Preparing the source and target

Cloud Manager

Ben Cammett
July 04, 2021

Table of Contents

Preparing the source and target

Prepare to sync data by verifying that your source and target are supported and setup.

Supported sync relationships

Cloud Sync enables you to sync data from a source to a target (this is called a *sync relationship*). You should understand the supported relationships before you get started.

Source location	Supported target locations
AWS EFS	• AWS EFS • AWS S3 • Azure Blob • Azure NetApp Files • Cloud Volumes ONTAP • Cloud Volumes Service • Google Cloud Storage • IBM Cloud Object Storage • NFS server • On-premises ONTAP cluster • SMB server • StorageGRID
AWS S3	• AWS EFS • AWS S3 • Azure Blob • Azure NetApp Files • Cloud Volumes ONTAP • Cloud Volumes Service • Google Cloud Storage • IBM Cloud Object Storage • NFS server • On-premises ONTAP cluster • SMB Server • StorageGRID

Source location	Supported target locations
Azure Blob	• AWS EFS • AWS S3 • Azure Blob • Azure NetApp Files • Cloud Volumes ONTAP • Cloud Volumes Service • Google Cloud Storage • IBM Cloud Object Storage • NFS server • On-premises ONTAP cluster • SMB Server • StorageGRID
Azure NetApp Files	• AWS EFS • AWS S3 • Azure Blob • Azure NetApp Files • Cloud Volumes ONTAP • Cloud Volumes Service • Google Cloud Storage • IBM Cloud Object Storage • NFS server • On-premises ONTAP cluster • SMB Server • StorageGRID

Source location	Supported target locations
Cloud Volumes ONTAP	• AWS EFS • AWS S3 • Azure Blob • Azure NetApp Files • Cloud Volumes ONTAP • Cloud Volumes Service • Google Cloud Storage • IBM Cloud Object Storage • NFS server • On-premises ONTAP cluster • SMB Server • StorageGRID
Cloud Volumes Service	• AWS EFS • AWS S3 • Azure Blob • Azure NetApp Files • Cloud Volumes ONTAP • Cloud Volumes Service • Google Cloud Storage • IBM Cloud Object Storage • NFS server • On-premises ONTAP cluster • SMB Server • StorageGRID

Source location	Supported target locations
Google Cloud Storage	• AWS EFS • AWS S3 • Azure Blob • Azure NetApp Files • Cloud Volumes ONTAP • Cloud Volumes Service • Google Cloud Storage • IBM Cloud Object Storage • NFS server • On-premises ONTAP cluster • ONTAP S3 Storage • SMB Server • StorageGRID
IBM Cloud Object Storage	• AWS EFS • AWS S3 • Azure Blob • Azure NetApp Files • Cloud Volumes ONTAP • Cloud Volumes Service • Google Cloud Storage • IBM Cloud Object Storage • NFS server • On-premises ONTAP cluster • SMB Server • StorageGRID

Source location	Supported target locations
NFS server	• AWS EFS • AWS S3 • Azure Blob • Azure NetApp Files • Cloud Volumes ONTAP • Cloud Volumes Service • Google Cloud Storage • IBM Cloud Object Storage • NFS server • On-premises ONTAP cluster • SMB Server • StorageGRID
On-prem ONTAP cluster	• AWS EFS • AWS S3 • Azure Blob • Azure NetApp Files • Cloud Volumes ONTAP • Cloud Volumes Service • Google Cloud Storage • IBM Cloud Object Storage • NFS server • On-premises ONTAP cluster • SMB Server • StorageGRID
ONTAP S3 Storage	• Google Cloud Storage • SMB server • StorageGRID • ONTAP S3 Storage
SFTP [1]	S3

Source location	Supported target locations
SMB server	• AWS EFS • AWS S3 • Azure Blob • Azure NetApp Files • Cloud Volumes ONTAP • Cloud Volumes Service • Google Cloud Storage • IBM Cloud Object Storage • NFS server • On-premises ONTAP cluster • ONTAP S3 Storage • SMB Server • StorageGRID
StorageGRID	• AWS EFS • AWS S3 • Azure Blob • Azure NetApp Files • Cloud Volumes ONTAP • Cloud Volumes Service • Google Cloud Storage • IBM Cloud Object Storage • NFS server • On-premises ONTAP cluster • ONTAP S3 Storage • SMB Server • StorageGRID

Notes:

1. Cloud Sync supports sync relationships from SFTP to S3 by using the API only.
2. You can choose a specific Azure Blob storage tier when a Blob container is the target:
    - Hot storage
    - Cool storage
3. You can choose a specific S3 storage class when AWS S3 is the target:
    - Standard (this is the default class)
    - Intelligent-Tiering

    - Standard-Infrequent Access
    - One Zone-Infrequent Access
    - Glacier
    - Glacier Deep Archive
4. You can choose a specific storage class when a Google Cloud Storage bucket is the target:
    - Standard
    - Nearline
    - Coldline
    - Archive

Source and target requirements

Verify that your source and targets meet the following requirements.

Networking

- The source and target must have a network connection to the data broker.

  For example, if an NFS server is in your data center and the data broker is in AWS, then you need a network connection (VPN or Direct Connect) from your network to the VPC.

- NetApp recommends configuring the source, target, and data broker to use a Network Time Protocol (NTP) service. The time difference between the three components should not exceed 5 minutes.

Target directory

When you create a sync relationship, Cloud Sync enables you to select an existing target directory and then optionally create a new folder inside that directory. So be sure that your preferred target directory already exists.

Permissions to read directories

In order to show every directory or folder in a source or target, Cloud Sync needs read permissions on the directory or folder.

NFS

Permissions must be defined on the source/target with uid/gid on files and directories.

Object storage

- For AWS and Google Cloud, the data broker must have list object permissions (these permissions are provided by default if you follow the data broker installation steps).
- For Azure, StorageGRID, and IBM, the credentials that you enter when setting up a sync relationship must have list object permissions.

SMB

The SMB credentials that you enter when setting up a sync relationship must have list folder permissions.

The data broker ignores the following directories by default: .snapshot, ~snapshot, .copy-offload

AWS S3 bucket requirements

Make sure that your AWS S3 bucket meets the following requirements.

Supported data broker locations for AWS S3

Sync relationships that include S3 storage require a data broker deployed in AWS or on your premises. In either case, Cloud Sync prompts you to associate the data broker with an AWS account during installation.

- Learn how to deploy the AWS data broker
- Learn how to install the data broker on a Linux host

Supported AWS regions

All regions are supported except for the China and GovCloud (US) regions.

Permissions required for S3 buckets in other AWS accounts

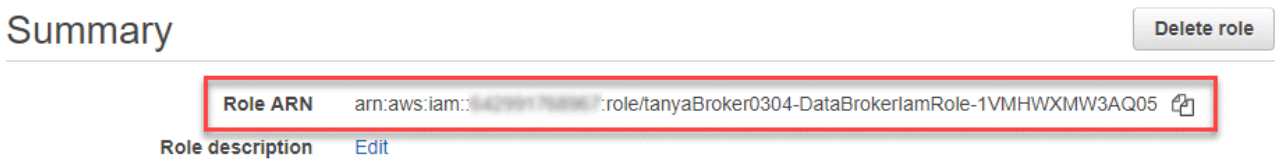
When setting up a sync relationship, you can specify an S3 bucket that resides in an AWS account that isn't associated with the data broker.

The permissions included in this JSON file must be applied to that S3 bucket so the data broker can access it. These permissions enable the data broker to copy data to and from the bucket and to list the objects in the bucket.

Note the following about the permissions included in the JSON file:

1. *<BucketName>* is the name of the bucket that resides in the AWS account that isn't associated with the data broker.
2. *<RoleARN>* should be replaced with one of the following:
   - If the data broker was manually installed on a Linux host, *RoleARN* should be the ARN of the AWS user for which you provided AWS credentials when deploying the data broker.
   - If the data broker was deployed in AWS using the CloudFormation template, *RoleARN* should be the ARN of the IAM role created by the template.

     You can find the Role ARN by going to the EC2 console, selecting the data broker instance, and clicking the IAM role from the Description tab. You should then see the Summary page in the IAM console that contains the Role ARN.



Azure Blob storage requirements

Make sure that your Azure Blob storage meets the following requirements.

Supported data broker locations for Azure Blob

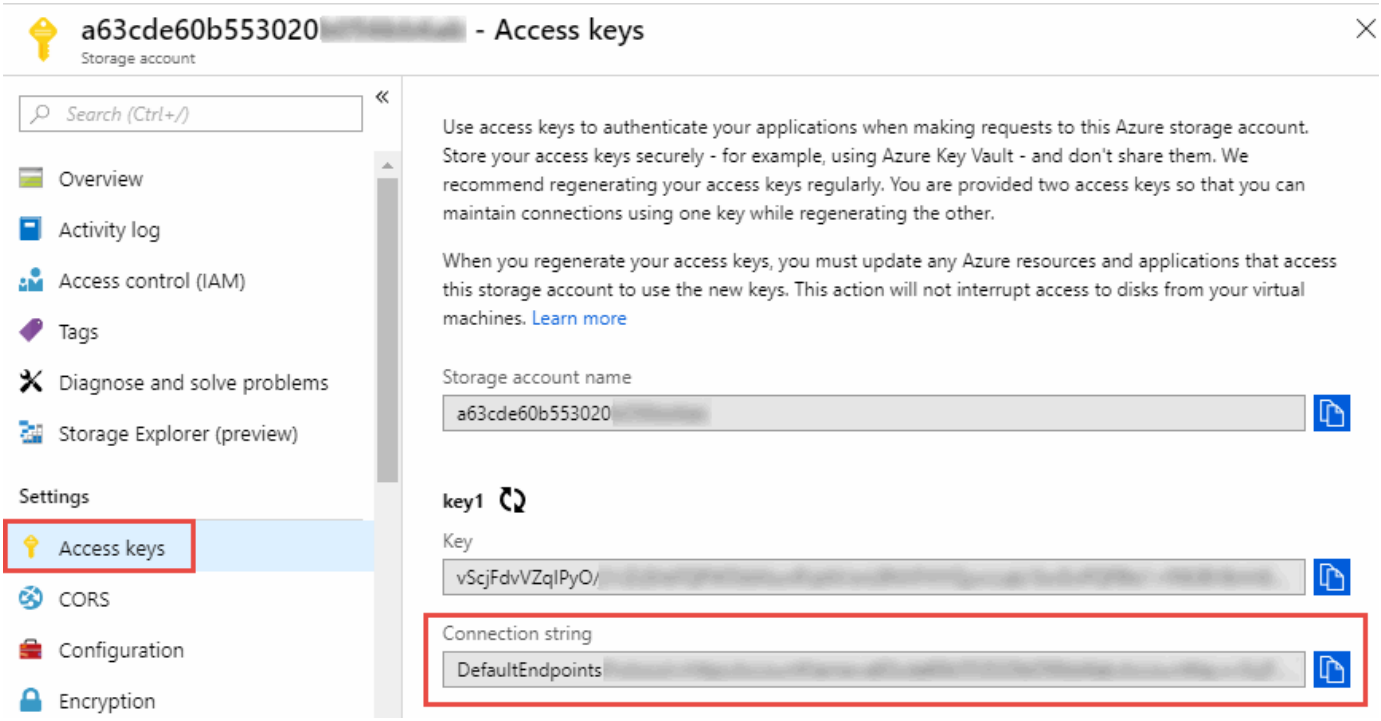
The data broker can reside in any location when a sync relationship includes Azure Blob storage.

Supported Azure regions

All regions are supported except for the China, US Gov, and US DoD regions.

Connection string required for relationships that include Azure Blob and NFS/SMB

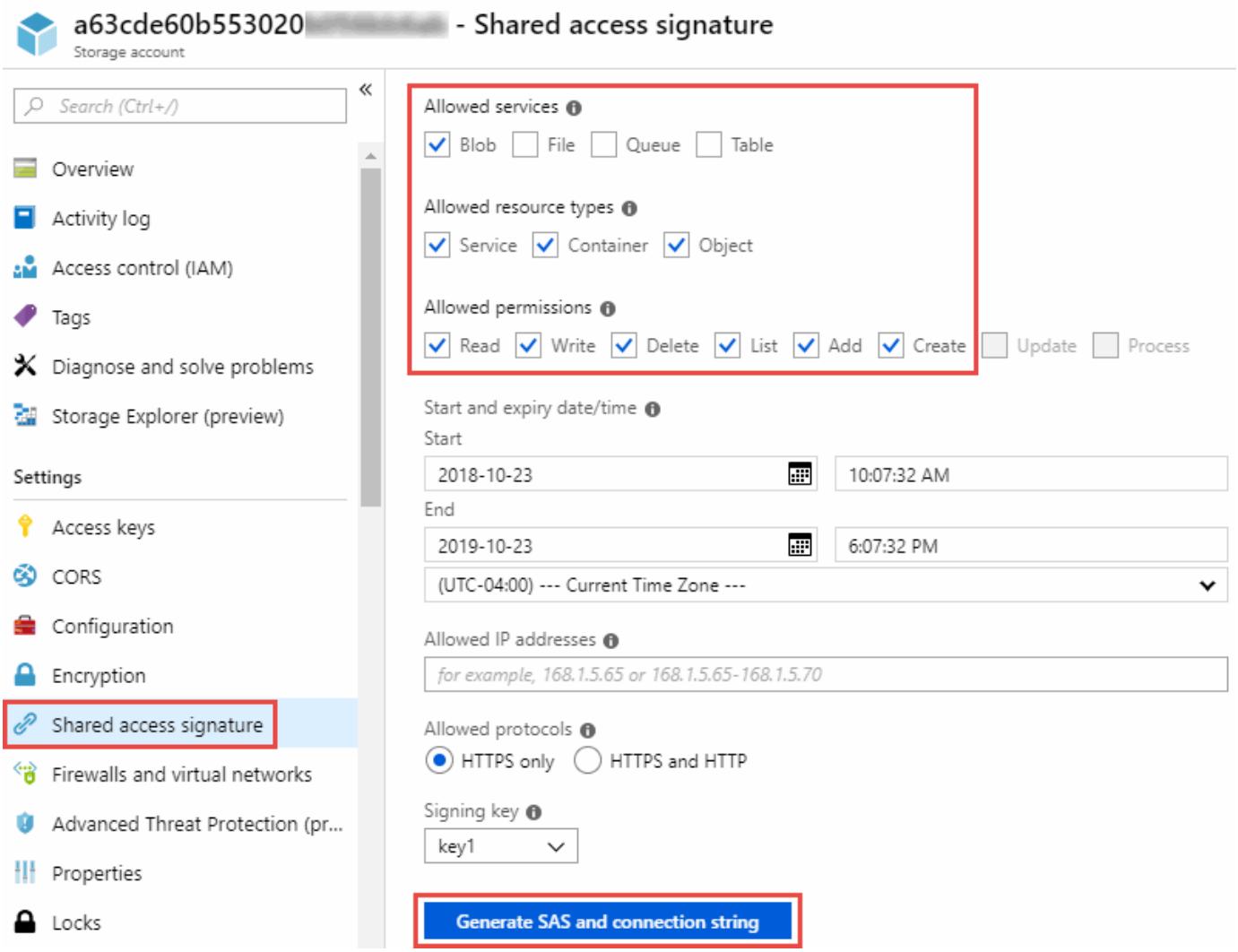
When creating a sync relationship between an Azure Blob container and an NFS or SMB server, you need to provide Cloud Sync with the storage account connection string:



If you want to sync data between two Azure Blob containers, then the connection string must include a shared access signature (SAS). You also have the option to use a SAS when syncing between a Blob container and an NFS or SMB server.

The SAS must allow access to the Blob service and all resource types (Service, Container, and Object). The SAS must also include the following permissions:

- For the source Blob container: Read and List
- For the target Blob container: Read, Write, List, Add, and Create



Azure NetApp Files requirement

Use the Premium or Ultra service level when you sync data to or from Azure NetApp Files. You might experience failures and performance issues if the disk service level is Standard.

Consult a solutions architect if you need help determining the right service level. The volume size and volume tier determines the throughput that you can get.

Learn more about Azure NetApp Files service levels and throughput.

Google Cloud Storage bucket requirements

Make sure that your Google Cloud Storage bucket meets the following requirements.

Supported data broker locations for Google Cloud Storage

Sync relationships that include Google Cloud Storage require a data broker deployed in GCP or on your premises. Cloud Sync guides you through the data broker installation process when you create a sync relationship.

- Learn how to deploy the GCP data broker

- Learn how to install the data broker on a Linux host

Supported GCP regions

All regions are supported.

Permissions required for buckets in other Google Cloud projects

When setting up a sync relationship, you can choose from Google Cloud buckets in different projects, if you provide the required permissions to the data broker's service account. Learn how to set up the service account.

ONTAP requirements

If the sync relationship includes Cloud Volumes ONTAP or an on-prem ONTAP cluster and you selected NFSv4 or later, then you'll need to enable NFSv4 ACLs on the ONTAP system. This is required to copy the ACLs.

Permissions for a SnapMirror destination

If the source for a sync relationship is a SnapMirror destination (which is read-only), "read/list" permissions are sufficient to sync data from the source to a target.

NFS server requirements

- The NFS server can be a NetApp system or a non-NetApp system.
- The file server must allow the data broker host to access the exports.
- NFS versions 3, 4.0, 4.1, and 4.2 are supported.

  The desired version must be enabled on the server.

- If you want to sync NFS data from an ONTAP system, ensure that access to the NFS export list for an SVM is enabled (vserver nfs modify -vserver *svm_name* -showmount enabled).

  The default setting for showmount is *enabled* starting with ONTAP 9.2.

ONTAP S3 Storage requirements

When you set up a sync relationship that includes ONTAP S3 Storage, you'll need to provide the following:

- The IP address of the LIF that's connected to ONTAP S3
- The access key and secret key that ONTAP is configured to use

SMB server requirements

- The SMB server can be a NetApp system or a non-NetApp system.
- The file server must allow the data broker host to access the exports.
- SMB versions 1.0, 2.0, 2.1, 3.0 and 3.11 are supported.
- Grant the "Administrators" group with "Full Control" permissions to the source and target folders.

  If you don't grant this permission, then the data broker might not have sufficient permissions to get the

ACLs on a file or directory. If this occurs, you'll receive the following error: "getxattr error 95"

SMB limitation for hidden directories and files

An SMB limitation affects hidden directories and files when syncing data between SMB servers. If any of the directories or files on the source SMB server were hidden through Windows, the hidden attribute isn't copied to the target SMB server.

SMB sync behavior due to case-insensitivity limitation

The SMB protocol is case-insensitive, which means uppercase and lowercase letters are treated as being the same. This behavior can result in overwritten files and directory copy errors, if a sync relationship includes an SMB server and data already exists on the target.

For example, let's say that there's a file named "a" on the source and a file named "A" on the target. When Cloud Sync copies the file named "a" to the target, file "A" is overwritten by file "a" from the source.

In the case of directories, let's say that there's a directory named "b" on the source and a directory named "B" on the target. When Cloud Sync tries to copy the directory named "b" to the target, Cloud Sync receives an error that says the directory already exists. As a result, Cloud Sync always fails to copy the directory named "b."

The best way to avoid this limitation is to ensure that you sync data to an empty directory.