Backing up on-premises ONTAP data to the public cloud

Cloud Manager

Tom Onacki
June 07, 2021

Table of Contents

Backing up on-premises ONTAP data to the public cloud

Complete a few steps to get started backing up data from your on-premises ONTAP systems to low-cost object storage in the public cloud. This includes creating backup files on Amazon S3, Azure Blob, and Google Cloud Storage.

If you have a NetApp StorageGRID system and want to create your backups there, see how to back up on-premises ONTAP data to the private cloud.

TIP

In most cases you'll use Cloud Manager for all backup and restore operations. However, starting with ONTAP 9.9.1 you can initiate volume backup operations of your on-premises ONTAP clusters using ONTAP System Manager. See how to use System Manager to back up your volumes to the cloud using Cloud Backup.

A Beta feature released in January 2021 allows you to run compliance scans on the backed up volumes from your on-premises systems. Typically, compliance scans are free up to 1 TB of data, and then a cost for the service is applied for data over 1 TB. When combining Backup and Data Sense for your on-premises volumes, the cost for scans on those on-prem volumes is free. Learn more about how Cloud Data Sense can get your business applications and cloud environments privacy ready.

Quick start

Get started quickly by following these steps, or scroll down to the remaining sections for full details.

Verify support for your configuration

- You have discovered the on-premises cluster and added it to a working environment in Cloud Manager. See Discovering ONTAP clusters for details.
  - The cluster is running ONTAP 9.7P5 or later.
  - The cluster has a SnapMirror license — it is included as part of the PREM or Data Protection bundle.
- You have subscribed to the Azure, the AWS, or the Google Cloud Manager Marketplace Backup offering, or you have purchased and activated a Cloud Backup BYOL license from NetApp.
- You have a valid cloud provider subscription for the object storage space where your backups will be located.
- For AWS and GCP, you need to have an account that has an access key and the required permissions so the ONTAP cluster can back up the data.

Enable Cloud Backup on the system

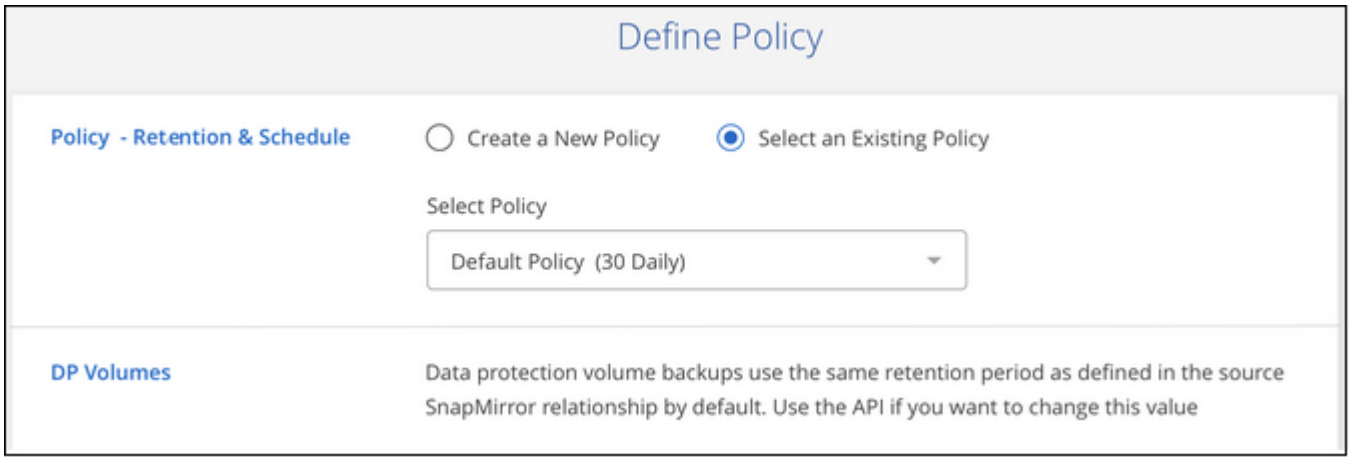
Select the working environment and click **Enable** next to the Backup & Compliance service in the right-panel, and then follow the setup wizard.

Select the cloud provider and enter the provider details

Select the provider and then enter the provider details. You also need to specify the IPspace in the ONTAP cluster where the volumes reside.

Define the backup policy

The default policy backs up volumes every day and retains the most recent 30 backup copies of each volume. Change to daily, weekly, or monthly backups, or select one of the system-defined policies that provide more options. You can also change the number of backup copies to retain.



Select the volumes that you want to back up

Identify which volumes you want to back up from the cluster.

Activate Compliance scans on the backed up volumes (optional)

Choose whether you want to have Cloud Data Sense scan the volumes that are backed up in the cloud.

Restore your data, as needed

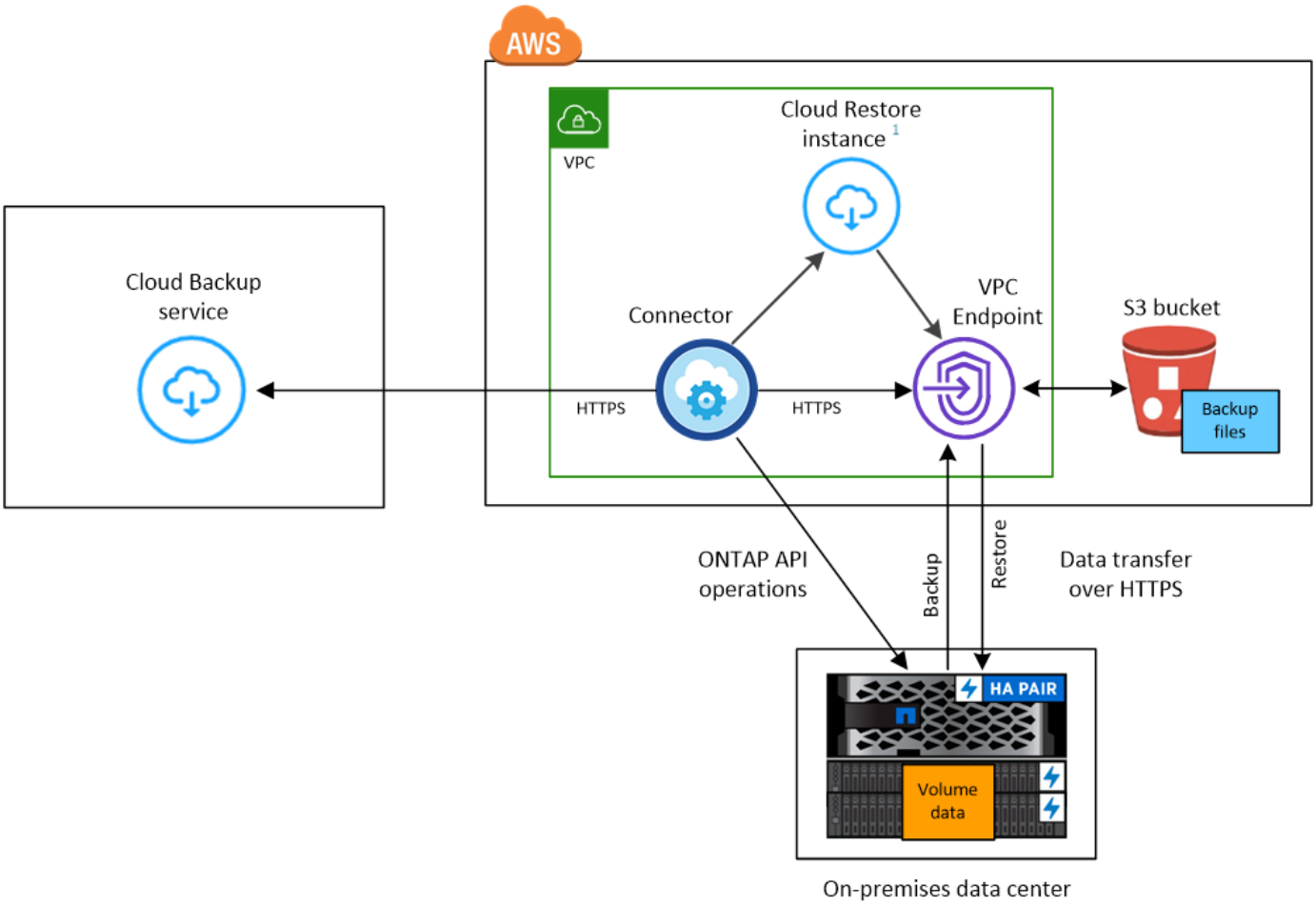
Choose to restore an entire backup to a new volume, or to restore individual files from the backup to an existing volume. You can restore data to a Cloud Volumes ONTAP system that is using the same cloud provider, or to an on-premises ONTAP system.

See Restoring volume data from backup files for details.

Overview diagrams

The following diagrams show each component when backing up an on-prem ONTAP system to object storage and the connections that you need to prepare between them.

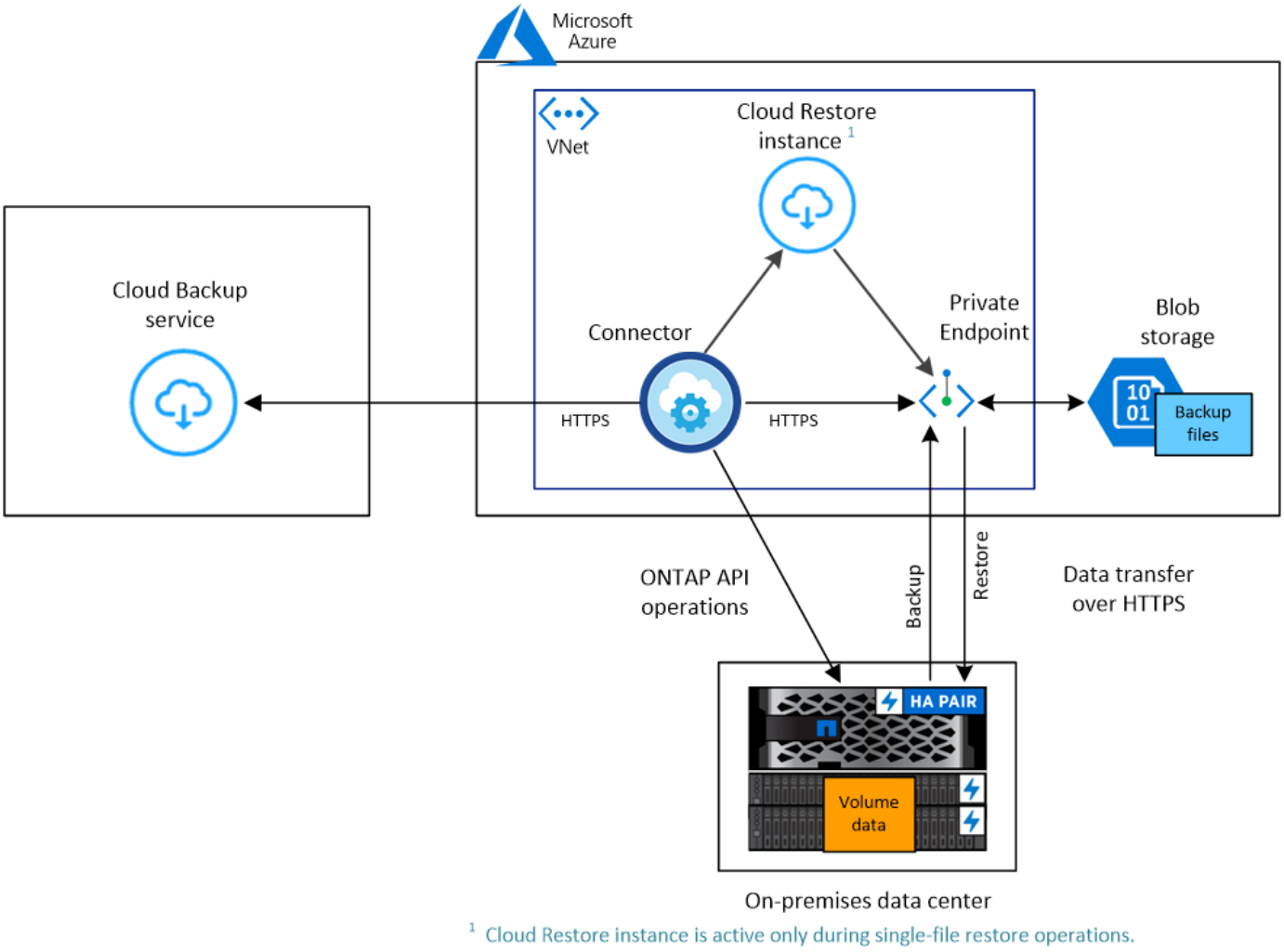
Amazon S3 configuration:



[1] Cloud Restore instance is active only during single-file restore operations.

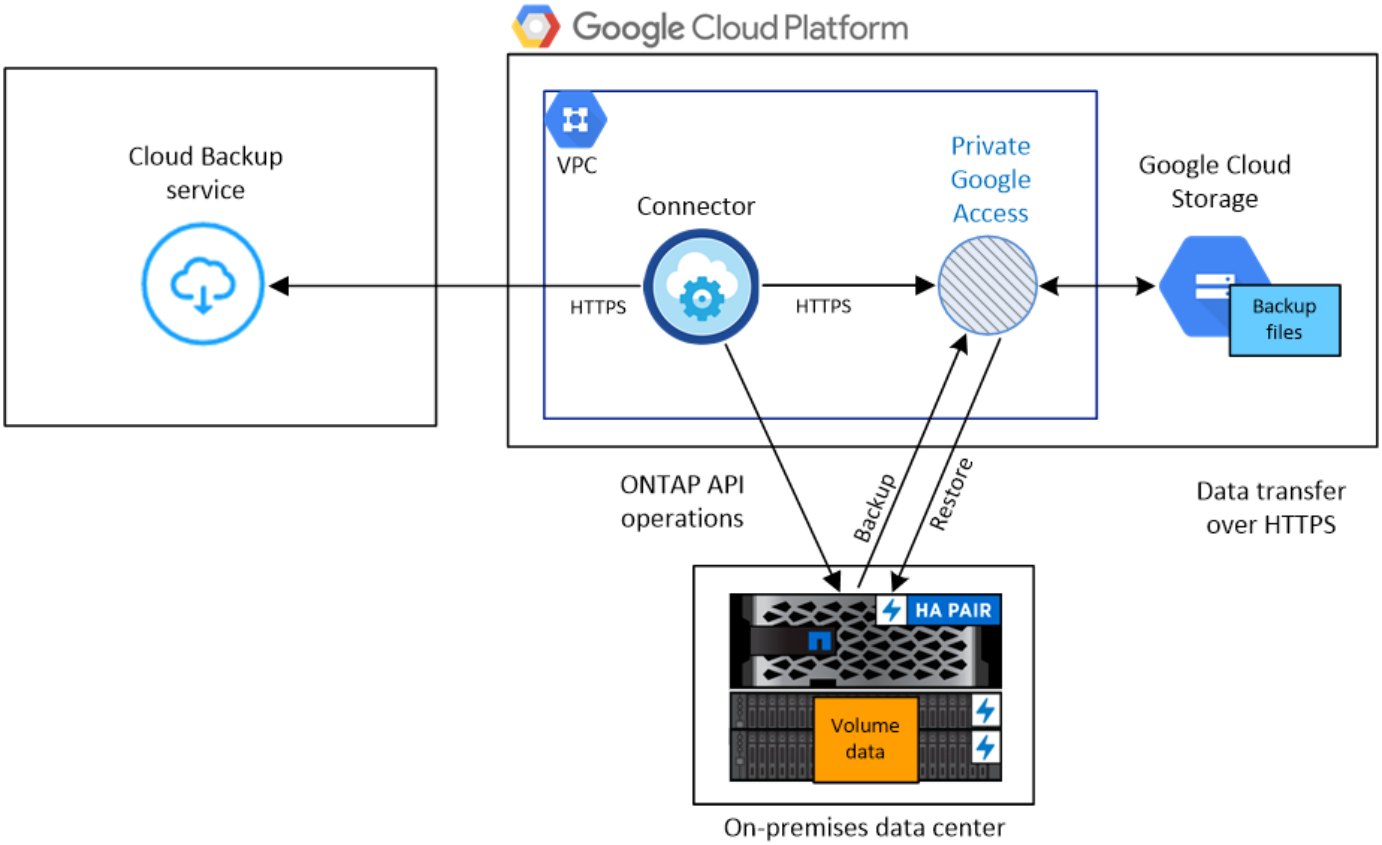
Note that when the Cloud Restore instance is deployed in the cloud, it is located in the same subnet as the Connector.

Azure Blob configuration:



Note that when the Cloud Restore instance is deployed in the cloud, it is located in the same subnet as the Connector.

Google Cloud Storage configuration:



Note that the Cloud Restore instance is not shown in this diagram because single-file restore is not currently supported in GCP.

Requirements

Read the following requirements to make sure you have a supported configuration before you start backing up on-premises volumes to object storage.

Preparing your ONTAP clusters

You need to discover your on-premises ONTAP clusters in Cloud Manager before you can start backing up volume data.

Learn how to discover a cluster.

ONTAP requirements

- ONTAP 9.7P5 and later.
- A SnapMirror license (included as part of the PREM or Data Protection bundle).

  See how to manage your cluster licenses.

- Time and time zone are set correctly.

  See how to configure your cluster time.

Cluster networking requirements

- The ONTAP cluster initiates an HTTPS connection over port 443 to the cloud object storage.

  ONTAP reads and writes data to and from object storage. The object storage never initiates, it just responds.

- An inbound connection is required from the Connector, which can reside in an AWS VPC, Azure VNet, or Google Cloud Platform VPC; depending on the object storage provider you are using.
- An intercluster LIF is required on each ONTAP node that hosts the volumes you want to back up. The LIF must be associated with the *IPspace* that ONTAP should use to connect to object storage. Learn more about IPspaces.

  When you set up Cloud Backup, you are prompted for the IPspace to use. You should choose the IPspace that each LIF is associated with. That might be the "Default" IPspace or a custom IPspace that you created.

- Node and intercluster LIFs are able to access the internet.
- DNS servers have been configured for the storage VM where the volumes are located.

  See how to configure DNS services for the SVM.

- Update firewall rules, if necessary, to allow Cloud Backup service connections from ONTAP to object storage through port 443 and name resolution traffic from the storage VM to the DNS server over port 53 (TCP/UDP).

Creating or switching Connectors

A Connector is required to back up data to the cloud, and the Connector must be in the same cloud provider as the destination object storage. For example, when backing up data to AWS S3 you must use a Connector that's in an AWS VPC. You cannot use a Connector that is deployed on-premises. You'll either need to create a new Connector or make sure that the currently selected Connector resides in the correct provider.

- Learn about Connectors
- Creating a Connector in AWS
- Creating a Connector in Azure
- Creating a Connector in GCP
- Switching between Connectors

Preparing networking for the Connector

Ensure that the Connector has the required networking connections.

Steps

1. Ensure that the network where the Connector is installed enables the following connections:
   - An outbound internet connection to the Cloud Backup service over port 443 (HTTPS)
   - An HTTPS connection over port 443 to your object storage (S3, Blob, or Google)
   - An HTTPS connection over port 443 to your ONTAP clusters
2. Enable an endpoint to your object storage:
   - For AWS: Enable a VPC Endpoint to S3. This is needed if you have a Direct Connect or VPN

connection from your ONTAP cluster to the VPC and you want communication between the Connector and S3 to stay in your AWS internal network.

- For Azure: Enable a VNet Private Endpoint to Azure storage. This is needed if you have an ExpressRoute or VPN connection from your ONTAP cluster to the VNet and you want communication between the Connector and Blob storage to stay in your virtual private network.
- For Google: Enable Private Google Access on the subnet where you plan to deploy the Service Connector. Private Google Access is needed if you have a direct connection from your ONTAP cluster to the VPC and you want communication between the Connector and Google Cloud Storage to stay in your virtual private network.

  Note that Private Google Access works with VM instances that have only internal (private) IP addresses (no external IP addresses).

Supported regions

You can create backups from on-premises systems to the public cloud in all regions where Cloud Volumes ONTAP is supported. You specify the region where the backups will be stored when you set up the service.

License requirements

For Cloud Backup PAYGO licensing, you'll need a subscription to the Azure, the AWS, or the Google Cloud Manager Marketplace Backup offering before you enable Cloud Backup. Billing for Cloud Backup is done through this subscription.

For Cloud Backup BYOL licensing, you do not need a subscription. You need the serial number from NetApp that enables you to use the service for the duration and capacity of the license. See Managing your Backup BYOL license.

And you need to have a subscription from your cloud provider for the object storage space where your backups will be located.

Preparing Amazon S3 for backups

When you are using Amazon S3, you must configure permissions for Cloud Manager to access the S3 bucket, and you must configure permissions so the on-premises ONTAP cluster can access the S3 bucket.

Steps

1. Provide the following S3 permissions (from the latest Cloud Manager policy) to the IAM role that provides Cloud Manager with permissions:

```
{
            "Sid": "backupPolicy",
            "Effect": "Allow",
            "Action": [
                "s3:DeleteBucket",
                "s3:GetLifecycleConfiguration",
                "s3:PutLifecycleConfiguration",
                "s3:PutBucketTagging",
                "s3:ListBucketVersions",
                "s3:GetObject",
                "s3:ListBucket",
                "s3:ListAllMyBuckets",
                "s3:GetBucketTagging",
                "s3:GetBucketLocation",
                "s3:GetBucketPolicyStatus",
                "s3:GetBucketPublicAccessBlock",
                "s3:GetBucketAcl",
                "s3:GetBucketPolicy",
                "s3:PutBucketPublicAccessBlock"
            ],
            "Resource": [
                "arn:aws:s3:::netapp-backup-*"
            ]
        },
```

2. Provide the following permissions to the IAM user so that the ONTAP cluster can back up data to S3.

```
"s3:ListAllMyBuckets",
"s3:ListBucket",
"s3:GetBucketLocation",
"s3:GetObject",
"s3:PutObject",
"s3:DeleteObject"
```

   See the AWS Documentation: Creating a Role to Delegate Permissions to an IAM User for details.

3. Provide the following permissions for the Cloud Restore instance:

```
            "Action": [
                "ec2:DescribeInstanceTypeOfferings",
                "ec2:StartInstances",
                "ec2:StopInstances",
                "ec2:TerminateInstances"
            ],
```

4. If your virtual or physical network uses a proxy server for internet access, ensure that the Cloud Restore instance has outbound internet access to contact the following endpoints.

Endpoints	Purpose
http://amazonlinux.us-east-1.amazonaws.com/2/extras/docker/stable/x86_64/4bf88ee77c395ffe1e0c3ca68530dfb3a683ec65a4a1ce9c0ff394be50e922b2/	CentOS package for the Cloud Restore Instance AMI.
http://cloudmanagerinfraprod.azurecr.io https://cloudmanagerinfraprod.azurecr.io	Cloud Restore Instance image repository.

5. Create or locate an access key.

   Cloud Backup passes the access key on to the ONTAP cluster. The credentials are not stored in the Cloud Backup service.

   See the AWS Documentation: Managing Access Keys for IAM Users for details.

Preparing Azure Blob storage for backups

The Cloud Restore virtual machine requires outbound internet access. If your virtual or physical network uses a proxy server for internet access, ensure that the instance has outbound internet access to contact the following endpoints.

Endpoints	Purpose
http://olcentgbl.trafficmanager.net https://olcentgbl.trafficmanager.net	Provides CentOS packages for the Cloud Restore virtual machine.
http://cloudmanagerinfraprod.azurecr.io https://cloudmanagerinfraprod.azurecr.io	Cloud Restore virtual machine image repository.

Preparing Google Cloud Storage for backups

When you set up backup, you need to provide storage access keys for a service account that has Storage Admin permissions. A service account enables Cloud Backup to authenticate and access Cloud Storage buckets used to store backups. The keys are required so that Google Cloud Storage knows who is making the request.

Steps

1. Create a service account that has the predefined Storage Admin role.

2. Go to GCP Storage Settings and create access keys for the service account:
    a. Select a project, and click **Interoperability**. If you haven't already done so, click **Enable interoperability access**.
    b. Under **Access keys for service accounts**, click **Create a key for a service account**, select the service account that you just created, and click **Create Key**.

       You'll need to enter the keys in Cloud Backup later when you configure the backup service.

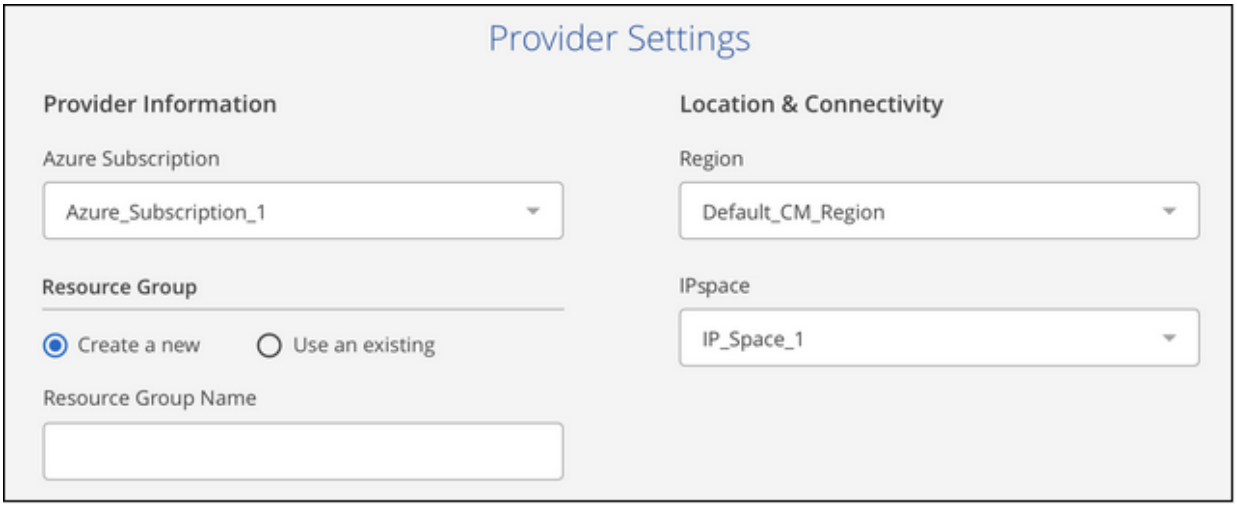
Enabling Cloud Backup

Enable Cloud Backup at any time directly from the on-premises working environment.

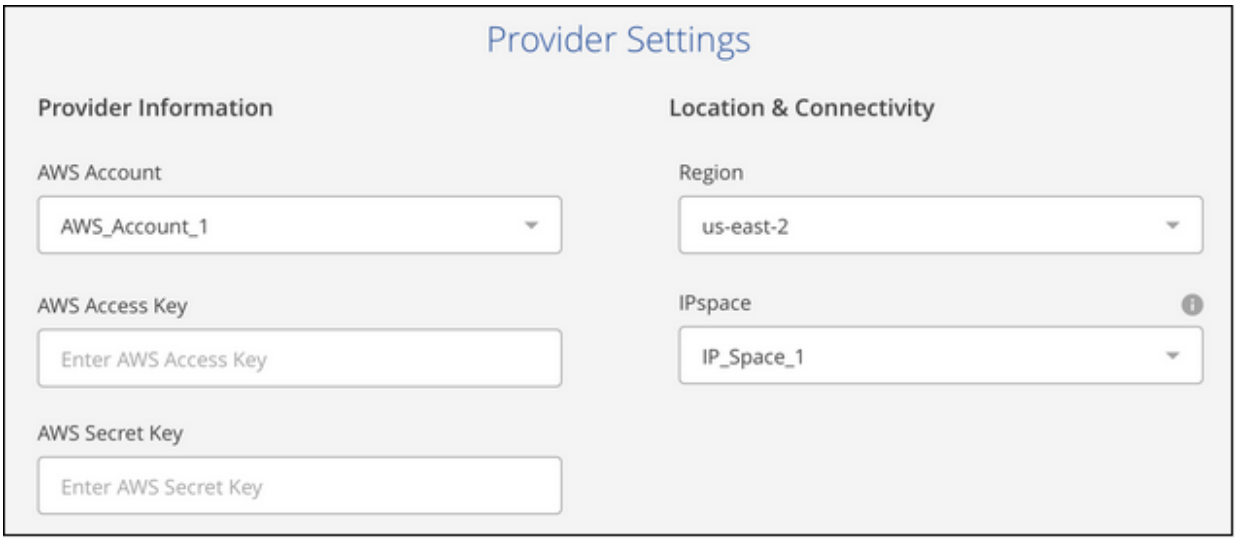
Steps

1. From the Canvas, select the working environment and click **Enable** next to the Backup & Compliance service in the right-panel.

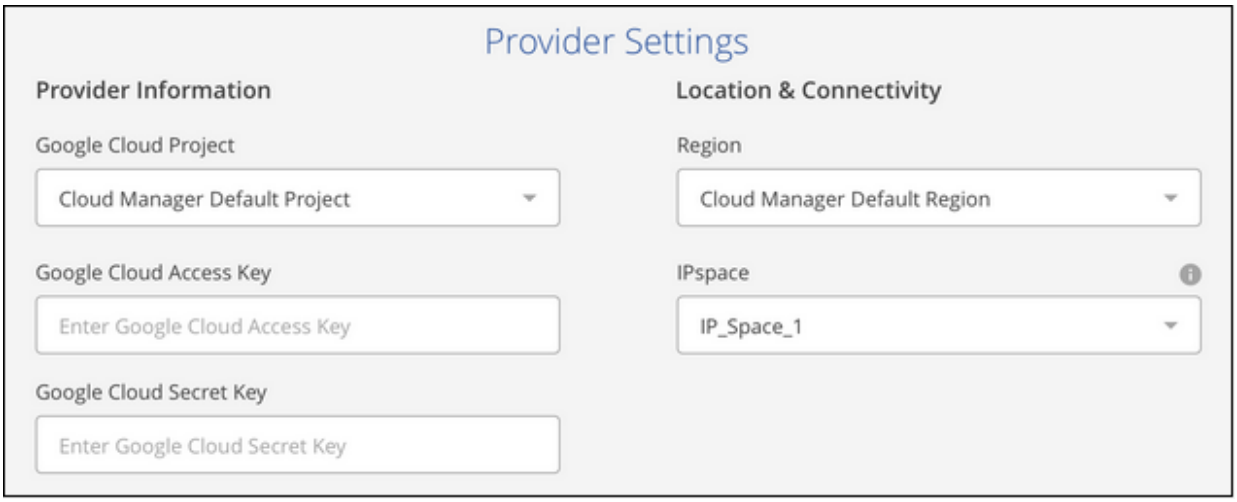
2. Select the provider, click **Next**, and then enter the provider details:
    ◦ For Azure, enter:
        a. The Azure subscription used for backups and the Azure region where the backups will be stored.
        b. The resource group - you can create a new resource group or select and existing resource group.
        c. The IPspace in the ONTAP cluster where the volumes you want to back up reside.



    ◦ For AWS, enter:
        a. The AWS Account, the AWS Access Key, and the Secret Key used to store the backups.
        b. The AWS region where the backups will be stored.
        c. The IPspace in the ONTAP cluster where the volumes you want to back up reside.



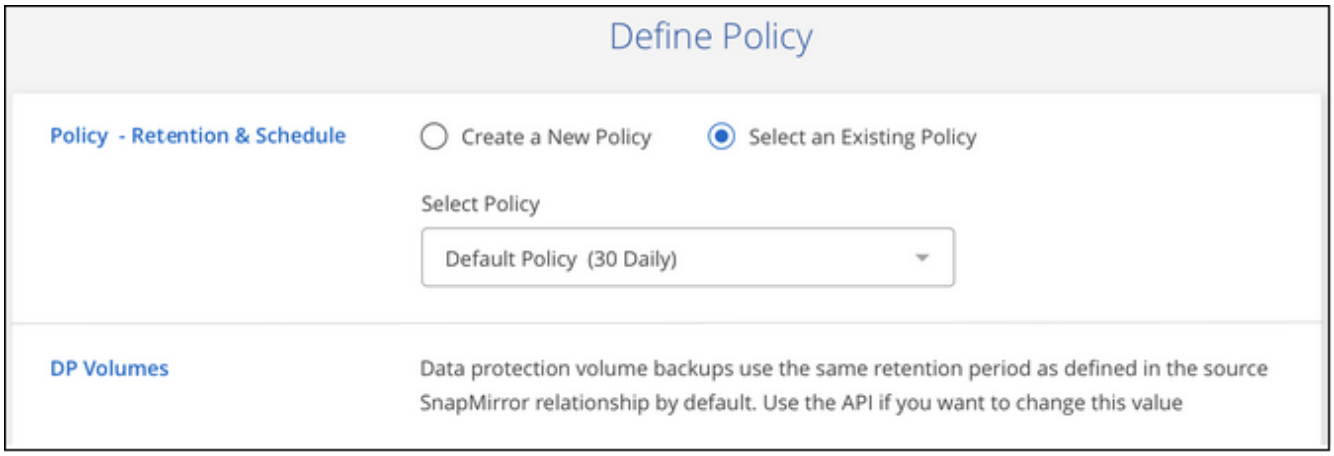
- For Google, enter:
  a. The Google Cloud Project where you want the Google Cloud Storage bucket to be created for backups. This can be a different Project than where Cloud Manager resides. (The Project must have a Service Account that has the predefined Storage Admin role.)
  b. The Google Access Key and Secret Key used to store the backups.
  c. The Google region where the backups will be stored. This can be a different region than where Cloud Manager resides.
  d. The IPspace in the ONTAP cluster where the volumes you want to back up reside.



- For StorageGRID, see how to back up on-premises ONTAP data to the private cloud.

  Note that you cannot change this information after the service has started.

3. Click **Next** after you've entered the provider details.
4. In the *Define Policy* page, select an existing backup schedule and retention value, or define a new backup policy, and click **Next**.

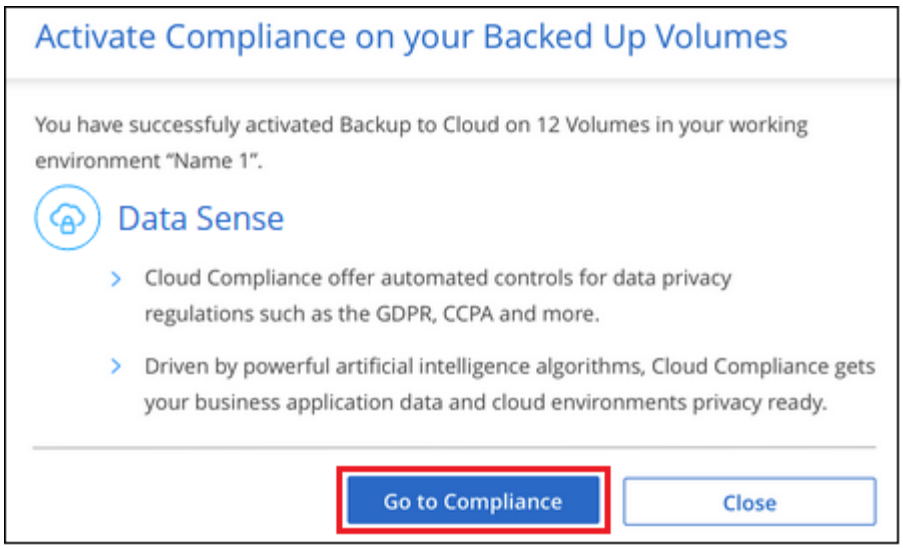


See the list of existing policies.

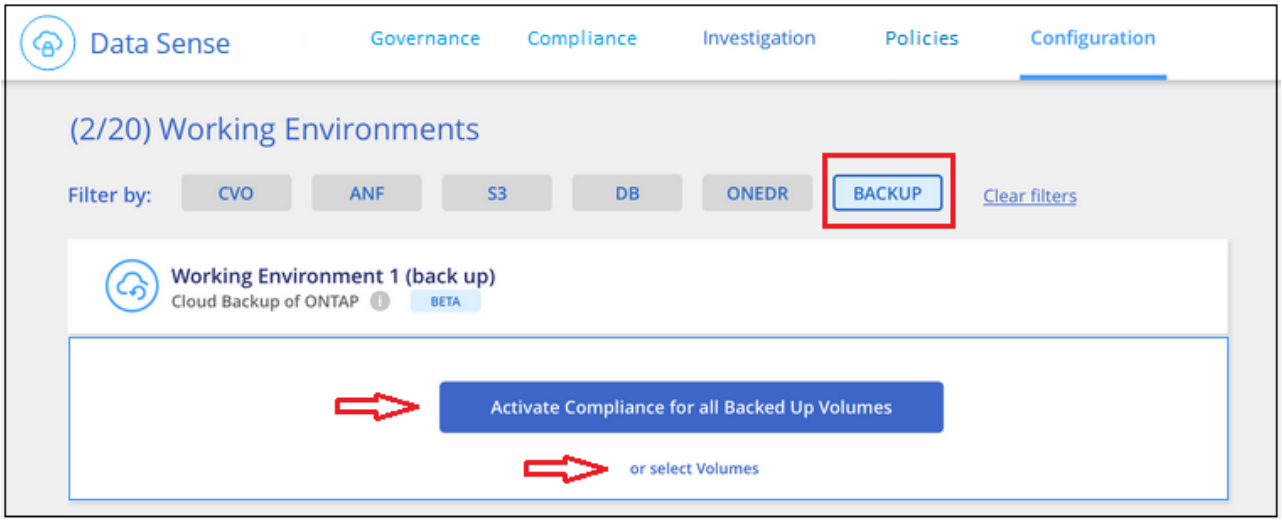
5. Select the volumes that you want to back up.
   - To back up all volumes, check the box in the title row (Volume Name).
   - To back up individual volumes, check the box for each volume (Volume_1).

6. Click **Activate Backup** and Cloud Backup starts taking the initial backups of your volumes.

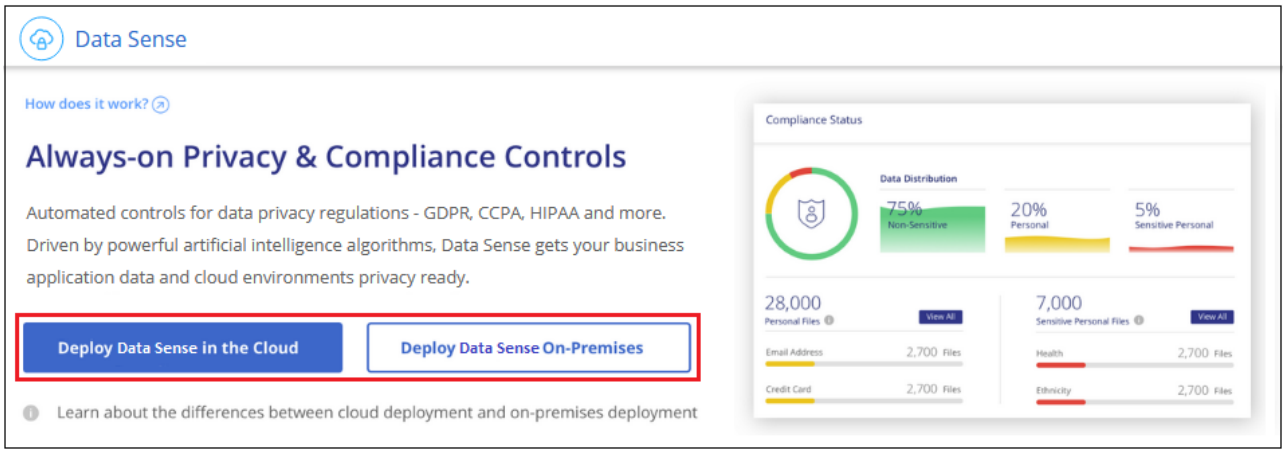
   When creating backup files in AWS or Azure, you are prompted whether you want to run compliance scans on the backed up volumes. Cloud Data Sense scans are free when you run them on the backed up volumes (except for the cost of the deployed Cloud Data Sense instance).



7. Click **Go to Compliance** to activate compliance scans on the volumes. (If you choose **Close** and not to scan these backed up volumes, you can always enable this functionality later from Cloud Data Sense.)
    - If an instance of Cloud Data Sense is already deployed in your environment, you are directed to the Configuration page to select the volumes you want to scan in each on-premises working environment that has backups. See how to choose the volumes.



    - If Cloud Data Sense has not been deployed, you are directed to the Compliance page where you can choose to deploy Compliance in the cloud or in your premises. We strongly recommend deploying it in the cloud. Go here for installation requirements and instructions.



After you have deployed Compliance you can choose the volumes you want to scan as described above.

Result

Cloud Backup backs up your volumes from the on-premises ONTAP system, and optionally, Cloud Data Sense runs compliance scans on the backed up volumes.

What's next?

You can start and stop backups for volumes or change the backup schedule and you can restore entire volumes or individual files from a backup file.

You can also view the results of the compliance scans and review other features of Cloud Data Sense that can help you understand data context and identify sensitive data in your organization.

The scan results are not available immediately because Cloud Backup has to finish creating the backups before Cloud Data Sense can start compliance scans.