column includes a list of file names. After a file is migrated, the name of the corresponding object consists of a prefix and the file name. Separate two columns with a tab (`\t`).

**Notice:**

You must specify a file rather than a file folder in a list file.

Each line includes a file name. Separate two lines with a line feed (`\n`).

For example, the content of a list file named list.txt is as follows:

```
http://docs-aliyun.cn-hangzhou.oss.aliyun-inc.com/docs/my.doc docs/my.doc
http://docs-aliyun.cn-hangzhou.oss.aliyun-inc.com/pics/my.jpg pics/my.jpg
```

**Notice:**

If a file name includes special characters or Chinese characters, you must perform URL encoding.

- If a file name includes special characters, such as spaces and tabs (`\t`), you must encode the URL and file name. For example, a file is named `#¥.jpg`. After the file

name is encoded, it is displayed as `﻿#%EF%BF%A5.jpg`. You need to specify the URL and file name in a list file as follows:

```
http://docs-aliyun.cn-hangzhou.oss.aliyun-inc.com/pics/﻿#%EF%BF%A5.jpg  
pics/﻿#%EF%BF%A5.jpg
```

- If a file name includes Chinese characters, you must encode the URL and keep the file name. For example, a file is named `<a Chinese file name>.jpg`. After the file name is encoded, it is displayed as `%e5%9b%be%e7%89%87.jpg`. You need to specify the URL and file name in a list file as follows:

```
http://docs-aliyun.cn-hangzhou.oss.aliyun-inc.com/pics/%e5%9b%be%e7%89%87.jpg  
pics/<a Chinese file name>.jpg
```

2. Upload the list file to OSS.

The path of the list file is `oss://{bucket}/{the name of the list file}`.

Alibaba Cloud Object Storage Service (OSS)

- Create an OSS bucket

Create a destination OSS bucket to store the migrated data. For more information, see [Create a bucket](#).

- Create and authorize a RAM user

1. Log on to the [RAM console](#).

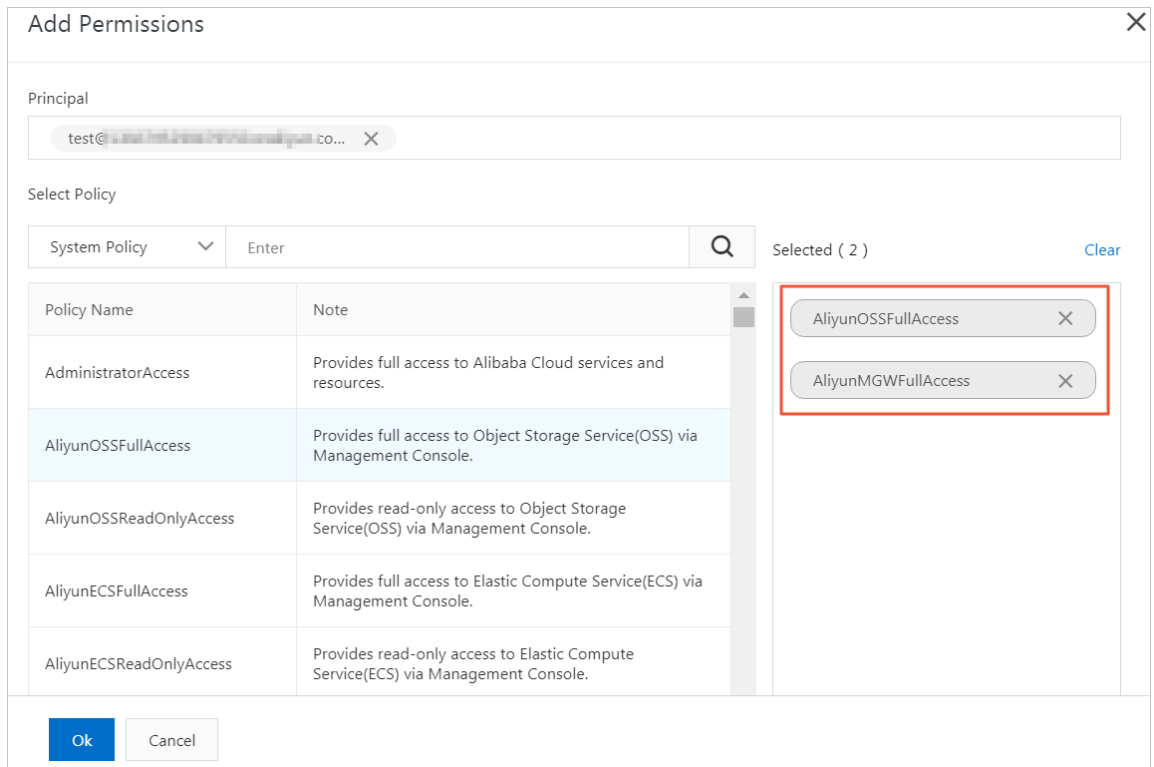
2. Choose **Identities > Users > Create User**.

3. Select **Console Password Logon** and **Programmatic Access**, and then enter the required RAM user information.

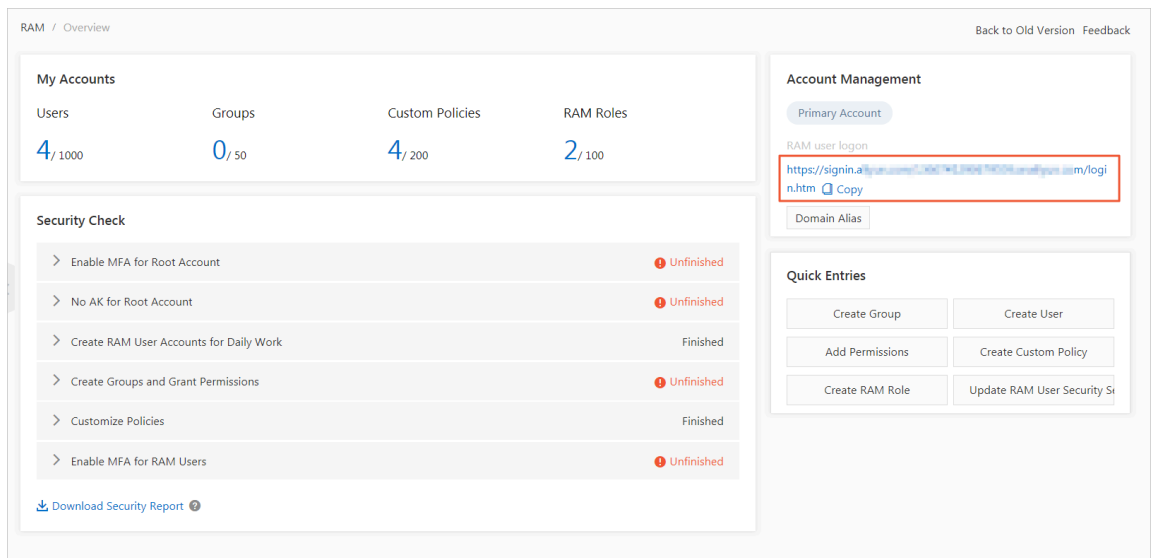
4. Click OK to save the generated RAM username, password, AccessKey ID, and AccessKey Secret.

5. Find the RAM user, and click **Add Permissions** to grant the OSS read/write permission (AliyunOSSFullAccess) and Data Transport management permission

(AliyunMGWFullAccess) to the RAM user. The following figure shows the authorization page.



6. In the left-side navigation pane, select **Overview**, click the link in the **RAM user logon** section, and enter the username and password of the new RAM user to log on to the console.



4 Manage migration jobs

This topic describes several subsequent operations that you can perform after you create a migration job.


View the status of a migration job

After you create a migration job, one of the following states is displayed for the job:

- **Migrating.** This state indicates that data is being migrated. Wait until the migration job is completed.
- **Create Failed.** This state indicates that the migration job failed to be created. You can view the cause of the failure and create a new migration job.
- **Completed.** This state indicates that the migration job is completed. You can view the migration report.
- **Failed.** This state indicates that the migration job failed. You can view the migration report and migrate the files that failed to be migrated.

Modify flow control settings

During migration, you can modify flow control settings at any time based on your needs.

1. Log on to the [Data Transport console](#).
2. Choose **Data Online Migration > Migration Jobs**. On the Migration Jobs page, find a migration job and click **Manage**.
3. Click **Stop** and ensure that the job is stopped.
4. On the Flow Control Time Schedule chart, click **Reset**.
 - To add a flow control setting, select an appropriate time range and maximum flow, and click **Add**.
 - To delete a flow control setting, click  next to the flow control setting.
 - To modify a flow control setting, you must delete the previous setting and add a new flow control setting.
5. Click **OK** and click **Start** to restart the job.

View a migration report

1. On the Migration Jobs page, find the required job and click **Manage**.

2. Click **Generate Migration Report**. After a report is generated, click **Export** to export the report.

In a migration report, the following files appear in the **File lists** section:

- The file whose name ends with `_total_list`. This file contains a list of files that need to be migrated.
 - The file whose name ends with `_completed_list`. This file contains a list of files that are successfully migrated.
 - The file whose name ends with `_error_list`. This file contains a list of files that failed to be migrated.
3. In the [OSS console](#), find the generated folder `aliyun_mgw_import_report/`. The folder includes the three files that appear in the migration report. You can download the files and view the detailed list included in each file. We recommend that you use the [ossbrowser](#) tool to view these files.

The file formats are described as follows:

- For the file whose name ends with `_total_list`, the four columns of the file list include the source data address, file name, file size (measured in bytes), and time when a file is last modified. The format of a source data address: `<vendor>://<bucketName>/<prefix>/<objectName>`. Example: `oss://bucket-test1022/myprefix/testfile.txt`.
- For the file whose name ends with `_completed_list`, the four columns of the file list include the file name, file size (measured in bytes), CRC-64 checksum, and time when a migration job is completed.
- For the file whose name ends with `_error_list`, the four columns of the file list include the file name, time when a migration job starts, time when the migration job ends, and error description.

Retry a failed migration job

If a migration job failed, you can view the generated file whose name ends with `_error_list` to identify the cause of the failure and troubleshoot the issue. On the Migration Jobs page, find the failed job, click **Manage**, and click **Retry** to remigrate the files that failed to be migrated.