



Activate scanning on your data sources

Cloud Manager

NetApp
October 11, 2021

This PDF was generated from https://docs.netapp.com/us-en/occm/task_getting_started_compliance.html on October 11, 2021. Always check docs.netapp.com for the latest.

Table of Contents

- Activate scanning on your data sources. 1
 - Getting started with Cloud Data Sense for Cloud Volumes ONTAP and on-premises ONTAP 1
 - Getting started with Cloud Data Sense for Azure NetApp Files 8
 - Get started with Cloud Data Sense for Amazon FSx for ONTAP 12
 - Getting started with Cloud Data Sense for Amazon S3 16
- Scanning database schemas 22
- Scanning OneDrive accounts. 26
- Scanning file shares 29
- Scanning object storage that uses S3 protocol 33

Activate scanning on your data sources

Getting started with Cloud Data Sense for Cloud Volumes ONTAP and on-premises ONTAP

Complete a few steps to get started with Cloud Data Sense for Cloud Volumes ONTAP and on-premises ONTAP systems.

Quick start

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



Discover the data sources that contain the data you want to scan

Before you can scan volumes, you must add the systems as working environments in Cloud Manager:

- For Cloud Volumes ONTAP systems, these working environments should already be available in Cloud Manager
- For on-premises ONTAP systems, [Cloud Manager must discover the ONTAP clusters](#)



Deploy the Cloud Data Sense instance

[Deploy Cloud Data Sense in Cloud Manager](#) if there isn't already an instance deployed.



Enable Cloud Data Sense and select the volumes to scan

Click **Data Sense**, select the **Configuration** tab, and activate compliance scans for volumes in specific working environments.



Ensure access to volumes

Now that Cloud Data Sense is enabled, ensure that it can access all volumes.

- The Cloud Data Sense instance needs a network connection to each Cloud Volumes ONTAP subnet or on-prem ONTAP system.
- Security groups for Cloud Volumes ONTAP must allow inbound connections from the Data Sense instance.
- Make sure these ports are open to the Data Sense instance:
 - For NFS – ports 111 and 2049.
 - For CIFS – ports 139 and 445.
- NFS volume export policies must allow access from the Data Sense instance.
- Data Sense needs Active Directory credentials to scan CIFS volumes.

Click **Compliance > Configuration > Edit CIFS Credentials** and provide the credentials.



Manage the volumes you want to scan

Select or deselect the volumes that you want to scan and Cloud Data Sense will start or stop scanning them.

Discovering the data sources that you want to scan

If the data sources you want to scan are not already in your Cloud Manager environment, you can add them to the canvas at this time.

Your Cloud Volumes ONTAP systems should already be available in the Canvas in Cloud Manager. For on-premises ONTAP systems you need to have [Cloud Manager discover these clusters](#).

Deploying the Cloud Data Sense instance

[Deploy Cloud Data Sense](#) if there isn't already an instance deployed.

Cloud Data Sense can be deployed in the cloud or in an on-premises location when scanning Cloud Volumes ONTAP or on-premises ONTAP systems.

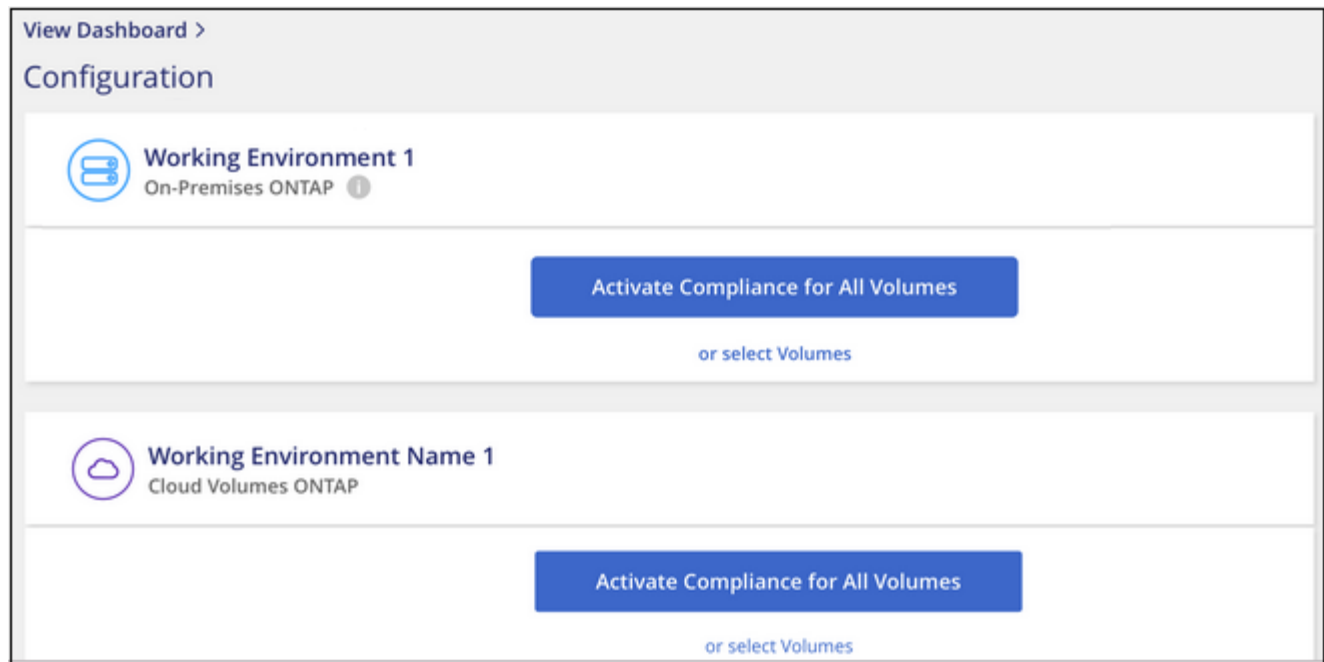
Enabling Cloud Data Sense in your working environments

You can enable Cloud Data Sense on Cloud Volumes ONTAP systems (in AWS or Azure) and on on-premises ONTAP clusters.



Following these steps for on-prem ONTAP systems scans the volumes directly on the on-prem ONTAP system. If you are already creating backup files from those on-prem systems using [Cloud Backup](#), you can run compliance scans on the backup files in the cloud instead. Go to [Scanning backup files from on-premises ONTAP systems](#) to scan the volumes by scanning the backup files.

1. At the top of Cloud Manager, click **Data Sense** and then select the **Configuration** tab.



2. To scan all volumes in a working environment, click **Activate Scanning for All Volumes**.

When enabled in this manner, full "mapping and classification" scanning is performed on all volumes.

If you want to enable scanning only for certain volumes, or if you only want to perform "mapping-only" scanning, click **or select Volumes** and then choose the volumes you want to scan.

See [Enabling and disabling compliance scans on volumes](#) for details.

Result

Cloud Data Sense starts scanning the volumes you selected in the working environment. Results will be available in the Compliance dashboard as soon as Cloud Data Sense finishes the initial scans. The time that it takes depends on the amount of data—it could be a few minutes or hours.

Verifying that Cloud Data Sense has access to volumes

Make sure that Cloud Data Sense can access volumes by checking your networking, security groups, and export policies. You'll need to provide Data Sense with CIFS credentials so it can access CIFS volumes.

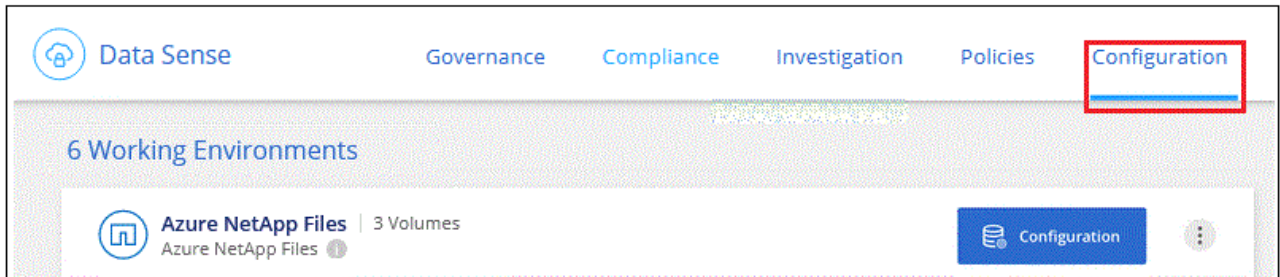
Steps

1. Make sure that there's a network connection between the Cloud Data Sense instance and each network that includes volumes for Cloud Volumes ONTAP or on-prem ONTAP clusters.
2. Ensure that the security group for Cloud Volumes ONTAP allows inbound traffic from the Data Sense instance.

You can either open the security group for traffic from the IP address of the Data Sense instance, or you can open the security group for all traffic from inside the virtual network.

3. Ensure the following ports are open to the Data Sense instance:
 - For NFS – ports 111 and 2049.
 - For CIFS – ports 139 and 445.

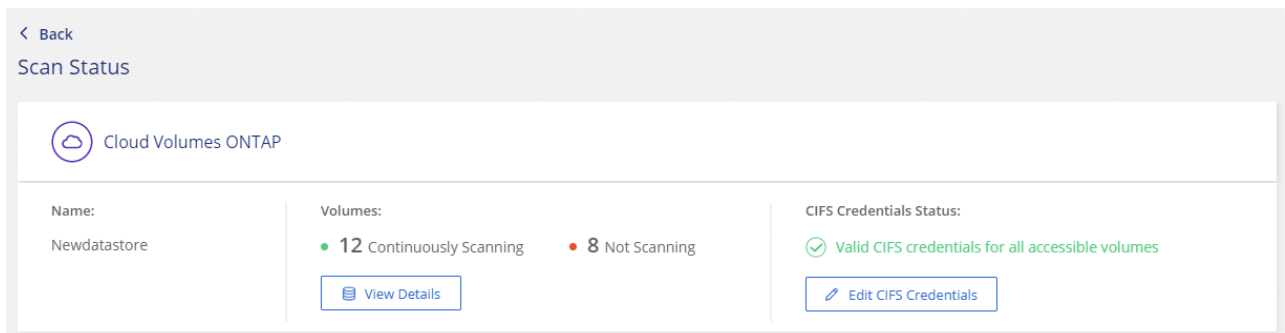
4. Ensure that NFS volume export policies include the IP address of the Data Sense instance so it can access the data on each volume.
5. If you use CIFS, provide Data Sense with Active Directory credentials so it can scan CIFS volumes.
 - a. At the top of Cloud Manager, click **Data Sense**.
 - b. Click the **Configuration** tab.



- c. For each working environment, click **Edit CIFS Credentials** and enter the user name and password that Data Sense needs to access CIFS volumes on the system.

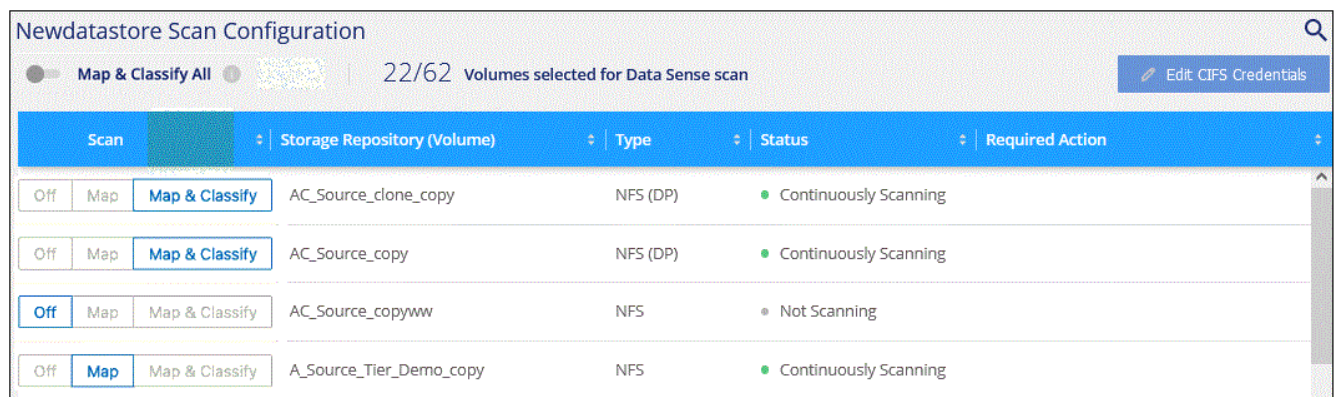
The credentials can be read-only, but providing admin credentials ensures that Data Sense can read any data that requires elevated permissions. The credentials are stored on the Cloud Data Sense instance.

After you enter the credentials, you should see a message that all CIFS volumes were authenticated successfully.



6. On the *Configuration* page, click **View Details** to review the status for each CIFS and NFS volume and correct any errors.

For example, the following image shows three volumes; one of which Cloud Data Sense can't scan due to network connectivity issues between the Data Sense instance and the volume.



Enabling and disabling compliance scans on volumes

You can stop or start mapping-only scans, or mapping and classification scans, in a working environment at any time from the Configuration page. We recommend that you scan all volumes.

cognitoWE Scan Configuration

Map & Classify All

4/79 Volumes selected for Data Sense scan

Edit CIFS Credentials

Scan	Storage Repository (Volume)	Type	Status	Required Action
<div>Off</div> <div>Map</div> <div>Map & Classify</div>	AdiNFSVol_copy	NFS	No Access	Access to the NFS volume was denied. Make sure tha...
<div>Off</div> <div>Map</div> <div>Map & Classify</div>	AdiProtest2501	NFS	Continuously Scanning	
<div>Off</div> <div>Map</div> <div>Map & Classify</div>	AlexTest	NFS	No Access	Access to the NFS volume was denied. Make sure tha...
<div>Off</div> <div>Map</div> <div>Map & Classify</div>	AlexTestSecond	NFS	Not Scanning	
<div>Off</div> <div>Map</div> <div>Map & Classify</div>	MoreDataNeed1000	NFS	Continuously Scanning	

To:	Do this:
Enable mapping-only scans on a volume	Click Map
Enable full scanning on a volume	Click Map & Classify
Enable full scanning on all volumes	Move the Map & Classify All slider to the right
Disable scanning on a volume	Click Off
Disable scanning on all volumes	Move the Map & Classify All slider to the left



New volumes added to the working environment are automatically scanned only when the **Map & Classify All** setting is enabled. When this setting is disabled, you'll need to activate mapping and/or full scanning on each new volume you create in the working environment.

Scanning backup files from on-premises ONTAP systems

If you don't want Cloud Data Sense to scan volumes directly on your on-prem ONTAP systems, a Beta feature released in January 2021 allows you to run compliance scans on backup files created from your on-prem ONTAP volumes. So if you're already creating backup files using [Cloud Backup](#), you can use this feature to run compliance scans on those backup files.

The Compliance scans you run on backup files are **free** - no Cloud Data Sense subscription or license is needed.

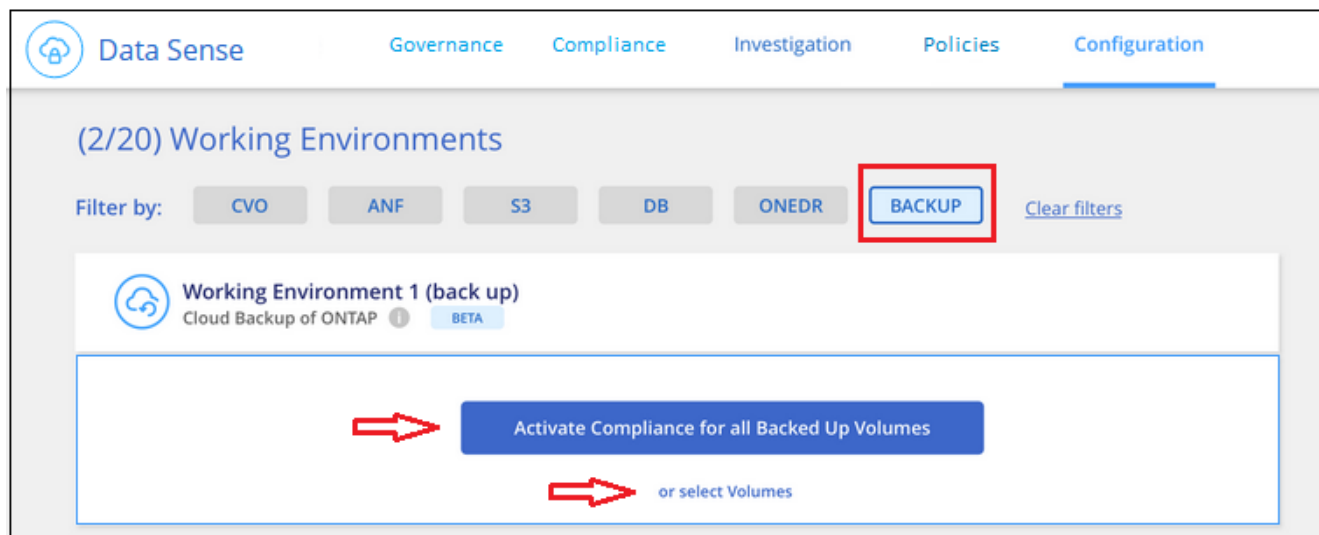
Note: When Data Sense scans backup files it uses permissions granted through the Cloud Restore instance to access the backup files. Typically the Restore instance powers down when not actively restoring files, but it remains **On** when scanning backup files. See [more information about the Restore instance](#).

Steps

If you want to scan the backup files from on-prem ONTAP systems:

1. At the top of Cloud Manager, click **Data Sense** and then select the **Configuration** tab.
2. From the list of working environments, click the **BACKUP** button from the list of filters.

All the on-premises ONTAP working environments that have backup files are listed. If you don't have any backup files from an on-prem system, then the working environment is not shown.



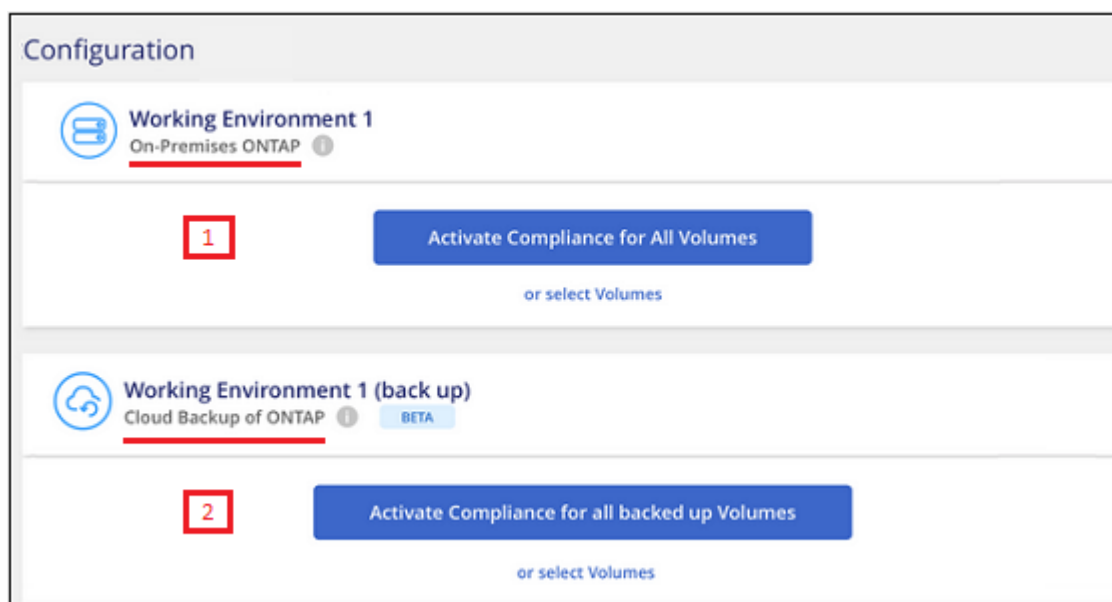
3. To scan all backed up volumes in a working environment, click **Activate Compliance for all backed up Volumes**.

To scan only certain backed up volumes in a working environment, click **or select Volumes** and then choose the backup files (volumes) that you want to scan.

See [Enabling and disabling compliance scans on volumes](#) for details.

Scanning on-prem volumes versus backups of those volumes

When you view the entire list of working environments you will see two listings for each on-prem cluster if they have backed up files.



The first item is the on-prem cluster and the actual volumes.

The second item is the backup files of those volumes from that same on-prem cluster.

Choose the first option to scan the volumes on the on-prem system. Choose the second option to scan the backup files from those volumes. Do not scan both on-prem volumes and backup files of the same cluster.

Scanning data protection volumes

By default, data protection (DP) volumes are not scanned because they are not exposed externally and Cloud Data Sense cannot access them. These are the destination volumes for SnapMirror operations from an on-premises ONTAP system or from a Cloud Volumes ONTAP system.

Initially, the volume list identifies these volumes as *Type DP* with the *Status Not Scanning* and the *Required Action Enable Access to DP volumes*.

The screenshot shows the 'Working Environment Name' Configuration page. At the top, there's a search bar and a button 'Enable Access to DP Volumes' which is highlighted with a red box. Below this is a table with columns: Scan, Storage Repository (Volume), Type, Status, and Required Action. The table lists three volumes: VolumeName1 (Type DP, Status Not Scanning), VolumeName2 (Type NFS, Status Continuously Scanning), and VolumeName3 (Type CIFS, Status Not Scanning). Each volume has a 'Scan' column with a toggle switch and buttons for 'Map' and 'Map & Classify'.

Steps

If you want to scan these data protection volumes:

1. Click **Enable Access to DP volumes** at the top of the page.
2. Review the confirmation message and click **Enable Access to DP volumes** again.
 - Volumes that were initially created as NFS volumes in the source ONTAP system are enabled.
 - Volumes that were initially created as CIFS volumes in the source ONTAP system require that you enter CIFS credentials to scan those DP volumes. If you already entered Active Directory credentials so that Cloud Data Sense can scan CIFS volumes you can use those credentials, or you can specify a different set of Admin credentials.

The screenshot shows the 'Provide Active Directory Credentials' dialog box. The 'Use existing CIFS Scanning Credentials (user1@domain2)' radio button is selected and highlighted with a red box. Below this are fields for 'Active Directory Domain' and 'DNS IP Address'. At the bottom, there are buttons for 'Enable Access to DP Volumes' and 'Cancel'.

The screenshot shows the 'Provide Active Directory Credentials' dialog box. The 'Use Custom Credentials' radio button is selected and highlighted with a red box. Below this are fields for 'Username', 'Password', 'Active Directory Domain', and 'DNS IP Address'. At the bottom, there are buttons for 'Enable Access to DP Volumes' and 'Cancel'.

3. Activate each DP volume that you want to scan [the same way you enabled other volumes](#), or use the **Activate Compliance for all Volumes** control to enable all volumes, including all DP volumes.

Result

Once enabled, Cloud Data Sense creates an NFS share from each DP volume that was activated for scanning. The share export policies only allow access from the Data Sense instance.

Note: If you had no CIFS data protection volumes when you initially enabled access to DP volumes, and later add some, the button **Enable Access to CIFS DP** appears at the top of the Configuration page. Click this button and add CIFS credentials to enable access to these CIFS DP volumes.



Active Directory credentials are only registered in the storage VM of the first CIFS DP volume, so all DP volumes on that SVM will be scanned. Any volumes that reside on other SVMs will not have the Active Directory credentials registered, so those DP volumes won't be scanned.

Getting started with Cloud Data Sense for Azure NetApp Files

Complete a few steps to get started with Cloud Data Sense for Azure NetApp Files.

Quick start

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



Discover the data sources that contain the data you want to scan

Before you can scan Azure NetApp Files volumes, [Cloud Manager must be set up to discover the configuration](#).



Deploy the Cloud Data Sense instance

[Deploy Cloud Data Sense in Cloud Manager](#) if there isn't already an instance deployed.



Enable Cloud Data Sense and select the volumes to scan

Click **Compliance**, select the **Configuration** tab, and activate compliance scans for volumes in specific working environments.



Ensure access to volumes

Now that Cloud Data Sense is enabled, ensure that it can access all volumes.

- The Cloud Data Sense instance needs a network connection to each Azure NetApp Files subnet.
- Make sure these ports are open to the Data Sense instance:
 - For NFS – ports 111 and 2049.
 - For CIFS – ports 139 and 445.
- NFS volume export policies must allow access from the Data Sense instance.

- Data Sense needs Active Directory credentials to scan CIFS volumes.

Click **Compliance > Configuration > Edit CIFS Credentials** and provide the credentials.



Manage the volumes you want to scan

Select or deselect the volumes that you want to scan and Cloud Data Sense will start or stop scanning them.

Discovering the data sources that you want to scan

If the data sources you want to scan are not already in your Cloud Manager environment, you can add them to the canvas at this time.

For Azure NetApp Files, [Cloud Manager must be set up to discover the configuration](#).

Deploying the Cloud Data Sense instance

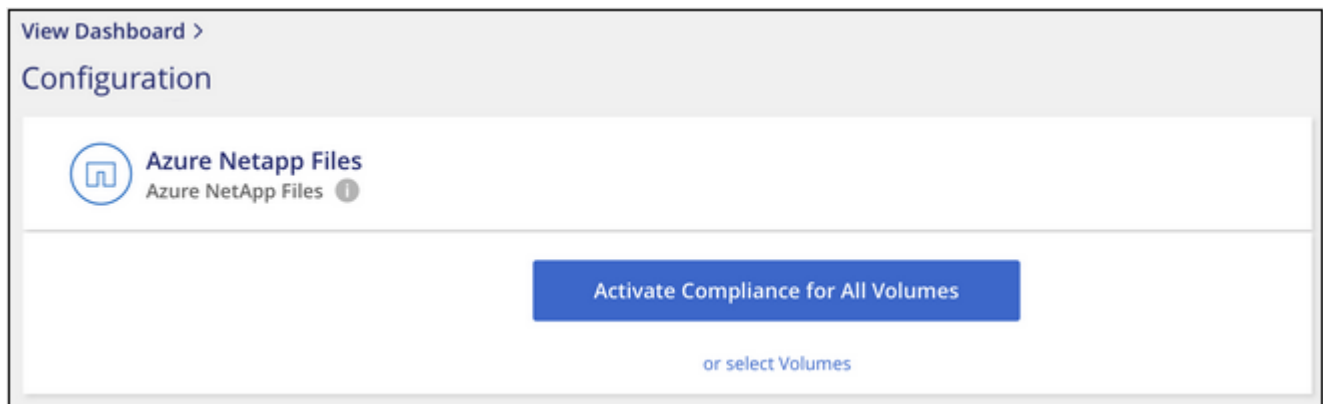
[Deploy Cloud Data Sense](#) if there isn't already an instance deployed.

Data Sense must be deployed in the cloud when scanning Azure NetApp Files volumes, and it must be deployed in the same region as the volumes you wish to scan.

Enabling Cloud Data Sense in your working environments

You can enable Cloud Data Sense on your Azure NetApp Files volumes.

1. At the top of Cloud Manager, click **Data Sense** and then select the **Configuration** tab.



2. To scan all volumes in a working environment, click **Activate Scanning for All Volumes**.

When enabled in this manner, full "mapping and classification" scanning is performed on all volumes.

If you want to enable scanning only for certain volumes, or if you only want to perform "mapping-only" scanning, click **or select Volumes** and then choose the volumes you want to scan.

See [Enabling and disabling compliance scans on volumes](#) for details.

Result

Cloud Data Sense starts scanning the volumes you selected in the working environment. Results will be available in the Compliance dashboard as soon as Cloud Data Sense finishes the initial scans. The time that it takes depends on the amount of data—it could be a few minutes or hours.

Verifying that Cloud Data Sense has access to volumes

Make sure that Cloud Data Sense can access volumes by checking your networking, security groups, and export policies. You'll need to provide Data Sense with CIFS credentials so it can access CIFS volumes.

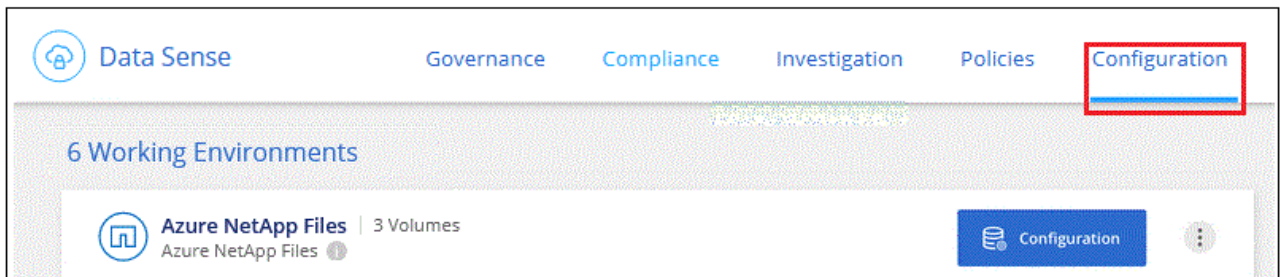
Steps

1. Make sure that there's a network connection between the Cloud Data Sense instance and each network that includes volumes for Azure NetApp Files.



For Azure NetApp Files, Cloud Data Sense can only scan volumes that are in the same region as Cloud Manager.

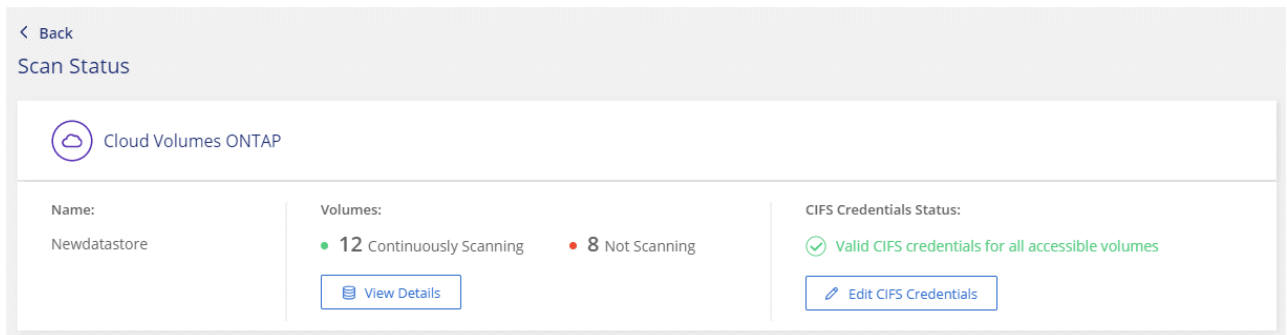
2. Ensure the following ports are open to the Data Sense instance:
 - For NFS – ports 111 and 2049.
 - For CIFS – ports 139 and 445.
3. Ensure that NFS volume export policies include the IP address of the Data Sense instance so it can access the data on each volume.
4. If you use CIFS, provide Data Sense with Active Directory credentials so it can scan CIFS volumes.
 - a. At the top of Cloud Manager, click **Data Sense**.
 - b. Click the **Configuration** tab.



- c. For each working environment, click **Edit CIFS Credentials** and enter the user name and password that Data Sense needs to access CIFS volumes on the system.

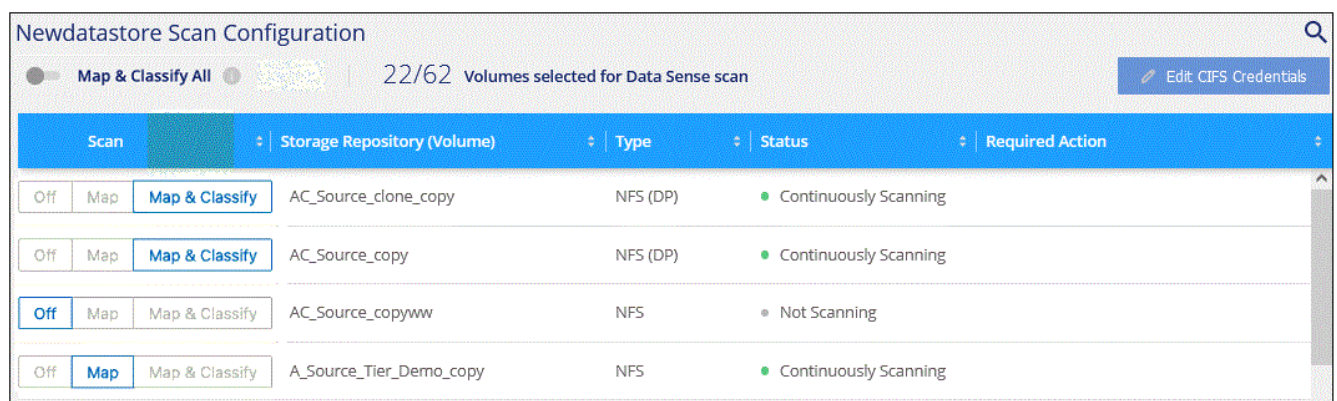
The credentials can be read-only, but providing admin credentials ensures that Data Sense can read any data that requires elevated permissions. The credentials are stored on the Cloud Data Sense instance.

After you enter the credentials, you should see a message that all CIFS volumes were authenticated successfully.



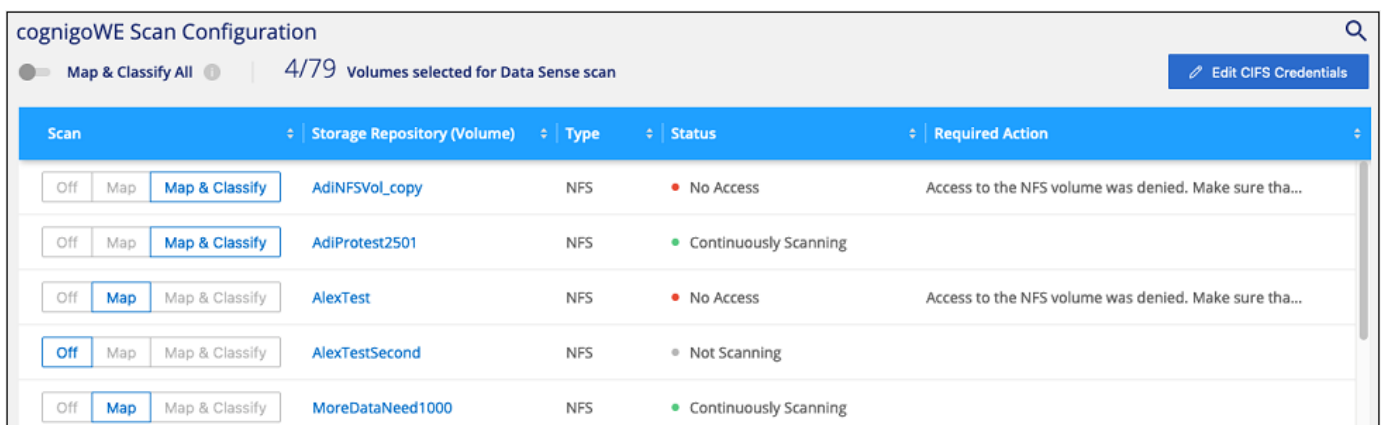
- On the *Configuration* page, click **View Details** to review the status for each CIFS and NFS volume and correct any errors.

For example, the following image shows three volumes; one of which Cloud Data Sense can't scan due to network connectivity issues between the Data Sense instance and the volume.



Enabling and disabling compliance scans on volumes

You can stop or start mapping scans, or mapping and classification scans, in a working environment at any time from the Configuration page. We recommend that you scan all volumes.



To:	Do this:
Enable mapping-only scans on a volume	Click Map
Enable full scanning on a volume	Click Map & Classify
Enable full scanning on all volumes	Move the Map & Classify All slider to the right

To:	Do this:
Disable scanning on a volume	Click Off
Disable scanning on all volumes	Move the Map & Classify All slider to the left



New volumes added to the working environment are automatically scanned only when the **Activate Compliance for all Volumes** setting is enabled. When this setting is disabled, you'll need to activate scanning on each new volume you create in the working environment.

Get started with Cloud Data Sense for Amazon FSx for ONTAP

Complete a few steps to get started with Cloud Data Sense for FSx for ONTAP.

Before you begin

- You need an active Connector in AWS to deploy and manage Data Sense.
- The security group you selected when creating the working environment must allow traffic from the Cloud Data Sense instance. You can find the associated security group using the ENI connected to the FSx for ONTAP file system and edit it using the AWS Management Console.

[AWS security groups for Linux instances](#)

[AWS security groups for Windows instances](#)

[AWS elastic network interfaces \(ENI\)](#)

Quick start

Get started quickly by following these steps or scroll down for full details.



Discover the data sources that contain the data you want to scan

Before you can scan FSx for ONTAP volumes, [you must have an FSx working environment with volumes configured](#).



Deploy the Cloud Data Sense instance

[Deploy Cloud Data Sense in Cloud Manager](#) if there isn't already an instance deployed.



Enable Cloud Data Sense and select the volumes to scan

Click **Data Sense**, select the **Configuration** tab, and activate compliance scans for volumes in specific working environments.



Ensure access to volumes

Now that Cloud Data Sense is enabled, ensure that it can access all volumes.

- The Cloud Data Sense instance needs a network connection to each FSx for ONTAP subnet.
- Make sure the following ports are open to the Data Sense instance.
 - For NFS – ports 111 and 2049.
 - For CIFS – ports 139 and 445.
- NFS volume export policies must allow access from the Data Sense instance.
- Data Sense needs Active Directory credentials to scan CIFS volumes.

Click **Compliance > Configuration > Edit CIFS Credentials** and provide the credentials.



Manage the volumes you want to scan

Select or deselect the volumes you want to scan and Cloud Data Sense will start or stop scanning them.

Discovering the data sources that you want to scan

If the data sources you want to scan are not already in your Cloud Manager environment, you can add them to the canvas at this time.

For FSx for ONTAP, [Cloud Manager must be set up to discover the configuration](#).

Deploying the Cloud Data Sense instance

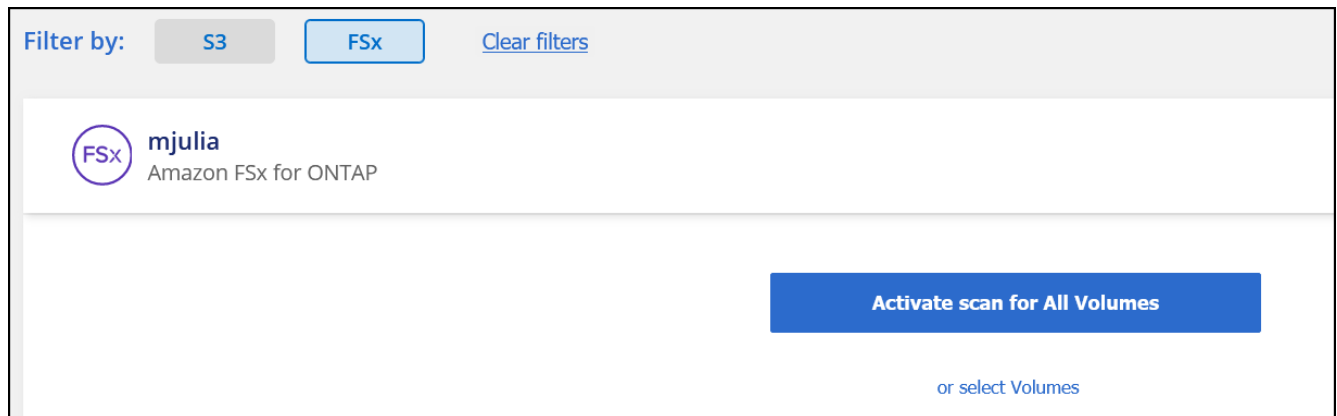
[Deploy Cloud Data Sense](#) if there isn't already an instance deployed.

Data Sense should be deployed in the same AWS network as the Connector for AWS and the FSx volumes you wish to scan.

Enabling Cloud Data Sense in your working environments

You can enable Cloud Data Sense for FSx for ONTAP volumes.

1. At the top of Cloud Manager, click **Data Sense** and then select the **Configuration** tab.



2. To scan all volumes in a working environment, click **Activate Scanning for All Volumes**.

When enabled in this manner, full "mapping and classification" scanning is performed on all volumes.

If you want to enable scanning only for certain volumes, or if you only want to perform "mapping-only" scanning, click **or select Volumes** and then choose the volumes you want to scan.

See [Enabling and disabling compliance scans on volumes](#) for details.

Result

Cloud Data Sense starts scanning the volumes you selected in the working environment. Results will be available in the Compliance dashboard as soon as Cloud Data Sense finishes the initial scans. The time that it takes depends on the amount of data—it could be a few minutes or hours.

Verifying that Cloud Data Sense has access to volumes

Make sure Cloud Data Sense can access volumes by checking your networking, security groups, and export policies.

You'll need to provide Data Sense with CIFS credentials so it can access CIFS volumes.

Steps

1. On the *Configuration* page, click **View Details** to review the status and correct any errors.

For example, the following image shows a volume Cloud Data Sense can't scan due to network connectivity issues between the Data Sense instance and the volume.

Scan	Storage Repository (Volume)	Type	Status	Required Action
<input type="button" value="Off"/> <input type="button" value="Map"/> <input type="button" value="Map & Classify"/>	jrmclone	NFS	● No Access	Check network connectivity between the Data Sense ...

2. Make sure there's a network connection between the Cloud Data Sense instance and each network that includes volumes for FSx for ONTAP.



For FSx for ONTAP, Cloud Data Sense can scan volumes only in the same region as Cloud Manager.

3. Ensure the following ports are open to the Data Sense instance.
 - For NFS – ports 111 and 2049.

- For CIFS – ports 139 and 445.
- 4. Ensure NFS volume export policies include the IP address of the Data Sense instance so it can access the data on each volume.
- 5. If you use CIFS, provide Data Sense with Active Directory credentials so it can scan CIFS volumes.
 - a. At the top of Cloud Manager, click **Data Sense**.
 - b. Click the **Configuration** tab.
 - c. For each working environment, click **Edit CIFS Credentials** and enter the user name and password that Data Sense needs to access CIFS volumes on the system.

The credentials can be read-only, but providing admin credentials ensures that Data Sense can read any data that requires elevated permissions. The credentials are stored on the Cloud Data Sense instance.

After you enter the credentials, you should see a message that all CIFS volumes were authenticated successfully.

Enabling and disabling compliance scans on volumes

You can stop or start mapping scans, or mapping and classification scans, in a working environment at any time from the Configuration page. We recommend that you scan all volumes.

The screenshot shows the 'cognigoWE Scan Configuration' interface. At the top, there's a slider for 'Map & Classify All' and a count '4/79 Volumes selected for Data Sense scan'. A button 'Edit CIFS Credentials' is visible. Below is a table with the following data:

Scan	Storage Repository (Volume)	Type	Status	Required Action
Off Map Map & Classify	AdINFSVol_copy	NFS	No Access	Access to the NFS volume was denied. Make sure tha...
Off Map Map & Classify	AdiProtest2501	NFS	Continuously Scanning	
Off Map Map & Classify	AlexTest	NFS	No Access	Access to the NFS volume was denied. Make sure tha...
Off Map Map & Classify	AlexTestSecond	NFS	Not Scanning	
Off Map Map & Classify	MoreDataNeed1000	NFS	Continuously Scanning	

To:	Do this:
Enable mapping-only scans on a volume	Click Map
Enable full scanning on a volume	Click Map & Classify
Enable full scanning on all volumes	Move the Map & Classify All slider to the right
Disable scanning on a volume	Click Off
Disable scanning on all volumes	Move the Map & Classify All slider to the left



New volumes added to the working environment are automatically scanned only when the **Activate Compliance for all Volumes** setting is enabled. When this setting is disabled, you'll need to activate scanning on each new volume you create in the working environment.

Getting started with Cloud Data Sense for Amazon S3

Cloud Data Sense can scan your Amazon S3 buckets to identify the personal and sensitive data that resides in S3 object storage. Cloud Data Sense can scan any bucket in the account, regardless if it was created for a NetApp solution.

Quick start

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



Set up the S3 requirements in your cloud environment

Ensure that your cloud environment can meet the requirements for Cloud Data Sense, including preparing an IAM role and setting up connectivity from Data Sense to S3. [See the complete list.](#)



Deploy the Cloud Data Sense instance

[Deploy Cloud Data Sense](#) if there isn't already an instance deployed.



Activate Data Sense on your S3 working environment

Select the Amazon S3 working environment, click **Enable**, and select an IAM role that includes the required permissions.



Select the buckets to scan

Select the buckets that you'd like to scan and Cloud Data Sense will start scanning them.

Reviewing S3 prerequisites

The following requirements are specific to scanning S3 buckets.

Set up an IAM role for the Cloud Data Sense instance

Cloud Data Sense needs permissions to connect to the S3 buckets in your account and to scan them. Set up an IAM role that includes the permissions listed below. Cloud Manager prompts you to select an IAM role when you enable Data Sense on the Amazon S3 working environment.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "s3:Get*",
        "s3:List*",
        "s3:PutObject",
      ],
      "Resource": "*"
    },
    {
      "Effect": "Allow",
      "Action": [
        "iam:GetPolicyVersion",
        "iam:GetPolicy",
        "iam:ListAttachedRolePolicies"
      ],
      "Resource": [
        "arn:aws:iam::*:policy/*",
        "arn:aws:iam::*:role/*"
      ]
    }
  ]
}
```

Provide connectivity from Cloud Data Sense to Amazon S3

Cloud Data Sense needs a connection to Amazon S3. The best way to provide that connection is through a VPC Endpoint to the S3 service. For instructions, see [AWS Documentation: Creating a Gateway Endpoint](#).

When you create the VPC Endpoint, be sure to select the region, VPC, and route table that corresponds to the Cloud Data Sense instance. You must also modify the security group to add an outbound HTTPS rule that enables traffic to the S3 endpoint. Otherwise, Data Sense can't connect to the S3 service.

If you experience any issues, see [AWS Support Knowledge Center: Why can't I connect to an S3 bucket using a gateway VPC endpoint?](#)

An alternative is to provide the connection by using a NAT Gateway.



You can't use a proxy to get to S3 over the internet.

Deploying the Cloud Data Sense instance

[Deploy Cloud Data Sense in Cloud Manager](#) if there isn't already an instance deployed.

You need to deploy the instance in an AWS Connector so that Cloud Manager automatically discovers the S3

buckets in this AWS account and displays them in an Amazon S3 working environment.

Note: Deploying Cloud Data Sense in an on-premises location is not currently supported when scanning S3 buckets.

Activating Data Sense on your S3 working environment

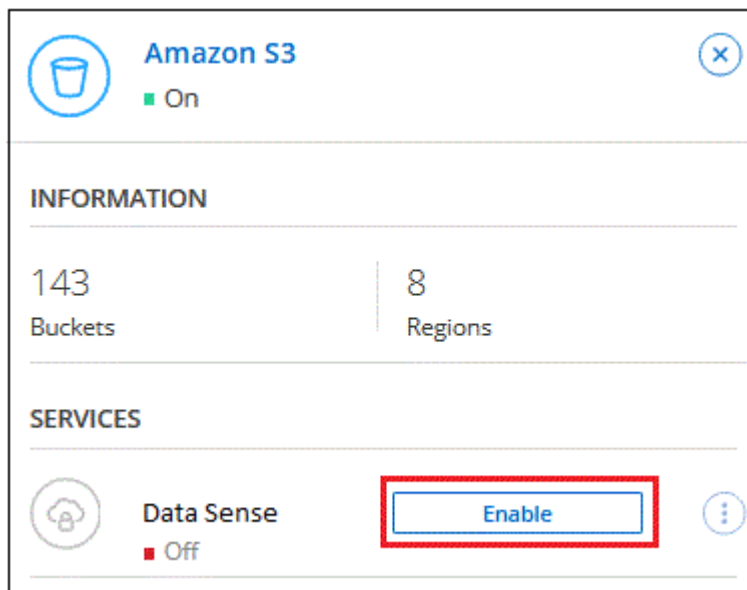
Enable Cloud Data Sense on Amazon S3 after you verify the prerequisites.

Steps

1. At the top of Cloud Manager, click **Canvas**.
2. Select the Amazon S3 working environment.



3. In the Data Sense pane on the right, click **Enable**.



4. When prompted, assign an IAM role to the Cloud Data Sense instance that has [the required permissions](#).

Assign an AWS IAM Role for Cloud Data Sense

To enable **Cloud Data Sense** on Amazon S3 buckets, select an existing IAM Role. Make sure that your AWS IAM Role has the permission defined in the [Policy Requirements](#).

Select IAM Role

occm

▼

VPC Endpoint for Amazon S3 Required

A VPC endpoint to the Amazon S3 service is required so **Data Sense** can securely scan the data.

Alternatively, ensure that the **Data Sense** instance has direct access to the internet via a NAT Gateway or Internet Gateway.

Free for the 1st TB


Over 1 TB you pay only for what you use. [Learn more about pricing.](#)

Enable

Cancel

5. Click **Enable**.



You can also enable compliance scans for a working environment from the Configuration page by clicking the  button and selecting **Activate Data Sense**.

Result

Cloud Manager assigns the IAM role to the instance.

Enabling and disabling compliance scans on S3 buckets

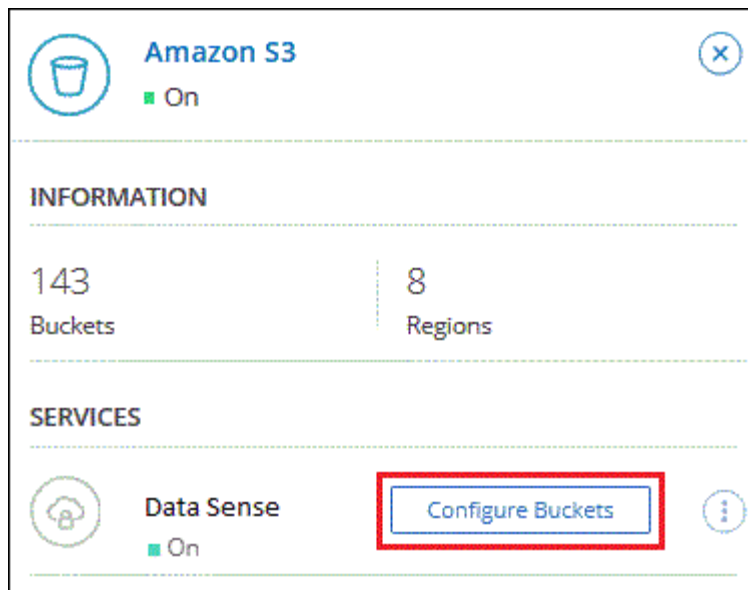
After Cloud Manager enables Cloud Data Sense on Amazon S3, the next step is to configure the buckets that you want to scan.

When Cloud Manager is running in the AWS account that has the S3 buckets you want to scan, it discovers those buckets and displays them in an Amazon S3 working environment.

Cloud Data Sense can also [scan S3 buckets that are in different AWS accounts](#).

Steps

1. Select the Amazon S3 working environment.
2. In the pane on the right, click **Configure Buckets**.



3. Enable mapping-only scans, or mapping and classification scans, on your buckets.

Amazon S3 Configuration			
15/28 Buckets in Scan Scope.			
Scan	Bucket Name	Status	Required Action
Off Map Map & Classify	BucketName1	● Not Scanning	Add Credentials
Off Map Map & Classify	BucketName2	● Continuously Scanning	
Off Map Map & Classify	BucketName3	● Not Scanning	

To:	Do this:
Enable mapping-only scans on a bucket	Click Map
Enable full scans on a bucket	Click Map & Classify
Disable scanning on a bucket	Click Off

Result

Cloud Data Sense starts scanning the S3 buckets that you enabled. If there are any errors, they'll appear in the Status column, alongside the required action to fix the error.

Scanning buckets from additional AWS accounts

You can scan S3 buckets that are under a different AWS account by assigning a role from that account to access the existing Cloud Data Sense instance.

Steps

1. Go to the target AWS account where you want to scan S3 buckets and create an IAM role by selecting **Another AWS account**.

Create role





1

2

3

4

Select type of trusted entity

 AWS service EC2, Lambda and others	 Another AWS account Belonging to you or 3rd party	 Web identity Cognito or any OpenID provider	 SAML 2.0 federation Your corporate directory
--	---	---	--


Allows entities in other accounts to perform actions in this account. [Learn more](#)

Specify accounts that can use this role

Account ID*



Options

- ☐ Require external ID (Best practice when a third party will assume this role)
- ☐ Require MFA 

Be sure to do the following:

- Enter the ID of the account where the Cloud Data Sense instance resides.
- Change the **Maximum CLI/API session duration** from 1 hour to 12 hours and save that change.
- Attach the Cloud Data Sense IAM policy. Make sure it has the required permissions.

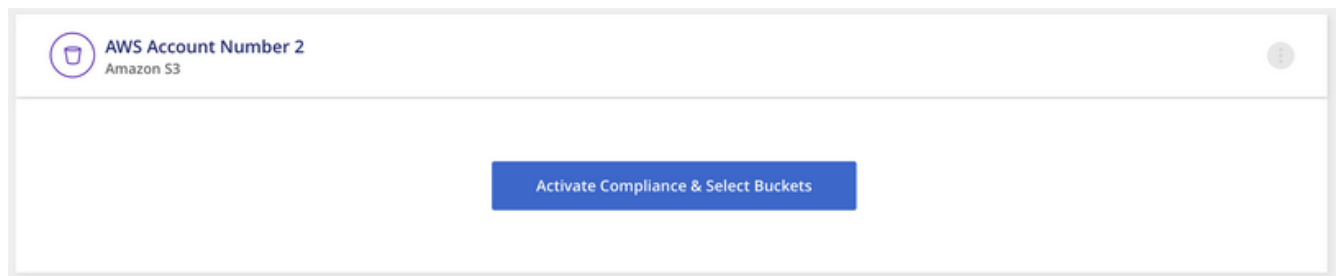
```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "s3:Get*",
        "s3:List*",
        "s3:PutObject",
      ],
      "Resource": "*"
    }
  ]
}
```

2. Go to the source AWS account where the Data Sense instance resides and select the IAM role that is attached to the instance.
 - a. Change the **Maximum CLI/API session duration** from 1 hour to 12 hours and save that change.
 - b. Click **Attach policies** and then click **Create policy**.
 - c. Create a policy that includes the "sts:AssumeRole" action and specify the ARN of the role that you created in the target account.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": "sts:AssumeRole",
      "Resource": "arn:aws:iam::<ADDITIONAL-ACCOUNT-ID>:role/<ADDITIONAL_ROLE_NAME>"
    },
    {
      "Effect": "Allow",
      "Action": [
        "iam:GetPolicyVersion",
        "iam:GetPolicy",
        "iam:ListAttachedRolePolicies"
      ],
      "Resource": [
        "arn:aws:iam::*:policy/*",
        "arn:aws:iam::*:role/*"
      ]
    }
  ]
}
```

The Cloud Data Sense instance profile account now has access to the additional AWS account.

3. Go to the **Amazon S3 Configuration** page and the new AWS account is displayed. Note that it can take a few minutes for Cloud Data Sense to sync the new account's working environment and show this information.



4. Click **Activate Data Sense & Select Buckets** and select the buckets you want to scan.

Result

Cloud Data Sense starts scanning the new S3 buckets that you enabled.

Scanning database schemas

Complete a few steps to start scanning your database schemas with Cloud Data Sense.

Quick start

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



Review database prerequisites

Ensure that your database is supported and that you have the information necessary to connect to the database.



Deploy the Cloud Data Sense instance

Deploy [Cloud Data Sense](#) if there isn't already an instance deployed.



Add the database server

Add the database server that you want to access.



Select the schemas

Select the schemas that you want to scan.

Reviewing prerequisites

Review the following prerequisites to make sure that you have a supported configuration before you enable Cloud Data Sense.

Supported databases

Cloud Data Sense can scan schemas from the following databases:

- Amazon Relational Database Service (Amazon RDS)
- MongoDB
- MySQL
- Oracle
- PostgreSQL
- SAP HANA
- SQL Server (MSSQL)



The statistics gathering feature **must be enabled** in the database.

Database requirements

Any database with connectivity to the Cloud Data Sense instance can be scanned, regardless of where it is hosted. You just need the following information to connect to the database:

- IP Address or host name
- Port
- Service name (only for accessing Oracle databases)
- Credentials that allow read access to the schemas

When choosing a user name and password, it's important to choose one that has full read permissions to all the schemas and tables you want to scan. We recommend that you create a dedicated user for the Cloud Data Sense system with all the required permissions.

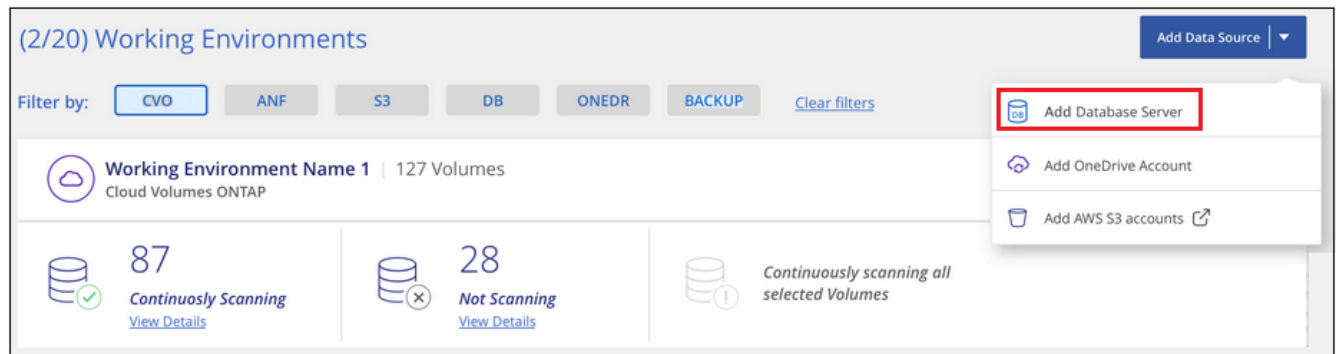
Note: For MongoDB, a read-only Admin role is required.

Adding the database server

You must have [deployed an instance of Cloud Data Sense in Cloud Manager already](#).

Add the database server where the schemas reside.

1. From the Working Environments Configuration page, click **Add Data Source > Add Database Server**.



2. Enter the required information to identify the database server.
 - a. Select the database type.
 - b. Enter the port and the host name or IP address to connect to the database.
 - c. For Oracle databases, enter the Service name.
 - d. Enter the credentials so that Cloud Data Sense can access the server.
 - e. Click **Add DB Server**.

Add DB Server

To activate Compliance on Databases, first add a Database Server. After this step, you'll be able to select which Database Schemas you would like to activate Compliance for.

Database

Database Type Host Name or IP Address

Port Service Name

Credentials

Username Password

The database is added to the list of working environments.

Enabling and disabling compliance scans on database schemas

You can stop or start full scanning of your schemas at any time.



There is no option to select mapping-only scans for database schemas.

1. From the *Configuration* page, click the **Configuration** button for the database you want to configure.

Configuration

Oracle DB 1 | 41 Schemas
Oracle

No Schemas selected for Compliance

7 Not Scanning
[View Details](#)

2. Select the schemas that you want to scan by moving the slider to the right.

'Working Environment Name' Configuration			
28/28 Schemas selected for compliance scan		Edit Credentials	
Scan	Schema Name	Status	Required Action
<input checked="" type="checkbox"/>	DB1 - SchemaName1	Not Scanning	Add Credentials ⓘ
<input checked="" type="checkbox"/>	DB1 - SchemaName2	Continuously Scanning	
<input checked="" type="checkbox"/>	DB1 - SchemaName3	Continuously Scanning	
<input checked="" type="checkbox"/>	DB1 - SchemaName4	Continuously Scanning	

Result

Cloud Data Sense starts scanning the database schemas that you enabled. If there are any errors, they'll appear in the Status column, alongside the required action to fix the error.

Scanning OneDrive accounts

Complete a few steps to start scanning files in your user's OneDrive folders with Cloud Data Sense.

Quick start

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

1

Review OneDrive prerequisites

Ensure that you have the Admin credentials to log into the OneDrive account.

2

Deploy the Cloud Data Sense instance

[Deploy Cloud Data Sense](#) if there isn't already an instance deployed.

3

Add the OneDrive account

Using Admin user credentials, log into the OneDrive account that you want to access so that it is added as a new working environment.

4

Add the users and select the users to scan

Add the list of users from the OneDrive account that you want to scan and select the type of scanning. You can add up to 100 users at time.

Reviewing OneDrive requirements

Review the following prerequisites to make sure that you have a supported configuration before you enable Cloud Data Sense.

- You must have the Admin login credentials for the OneDrive for Business account that provides read access to all user files.
- You will need a line-separated list of the email addresses for all the users whose OneDrive folders you want to scan.

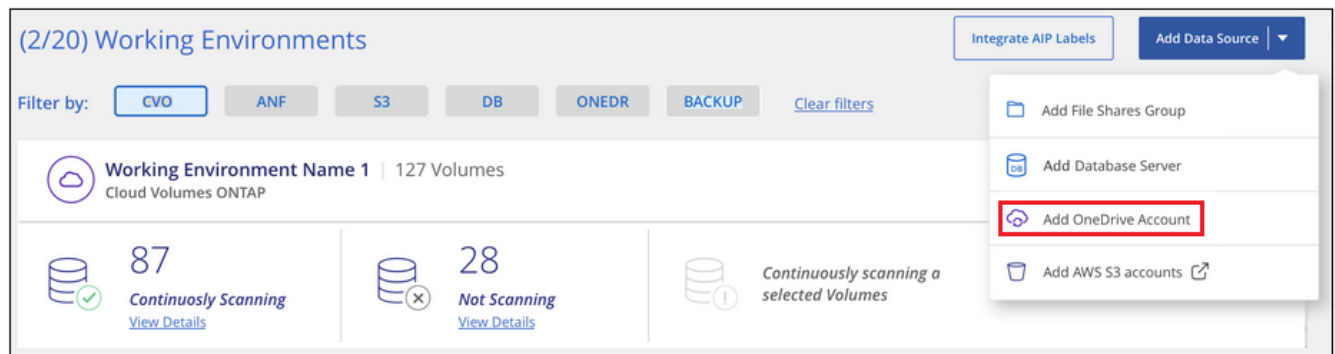
Adding the OneDrive account

You must have [deployed an instance of Cloud Data Sense in Cloud Manager already](#).

Add the OneDrive account where the user files reside.

Steps

1. From the Working Environments Configuration page, click **Add Data Source > Add OneDrive Account**.



2. In the Add a OneDrive account dialog, click **Sign in to OneDrive**.
3. In the Microsoft page that appears, select the OneDrive account and enter the required Admin user and password, then click **Accept** to allow Cloud Data Sense to read data from this account.

The OneDrive account is added to the list of working environments.

Adding OneDrive users to compliance scans

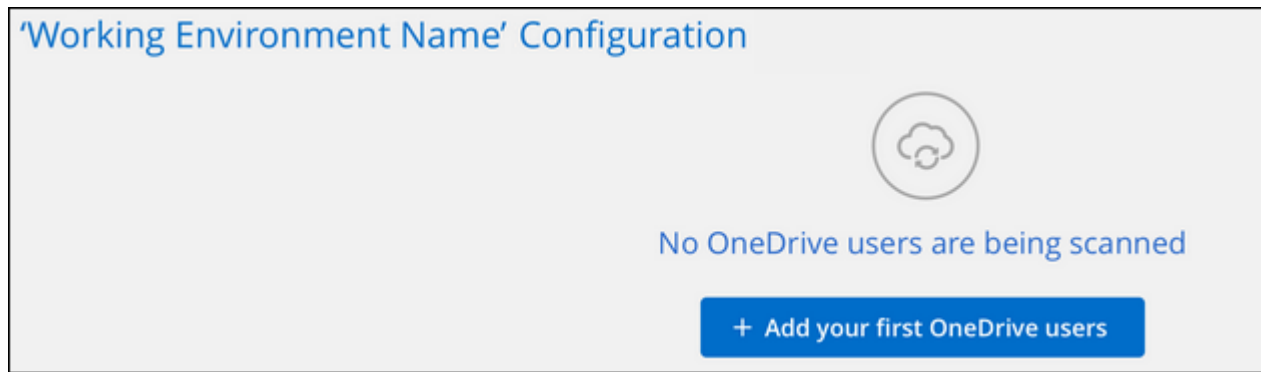
You can add individual OneDrive users, or all of your OneDrive users, so that their files will be scanned by Cloud Data Sense.

Steps

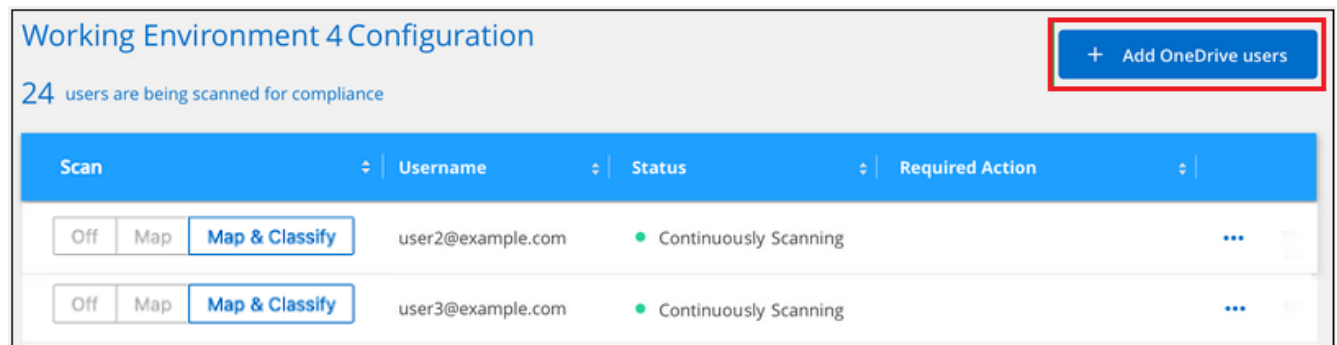
1. From the *Configuration* page, click the **Configuration** button for the OneDrive account.



2. If this is the first time adding users for this OneDrive account, click **Add your first OneDrive users**.



If you are adding additional users from a OneDrive account, click **Add OneDrive users**.



3. Add the email addresses for the users whose files you want to scan - one email address per line (up to 100 maximum per session) - and click **Add Users**.

The screenshot shows a dialog box titled "Add OneDrive users". It contains the instruction: "Provide a list of OneDrive users for Cloud Data Sense to scan their data, line-separated. You can add up to 100 users at a time." Below this, there is a section titled "Type or paste below the OneDrive user accounts to add". Underneath, a text area labeled "User Accounts" contains several lines of placeholder text: "user@example.com". At the bottom of the dialog, there are two buttons: "Add Users" and "Cancel".

A confirmation dialog displays the number of users who were added.

If the dialog lists any users who could not be added, capture this information so that you can resolve the issue. In some cases you can re-add the user with a corrected email address.

4. Enable mapping-only scans, or mapping and classification scans, on user files.

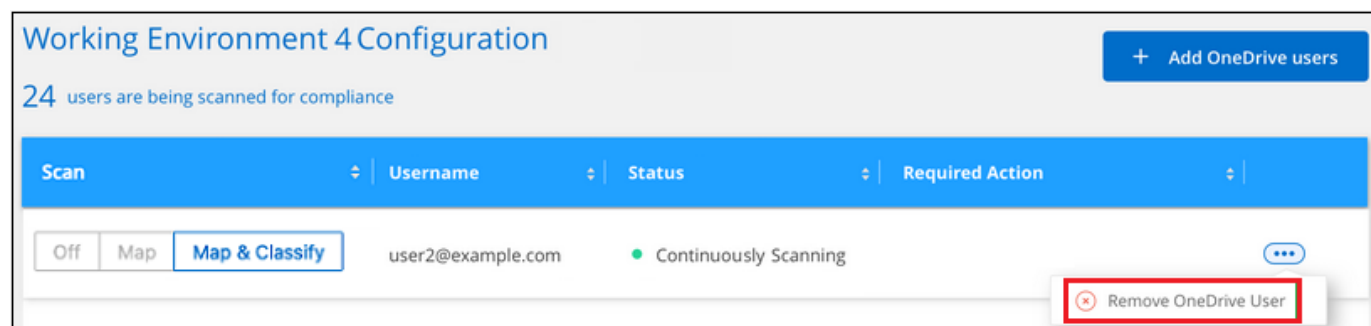
To:	Do this:
Enable mapping-only scans on user files	Click Map
Enable full scans on user files	Click Map & Classify
Disable scanning on user files	Click Off

Result

Cloud Data Sense starts scanning the files for the users you added, and the results are displayed in the Dashboard and in other locations.

Removing a OneDrive user from compliance scans

If users leave the company or if their email address changes, you can remove individual OneDrive users from having their files scanned at any time. Just click **Remove OneDrive User** from the Configuration page.



Scanning file shares

Complete a few steps to start scanning non-NetApp NFS or CIFS file shares directly with Cloud Data Sense. These file shares can reside on-premises or in the cloud.

Quick start

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



Review file share prerequisites

For CIFS (SMB) shares, ensure that you have credentials to access the shares.



Deploy the Cloud Data Sense instance

Deploy [Cloud Data Sense](#) if there isn't already an instance deployed.



Create a group to hold the file shares

The group is a container for the file shares that you want to scan, and it is used as the working environment name for those file shares.



Add the file shares and select the shares to scan

Add the list of file shares that you want to scan and select the type of scanning. You can add up to 100 file shares at a time.

Reviewing file share requirements

Review the following prerequisites to make sure that you have a supported configuration before you enable Cloud Data Sense.

- The shares can be hosted anywhere, including in the cloud or on-premises. These are file shares that reside on non-NetApp storage systems.
- There needs to be network connectivity between the Data Sense instance and the shares.
- Make sure these ports are open to the Data Sense instance:
 - For NFS – ports 111 and 2049.
 - For CIFS – ports 139 and 445.
- You will need the list of shares you want to add in the format `<host_name>:/<share_path>`. You can enter the shares individually, or you can supply a line-separated list of the file shares you want to scan.
- For CIFS (SMB) shares, ensure that you have Active Directory credentials that provide read access to the shares. Admin credentials are preferred in case Cloud Data Sense needs to scan any data that requires elevated permissions.

Creating the group for the file shares

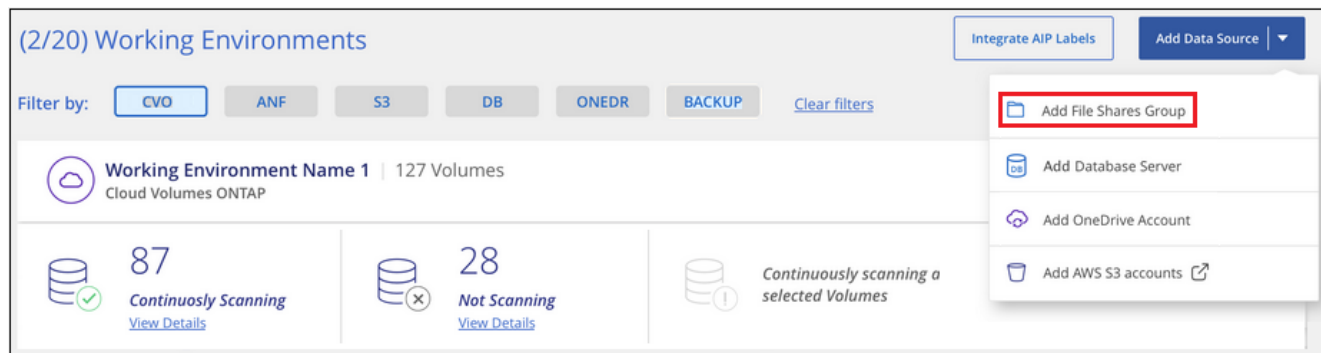
You must have [deployed an instance of Cloud Data Sense in Cloud Manager already](#).

You must add a file shares "group" before you can add your file shares. The group is a container for the file shares that you want to scan, and the group name is used as the working environment name for those file shares.

You can mix NFS and CIFS shares in the same group, however, all CIFS file shares in a group need to be using the same Active Directory credentials. If you plan to add CIFS shares that use different credentials, you must make a separate group for each unique set of credentials.

Steps

1. From the Working Environments Configuration page, click **Add Data Source > Add File Shares Group**.



2. In the Add Files Shares Group dialog, enter the name for the group of shares and click **Continue**.

The new File Shares Group is added to the list of working environments.

Adding file shares to a group

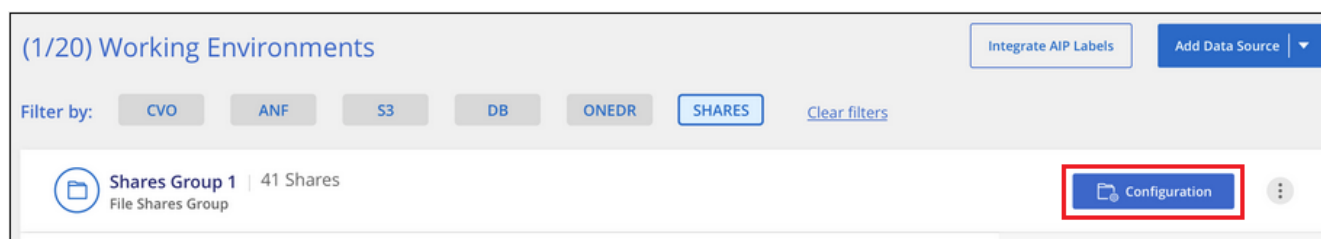
You add file shares to the File Shares Group so that the files in those shares will be scanned by Cloud Data Sense. You add the shares in the format <host_name>:/<share_path>.

You can add individual file shares, or you can supply a line-separated list of the file shares you want to scan. You can add up to 100 shares at a time.

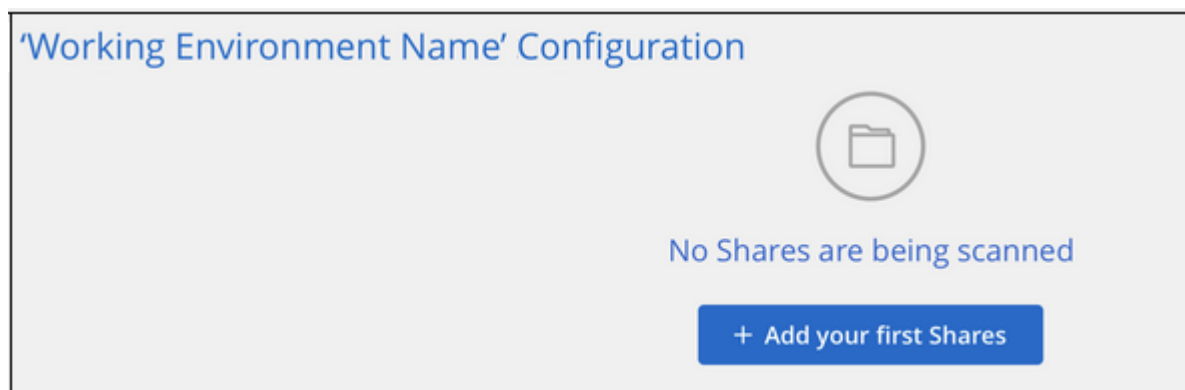
When adding both NFS and CIFS shares in a single group, you'll need to run through the process twice - once adding NFS shares, and then again adding the CIFS shares.

Steps

1. From the *Working Environments* page, click the **Configuration** button for the File Shares Group.



2. If this is the first time adding file shares for this File Shares Group, click **Add your first Shares**.



If you are adding file shares to an existing group, click **Add Shares**.

Working Environment 2 Configuration

+ Add Shares

Edit CIFS Credentials

2/22 Shares selected for compliance scan

Scan	Share name	Protocol	Status	Required Action
<div>Off</div> <div>Map</div> <div>Map & Classify</div>	Sharepath2	CIFS	<div></div> <div>Continuously Scanning</div>	...
<div>Off</div> <div>Map</div> <div>Map & Classify</div>	Sharepath3	NFS	<div></div> <div>Continuously Scanning</div>	...

- Select the protocol for the file shares you are adding, add the file shares that you want to scan - one file share per line - and click **Continue**.

When adding CIFS (SMB) shares, you need to enter the Active Directory credentials that provide read access to the shares. Admin credentials are preferred.

Adding Shares

Directly add any NFS or CIFS (SMB) File Shares, located in the cloud or on-premises.

Select Protocol

You'll be able to add additional shares from the other protocol later.

☒ NFS

☐ CIFS (SMB)

Type or paste below the Shares to add

Provide a list of shares, line-separated. You can add up to 100 at a time (you can add more later).

Hostname:/SHAREPATH

Hostname:/SHAREPATH

Hostname:/SHAREPATH

Continue

Cancel

☐ NFS

☒ CIFS (SMB)

Provide CIFS Credentials ⓘ

Username ⓘ

Password

A confirmation dialog displays the number of shares that were added.

If the dialog lists any shares that could not be added, capture this information so that you can resolve the issue. In some cases you can re-add the share with a corrected host name or share name.

- Enable mapping-only scans, or mapping and classification scans, on each file share.

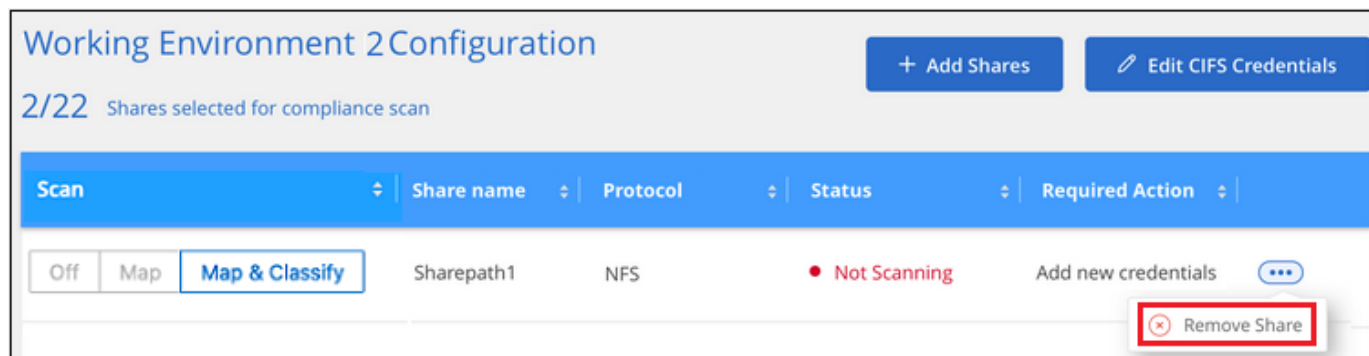
To:	Do this:
Enable mapping-only scans on file shares	Click Map
Enable full scans on file shares	Click Map & Classify
Disable scanning on file shares	Click Off

Result

Cloud Data Sense starts scanning the files in the file shares you added, and the results are displayed in the Dashboard and in other locations.

Removing a file share from compliance scans

If you no longer need to scan certain file shares, you can remove individual file shares from having their files scanned at any time. Just click **Remove Share** from the Configuration page.



Scanning object storage that uses S3 protocol

Complete a few steps to start scanning data within object storage directly with Cloud Data Sense. Data Sense can scan data from any Object Storage service which uses the Simple Storage Service (S3) protocol. This includes NetApp StorageGRID, IBM Cloud Object Store, Azure Blob (using MinIO), Linode, B2 Cloud Storage, and more.

Quick start

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



Review object storage prerequisites

You need to have the endpoint URL to connect with the object storage service.

You need to have the Access Key and Secret Key from the object storage provider so that Cloud Data Sense can access the buckets.



Deploy the Cloud Data Sense instance

Deploy [Cloud Data Sense](#) if there isn't already an instance deployed.



Add the Object Storage Service

Add the object storage service to Cloud Data Sense.

4

Select the buckets to scan

Select the buckets that you'd like to scan and Cloud Data Sense will start scanning them.

Reviewing object storage requirements

Review the following prerequisites to make sure that you have a supported configuration before you enable Cloud Data Sense.

- You need to have the endpoint URL to connect with the object storage service.
- You need to have the Access Key and Secret Key from the object storage provider so that Data Sense can access the buckets.
- Support for Azure Blob requires that you use the [MinIO service](#).

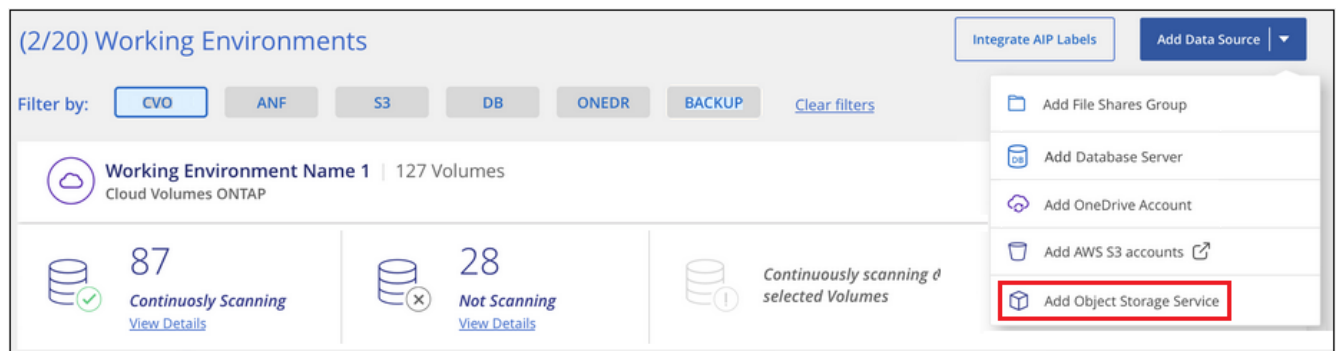
Adding the object storage service to Cloud Data Sense

You must have [deployed an instance of Cloud Data Sense in Cloud Manager already](#).

Add the object storage service.

Steps

1. From the Working Environments Configuration page, click **Add Data Source > Add Object Storage Service**.



2. In the Add Object Storage Service dialog, enter the details for the object storage service and click **Continue**.
 - a. Enter the name you want to use for the Working Environment. This name should reflect the name of the object storage service to which you are connecting.
 - b. Enter the Endpoint URL to access the object storage service.
 - c. Enter the Access Key and Secret Key so that Cloud Data Sense can access the buckets in the object storage.

Add Object Storage Service

Cloud Data Sense can scan data from any Object Storage service which uses the S3 protocol. This includes NetApp StorageGRID, IBM Object Store, and more.

To continue, enter the following information. In the next steps you'll need to select the buckets you want to scan.

Name the Working Environment	Endpoint URL
<input type="text" value="object_myIBM"/>	<input type="text" value="http://my.endpoint.com"/>
Access Key	Secret Key
<input type="text" value="AJUKDO574NDJG86795"/>	<input type="text" value="....."/>

Result

The new Object Storage Service is added to the list of working environments.

Enabling and disabling compliance scans on object storage buckets

After you enable Cloud Data Sense on your Object Storage Service, the next step is to configure the buckets that you want to scan. Data Sense discovers those buckets and displays them in the working environment you created.

Steps


1. In the Configuration page, click **Configuration** from the Object Storage Service working environment.

(1/20) Working Environments

Integrate AIP Labels

Add Data Source


Filter by: CVO ANF S3 DB ONEDR BCKUP **OB.STG** [Clear filters](#)



Rstor Integrated | 41 Buckets

Object Storage Service


Configuration




23

Continuously Scanning

[View Details](#)



All Buckets selected for Compliance



Continuously scanning all selected Buckets

2. Enable mapping-only scans, or mapping and classification scans, on your buckets.

35

Rstor Integrated Configuration

3/55 Buckets selected for Compliance scan

Q

Scan		Storage Repository (Bucket) ↓↑	Status ↓↑	Required Action ↓↑
Off	Map	Map & Classify	logs-759995470648-us-east-1	● Not Scanning
Off	Map	Map & Classify	logs-759995470648-us-west-2	● Not Scanning
Off	Map	Map & Classify	carstock	● Continuously Scanning

To:	Do this:
Enable mapping-only scans on a bucket	Click Map
Enable full scans on a bucket	Click Map & Classify
Disable scanning on a bucket	Click Off

Result

Cloud Data Sense starts scanning the buckets that you enabled. If there are any errors, they'll appear in the Status column, alongside the required action to fix the error.

Copyright Information

Copyright © 2021 NetApp, Inc. All rights reserved. Printed in the U.S. No part of this document covered by copyright may be reproduced in any form or by any means-graphic, electronic, or mechanical, including photocopying, recording, taping, or storage in an electronic retrieval system-without prior written permission of the copyright owner.

Software derived from copyrighted NetApp material is subject to the following license and disclaimer:

THIS SOFTWARE IS PROVIDED BY NETAPP "AS IS" AND WITHOUT ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE, WHICH ARE HEREBY DISCLAIMED. IN NO EVENT SHALL NETAPP BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

NetApp reserves the right to change any products described herein at any time, and without notice. NetApp assumes no responsibility or liability arising from the use of products described herein, except as expressly agreed to in writing by NetApp. The use or purchase of this product does not convey a license under any patent rights, trademark rights, or any other intellectual property rights of NetApp.

The product described in this manual may be protected by one or more U.S. patents, foreign patents, or pending applications.

RESTRICTED RIGHTS LEGEND: Use, duplication, or disclosure by the government is subject to restrictions as set forth in subparagraph (c)(1)(ii) of the Rights in Technical Data and Computer Software clause at DFARS 252.277-7103 (October 1988) and FAR 52-227-19 (June 1987).

Trademark Information

NETAPP, the NETAPP logo, and the marks listed at <http://www.netapp.com/TM> are trademarks of NetApp, Inc. Other company and product names may be trademarks of their respective owners.