

Object Storage Service

FAQ

FAQ

Data migration/synchronization

How to synchronize data to OSS

If you want the incremental data to be automatically synchronized to OSS buckets, you can use the following methods.

Methods

- OSS cross-region replication: Synchronize the data in bucket A to bucket B in another region. You can **set up cross-region replication** for the source bucket in the console.
- Mirroring: Seamlessly migrate data from the source site to OSS, that is, migrate a service that is already running on the source site or other cloud products to OSS without stopping it. If the data to be accessed cannot be found in OSS after historical data of the service has been migrated to OSS by OSSimport, mirroring is triggered to capture the data from the source site to OSS.



- OSS APIs or SDKs: Migrate data by coding using OSS APIs or SDKs. See [Copy Object](#) and [Upload Part Copy](#) for more information.

Scenario 1: Data synchronization between different OSS buckets

To synchronize the data in bucket A to bucket B, use the following methods:

- Set up cross-region replication
 - Apply to the scenario where the source bucket A and the target bucket B are in different regions and neither of them is of the Archive storage type.
- Mirroring
 - Apply to the scenario where the required data cannot be found in bucket B by the client. In this case, mirroring is triggered to capture and download the data from bucket A to bucket B.
- OSS APIs/SDKs Copy Object and Upload Part Copy
 - Apply to users who have special requirements and strong development capabilities.

FAQ

Can I configure to synchronize a subset of directories of a bucket rather than the entire bucket?

Yes. You can specify a file name prefix in the configuration by setting up the cross-region replication feature to synchronize only directories with that prefix.

Can I use cross-region replication for archive storage?

Cross-region replication is not supported if the source or target bucket is of the Archive storage type. You can synchronize the data by coding using OSS APIs/SDKs.

It takes time to restore data from an archive storage. If you have to synchronize data from an archive storage to a standard or infrequent access storage, we recommend that you use a more reasonable architecture by changing the storage type of the source bucket to standard or infrequent access and that of the target bucket to archive, reducing costs and improving business experience.

Scenario 2: Data synchronization from a third-party data source to OSS

To automatically synchronize data from a third-party data source to OSS, use the following methods:

- Mirroring
 - If the client cannot find the required data in OSS, mirroring is triggered to capture and download the data from the third-party source site to OSS.



- OSS APIs/SDKs Copy Object and Upload Part Copy

- Apply to users who have special requirements and strong development capabilities.
- OSSImport incremental migration
- Use the OSSIMPORT incremental data migration mode to query data in the source site at a fixed interval and synchronize the data to OSS.

FAQ

To migrate or replicate historical data to OSS in batch, see [data migration document](#).

Others

If the problem persists, contact [Alibaba Cloud After-Sales Technical Support](#).

How to migrate data to OSS

The user wants to migrate historical data to one of their target bucket in OSS. The source data to be migrated may be stored in an OSS bucket, a local storage, a third-party cloud storage (for example, AWS S3, Azure Blob, Tencent Cloud COS, Qiniu, Baidu BOS, UPYUN, and Kingsoft S3), or HTTP.

Methods

OssImport

Use the OssImport tool to migrate the data. This tool can migrate historical data of various data sources to OSS in batches. See [OssImport documentation](#).

Image origin retrieval

This method can seamlessly migrate data from an origin site to OSS. It can be used when services are running on the origin site established by the user or another cloud product, and data needs to be migrated to OSS without stopping the services. For more information about how to migrate new data after migrating historical data using the OssImport tool, see [Image origin retrieval](#) and [Seamless data migration to OSS](#).

Cross-region replication

This method can replicate data from Bucket A to Bucket B of the same user account. For more information, see [Set cross-region replication](#). Set the historical data synchronization function when using this method. In addition, if you do not want to synchronize the deletion operations in Bucket A to Bucket B, set the synchronization policy to "write

synchronization” .

Lightning cube

This method can migrate a large amount (TB or PB level) of offline data in a data center to a cloud. See [Lightning Cube](#) for more information.

OSS API or SDK

Users with high development capabilities can use OSS APIs or SDK code to migrate data. For more information, see [Copy Object](#) and [Upload Part Copy](#).

Scenario 1: migration of non-OSS data to OSS

In this scenario, the source data is to be migrated to an OSS bucket, but this data is stored in a non-OSS storage, such as a local storage, HTTP, or a third-party cloud storage (for example, AWS S3 and Azure Blob).

The methods are used as follows:

OssImport

Applicable to batch historical data migration to OSS, especially when the data volume is large. The distributed version of OssImport can be used for large-scale data migration.

Image original retrieval

Applicable to in-service hot data migration from an origin site to OSS. This method can meet the requirements for data migration of data and redirection of specific requests.

Set up cross-region replication

Not applicable.

Lightning Cube

- Any demand to migrate TBs or PBs of data from local data centers to the cloud
- Any demand for a professional solution that migrates data from local data centers to the cloud in a faster, safer, more efficient, and more cost-effective manner

OSS APIs/SDKs Copy Object and Upload Part Copy

- Source data is stored locally: Use OSS APIs/SDKs for code development.
- If the source data is stored in HTTP or a third-party cloud storage, you must write code to download the data from the data source, and then write code, by using OSS APIs/SDKs, to upload the data to OSS.

Scenario 2: data migration between OSSs

In this scenario, the data is to be migrated from a source OSS bucket to a target OSS bucket. The migration occurs only between two OSS buckets.

The methods are used as follows:

OssImport

Applicable for any situation, including data migration between different directories of a bucket or between different buckets (including buckets of the same account/region or of different accounts/regions). OssImport is particularly suitable for large-scale data migration, for which the distributed version is recommended.

OSS Cross-region Replication

- The source and target buckets belong to the same user but are in different regions.
- The source and target buckets are of one of the following storage types:
 - Standard for both the source and target buckets.
 - Standard for the source bucket and Infrequent Access for the target bucket.
 - Infrequent Access for the source bucket and Standard for the target bucket.
 - Infrequent Access for both the source and target buckets.

Note: If you do not want to synchronize any changes (additions, updates, and deletions) to the source bucket once the historical data of the source bucket is migrated (replicated) to the target bucket, you can disable cross-region replication for the source bucket in the console.

Lightning Cube

Not applicable.

OSS APIs/SDKs Copy Object and Upload Part Copy

- Use of OssImport or cross-region replication is preferred, unless you have more refined or personalized requirements and a certain degree of capability in development.
- Other relatively applicable situations:
 - The source and target buckets are in the same region.
 - Either the source or target bucket is of the Archive storage type.

Note: If the source and target buckets do not belong to the same user, OssImport takes priority over OSS APIs/SDKs, and cross-region replication is not applicable.

Others

If the problem persists, contact Alibaba Cloud After-Sales Technical Support.

Data protection

How does OSS verify the integrity of uploaded files?

ETags are created for objects in OSS. ETag identifies whether server data changes. However, the label is not equivalent to the MD5 value of an object file and it is not recommended using the label as a basis for verifying data consistency.

To check whether the file uploaded to OSS is consistent with the local file, attach the Content-MD5 value to the file while uploading and then the OSS server helps users with the MD5 check. Only when the MD5 of the file received by OSS server is equal to the Content-MD5 value can the uploading process be completed. This can guarantee the consistency of uploaded data.

If the problem persists, contact the after-sales technical support.

Domain/network

How do ECS users use OSS intranet addresses?

If the ECS instance you purchased is in the same region as the OSS bucket, your ECS can access OSS over the intranet without incurring any traffic cost. Whether the access occurs over the intranet depends on the domain name you configured. If the domain name you configured is the domain name of the intranet, the access occurs over the intranet. If the domain name you configured is not the domain name of the intranet, the access occurs over the Internet.

Note:

- ECS in the same region can access OSS over the intranet.
- ECS and OSS of different accounts can communicate with each other over the intranet.
- ECS in a different region cannot access OSS over the intranet.

Intranet address format

- VPC/Classic network ECS: BucketName.region-internal.aliyuncs.com

Example:

- East China 2 VPC/classic network ECS: aliyun-abc.oss-cn-shanghai-internal.aliyuncs.com

See the table in [Regions and endpoints](#) for available regions.

Quickly obtain the intranet address of the bucket in the console

Log on to the OSS console, and select the bucket to open the **Overview** page of the bucket. The bucket address is displayed in the **Access Domain Name** area.

Rules of using OSS access domain names

For more information about the spelling rules and the Internet and intranet accessing rules for OSS, see [Endpoints](#).

How a website uses OSS resources over the intranet

If a website is deployed on ECS, directly embed the intranet URL of the object, for example, `http://ali.oss-cn-hangzhou-internal.aliyuncs.com/1.jpg`, for the customers to access. Customers who directly access the OSS intranet address by the webpage on their local PC, or an Internet environment, cannot obtain the object.

If a website on ECS in the same region calls OSS resources over the intranet, you must set the access link of the webpage as ECS resources. ECS obtains OSS resources over the intranet for the customers to access. How to obtain: API/SDK or reverse proxy. ECS bandwidth may become a bottleneck when the website is visited heavily.

How an SDK/API uses OSS over the intranet

When you configure an SDK to initialize the client, set the OSS intranet endpoint as the client endpoint. See [Endpoints and Data Centers](#).

Java

```
String endpoint = "http://oss-cn-hangzhou-internal.aliyuncs.com"; //for example, East China 1
String accessKeyId = "<yourAccessKeyId>";
String accessKeySecret = "<yourAccessKeySecret>";
OSSClient client = new OSSClient(endpoint, accessKeyId, accessKeySecret);
```

For more information, see [Java SDK](#).

PHP

```
$accessKeyId = "<the AccessKeyID you obtained from OSS>";
$accessKeySecret = "<the AccessKeySecret you obtained from OSS>";
$endpoint = "<The endpoint of the OSS data center you selected, for example, http://oss-cn-hangzhou-internal.aliyuncs.com>";
```

For more information, see [PHP SDK](#).

Python

```
auth = oss2.Auth('your AccessKeyID', 'your AccessKeySecret')
endpoint = 'http://oss-cn-hangzhou-internal.aliyuncs.com' # Suppose that your bucket is in the Hangzhou region
bucket = oss2.Bucket(auth, endpoint, 'your bucket name')
```

For more information, see [Python SDK](#).

.Net

```
const string accessKeyId = "<your AccessKeyId>";
const string accessKeySecret = "<your AccessKeySecret>";
const string endpoint = "http://oss-cn-hangzhou-internal.aliyuncs.com";
var ossClient = new OssClient(endpoint, accessKeyId, accessKeySecret);
```

For more information, see .NET SDK.

C

```
ptions->config = oss_config_create(options->pool);
aos_str_set(&options->config->endpoint, "http://oss-cn-hangzhou-internal.aliyuncs.com");
aos_str_set(&options->config->access_key_id, "<your AccessKeyId>");
aos_str_set(&options->config->access_key_secret, "<your AccessKeySecret>");
options->config->is_cname = 0;
options->ctl = aos_http_controller_create(options->pool, 0);
```

For more information, see C SDK.

Configure HTTPS for your own domain name in OSS through reverse proxy

Domain name: oss.a.com

Bucket: bu

Requirement: Access objects under bu through https://oss.a.com

Configure reverse proxy for **oss.a.com** on a server.

For reverse proxy configuration, see the following (The nginx configuration only serves as a reference. Perform the configuration based on your requirements).

```
server {

listen 443 ; #default_server;

server_name oss.a.com;

location / {

root /alidata/www/ www-a-com/;

//If the host and the bucket are not in the same region, use Internet.

proxy_pass http://aialiyun.oss-cn-hangzhou-internal.aliyuncs.com/;
```

```
proxy_set_header Referer http://www.test.com;  
  
}  
  
}
```

Resolve the records for oss.a.com to the IP address of the server.

Test whether the objects can be accessed successfully. If so, you can add the SSL certificate for this domain name on the server.

If the problem persists, contact after-sales technical support.

Low OSS upload speed

In a scenario where a single client uploads data through Internet, the upload speed is mainly limited by the local bandwidth of the client and the quality of the network provided by operators.

If you want to upload buckets in the same region through ECS, we recommend you upload them with intranet IP addresses, which not only takes advantage of the stability of intranet but also reduces the load on the network card of ECS for Internet. For more information, see [How to use OSS intranet IP addresses](#).

Bucket management

How to move and modify the object and directory in OSS?

Actually, OSS has no concept of directory structure. The OSS directory structure is simulated by using a forward slash (/).

If an object whose name ends with `"/`, it is recognized as a directory.

You cannot rename an object. However, you can move the object to rename it. To move an object, you must manually upload the object to a new directory and delete it from the original directory. Before deleting a directory, make sure that all objects in the directory are deleted.

If the problem persists, contact Alibaba Cloud after-sales technical support.

Attaching a Bucket to OSSFS Through the www Account

If you are the root user, you can use the `"umount /attached directory"` command to detach a directory.

If you are a non-root user, you can use the following command to detach a directory:

```
fusermount -u your_mountpoint
```

Prerequisites: directory to attach: test1; account name: www (no logon permission); bucket to attach: dcces

"Which ossfs" is used to obtain the OSSFS installation path, because a regular account does not have the required permission.

1. Log on using the root account.
2. Run the `"id www"` command to obtain the UID and GID.
3. Create the directory to which you want to attach a bucket through the root account or other account, and run the following command to change the directory owner: `chown -R www:root test1`
4. Run the following command to modify the permission on the configuration file: `chmod 600 passwd-ossfs`

Run the following command to change the owner of the configuration file and the group where it belongs:

```
chown -R www:root passwd-ossfs
```

5. `sudo -u www /usr/local/bin/ossfs dcces test1 -ourl=https://oss-cn-hangzhou.aliyuncs.com -oallow_other -omp_umask=007 -ouid=504 -ogid=0 -o passwd_file=passwd-ossfs -odefault_permission=750`

6. Run the following command to attach the bucket to the directory:

```
sudo -u www ls -ll test1
```

For more information, contact Alibaba Cloud After-Sales Technical Support.

How to find access logs to the OSS object

To view the access records to the object in OSS, you must enable Logging in the OSS. In the bucket attribute in OSS, you can enable logging and set the bucket for storing the logs. Steps for the settings can be found.

OSS logging can record the period of time when the object was accessed. Users can find information about operations on the object within the period of time. In each operation log, the "Remote IP" field records the IP address of the request initiator, and the "Operation" field records the type of access request. The two fields and the time can be used to identify the IP address of the request initiator and the request type used to access the object file.

For information about the log file format, see [Set access logging](#).

Object management

How to rename an object?

You have two ways to rename an object:

Delete the old files and upload the renamed files.

Copy the source file of the object and rename the copied file. Then, delete the source file.

How to use OSS gzip?

OSS supports gzip compression for static files (such as HTML, Javascript, XML, and json files), but the following conditions must be met:

- File size must be equal to or greater than 1 KB.
- Content-Type must be text/cache-manifest, text/xml, text/plain, text/css, application/javascript, application/x-javascript, application/rss+xml, application/json, or text/json.

Usage:

- Console: Upload a file through the OSS console, and use the generated URL to obtain the file.
- API: Take Java SDK for example, Multipart upload and PutObject support gzip compression.

Note: In the OSS console > Files > Set HTTP Header, if Content-Encoding is set to gzip, then the file must be compressed with gzip before being uploaded to OSS. Otherwise, some browsers cannot open it.

How to get the link address of the object in OSS

If you upload file 123.jpg in the Bucket of aliyun-abc in China East 1 (Hangzhou), how can you obtain the URL of 123.jpg after it is successfully uploaded? The procedure is as follows:

Obtain the access domain name of the Bucket: **Bucket name.EndPoint**.

- Internet access domain name: aliyun-abc.oss-cn-hangzhou.aliyuncs.com
- Intranet access domain name: aliyun-abc.oss-cn-hangzhou-internal.aliyuncs.com

Confirm the permissions of the Object.

If the Object has the **Public Read** or **Public Read/Write** permission, the format of the access URL is **domain name/Object name**.

- URL for access from a public network: aliyun-abc.oss-cn-

hangzhou.aliyuncs.com/123.jpg

- URL for access from Alibaba Cloud intranet ECS/VPC: aliyun-abc.oss-cn-hangzhou-internal.aliyuncs.com/123.jpg

If the Object has the **Private** permission, perform the signature operation **domain name/Object name? signature parameter**. The procedure is as follows:

Obtain the Object URL through the console. For more information, see [Obtain the file access address](#).

Call the API signature algorithm. For more information, see [Include a signature in the URL](#).

Use the SDK to generate a signature:

- Java: Authorized access
- Python: Authorized access
- Android: Access control
- iOS: Access control
- .Net: Authorized access
- JavaScript: Authorized access
- PHP: Authorized access
- C: Authorized access
- Ruby: Authorized access
- Go: Authorized access

Signature

Is there a JAVA code example of presigned URL?

Example codes of the presigned URL are as follows:

```
// Generate a presigned URL
```

```
Date expires = new Date (new Date().getTime() + 1000 * 60); // 1 minute to expire
```

```
GeneratePresignedUrlRequest generatePresignedUrlRequest = new GeneratePresignedUrlRequest(bucketName,
key);

generatePresignedUrlRequest.setExpiration(expires);

URL url = client.generatePresignedUrl(generatePresignedUrlRequest);

System.out.println(url.toString());
```

This section of code assumes that the object with the specified bucketName and key has been uploaded to the OSS, and the users can set the expiration time of expires. More detailed operations can be found in help on `OSSClient#generatePresignedUrl` and `GeneratePresignedUrlRequest`.

API/SDK

What is the instruction on the "Date" field of GMT format in OSS API Reference?

In the OSS API Reference, fields with "DATE" to indicate the time must follow the GMT format supported by HTTP1.1.

HTTP1.1 protocol instructions: <http://www.w3.org/Protocols/rfc2616/rfc2616-sec3.html>

date1 = 2DIGIT SP month SP 4DIGIT

; day month year (e.g., 02 Jun 1982)

date2 = 2DIGIT "-" month "-" 2DIGIT

; day-month-year (e.g., 02-Jun-82)

date3 = month SP (2DIGIT | (SP 1DIGIT))

; month day (e.g., Jun 2) <----- Attention: No leading zero is located before the date and a space is required to hold the place.