



## **Gain insight into data privacy**

### **Cloud Manager**

NetApp  
April 22, 2021

# Table of Contents

- Gain insight into data privacy . . . . . 1
  - Learn about Cloud Compliance . . . . . 1
  - Get started . . . . . 5
  - Viewing governance details about the data stored in your organization . . . . . 44
  - Viewing compliance details about the data stored in your organization. . . . . 47
  - Managing your private data . . . . . 60
  - Adding personal data identifiers using Data Fusion. . . . . 70
  - Viewing compliance reports . . . . . 72
  - Responding to a Data Subject Access Request. . . . . 77
  - Categories of private data . . . . . 79
  - Removing data sources from Cloud Compliance. . . . . 84
  - Frequently asked questions about Cloud Compliance. . . . . 86

# Gain insight into data privacy

## Learn about Cloud Compliance

Cloud Compliance is a data privacy and compliance service for Cloud Manager that scans your volumes, Amazon S3 buckets, databases, OneDrive accounts, and other data sources to identify the personal and sensitive data that resides in those files. Using Artificial Intelligence (AI) driven technology, Cloud Compliance helps organizations understand data context and identify sensitive data.

[Learn about the use cases for Cloud Compliance.](#)

### Features

Cloud Compliance provides several tools that can help you with your compliance efforts. You can use Cloud Compliance to:

- Identify Personal Identifiable Information (PII)
- Identify a wide scope of sensitive information as required by GDPR, CCPA, PCI, and HIPAA privacy regulations
- Respond to Data Subject Access Requests (DSAR)
- Notify Cloud Manager users through email when files contain certain PII (you define this criteria using [Policies](#))
- View and modify [Azure Information Protection \(AIP\) labels](#) in your files
- Delete individual files

Cloud Compliance also provides tools that can help with your governance efforts. You can use Cloud Compliance to:

- Identify the stale data, non-business data, duplicate files, and very large files in your systems.

You can use this information to decide whether you want to move, delete, or tier some files to less expensive object storage.

- View the size of data and whether any of the data contains sensitive information prior to moving it.

This is useful if you are planning to migrate data from on-premises locations to the cloud.

### Supported working environments and data sources

Cloud Compliance can scan data from the following types of working environments and data sources:

- Cloud Volumes ONTAP in AWS
- Cloud Volumes ONTAP in Azure
- On-premises ONTAP clusters
- Azure NetApp Files
- Amazon S3

- Non-NetApp file shares
- Object storage (that uses S3 protocol)
- Databases
- OneDrive accounts



A Beta feature released in January 2021 allows you to run Compliance scans *for free* on the backup files created from your on-prem ONTAP volumes (created using [Cloud Backup](#)). This gives you a choice whether you want to have Cloud Compliance scan your on-prem ONTAP volumes directly, or scan the backup files made from those volumes.

## Cost

- The cost to use Cloud Compliance depends on the amount of data that you're scanning. The first 1 TB of data that Cloud Compliance scans in a Cloud Manager workspace is free. This includes all data from all working environments and data sources. A subscription to the AWS or Azure Marketplace is required to continue scanning data after that point. See [pricing](#) for details.

[Learn how to subscribe.](#)

**Note:** This subscription is not needed to scan backup files created from your on-prem ONTAP systems.

- Installing Cloud Compliance in the cloud requires deploying a cloud instance, which results in charges from the cloud provider where it is deployed. See [the type of instance that is deployed for each cloud provider](#). There is no cost if you install Cloud Compliance on an on-premises system.
- Cloud Compliance requires that you have deployed a Connector. In many cases you already have a Connector because of other storage and services you are using in Cloud Manager. The Connector instance results in charges from the cloud provider where it is deployed. See the [type of instance that is deployed for each cloud provider](#).

## Data transfer costs

Data transfer costs depend on your setup. If the Cloud Compliance instance and data source are in the same Availability Zone and region, then there are no data transfer costs. But if the data source, such as a Cloud Volumes ONTAP cluster or S3 Bucket, is in a *different* Availability Zone or region, then you'll be charged by your cloud provider for data transfer costs. See these links for more details:

- [AWS: Amazon EC2 Pricing](#)
- [Microsoft Azure: Bandwidth Pricing Details](#)

## How Cloud Compliance works

At a high-level, Cloud Compliance works like this:

1. You deploy an instance of Cloud Compliance in Cloud Manager.
2. You enable it on one or more working environments or data sources.
3. Cloud Compliance scans the data using an AI learning process.
4. You click **Compliance** and use the provided dashboards and reporting tools to help in your compliance efforts.

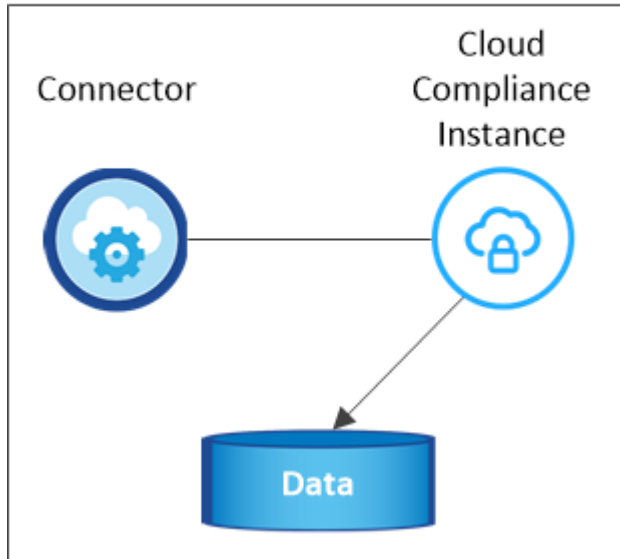
## The Cloud Compliance instance

When you deploy Cloud Compliance in the cloud, Cloud Manager deploys the instance in the same subnet as the Connector. [Learn more about Connectors.](#)



If the Connector is installed on-prem, it deploys the Cloud Compliance instance in same VPC or VNet as the first Cloud Volumes ONTAP system in the request.

### VPC or VNet



Note the following about the instance:

- In Azure, Cloud Compliance runs on a [Standard\\_D16s\\_v3 VM](#) with a 512 GB disk.
- In AWS, Cloud Compliance runs on an [m5.4xlarge instance](#) with a 500 GB GP2 disk.

In regions where m5.4xlarge isn't available, Cloud Compliance runs on an m4.4xlarge instance instead.



Changing or resizing the instance/VM type isn't supported. You need to use the size that's provided.

- The instance is named *CloudCompliance* with a generated hash (UUID) concatenated to it. For example: *CloudCompliance-16bb6564-38ad-4080-9a92-36f5fd2f71c7*
- Only one Cloud Compliance instance is deployed per Connector.
- Upgrades of Cloud Compliance software is automated—you don't need to worry about it.



The instance should remain running at all times because Cloud Compliance continuously scans the data.

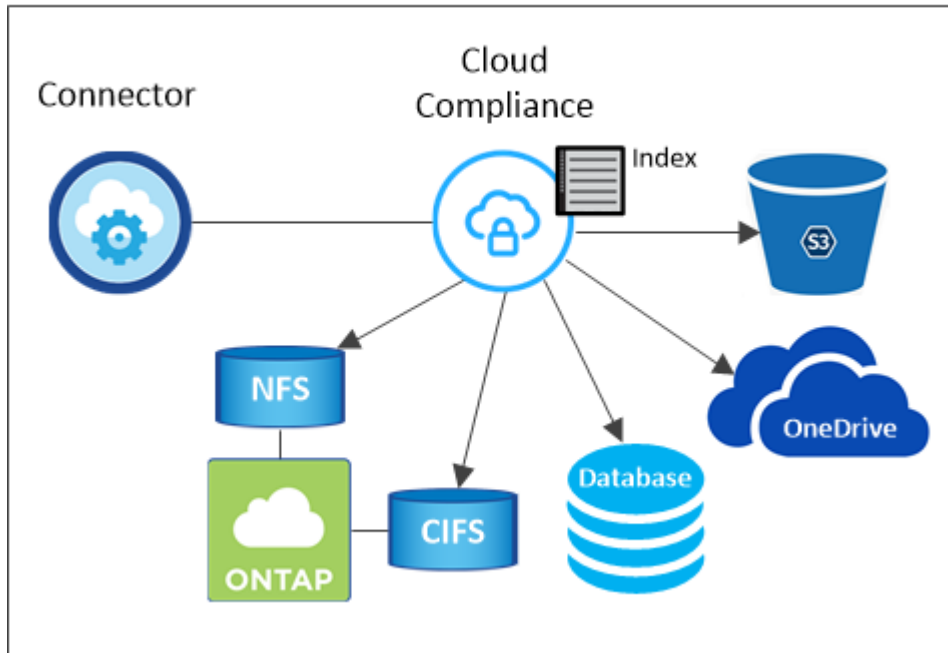
## How scans work

After you enable Cloud Compliance and select the volumes, buckets, database schemas, or OneDrive users you want to scan, it immediately starts scanning the data to identify personal and sensitive data. It maps your organizational data, categorizes each file, and identifies and extracts entities and predefined patterns in the data. The result of the scan is an index of personal information, sensitive personal information, data categories,

and file types.

Cloud Compliance connects to the data like any other client by mounting NFS and CIFS volumes. NFS volumes are automatically accessed as read-only, while you need to provide Active Directory credentials to scan CIFS volumes.

## VPC or VNet



After the initial scan, Cloud Compliance continuously scans your data to detect incremental changes (this is why it's important to keep the instance running).

You can enable and disable scans at the volume level, at the bucket level, at the database schema level, and at the OneDrive user level.

## Information that Cloud Compliance indexes

Cloud Compliance collects, indexes, and assigns categories to your data (files). The data that Cloud Compliance indexes includes the following:

### Standard metadata

Cloud Compliance collects standard metadata about files: the file type, its size, creation and modification dates, and so on.

### Personal data

Personally identifiable information such as email addresses, identification numbers, or credit card numbers. [Learn more about personal data.](#)

### Sensitive personal data

Special types of sensitive information, such as health data, ethnic origin, or political opinions, as defined by GDPR and other privacy regulations. [Learn more about sensitive personal data.](#)

### Categories

Cloud Compliance takes the data that it scanned and divides it into different types of categories. Categories are topics based on AI analysis of the content and metadata of each file. [Learn more about categories.](#)

## Types

Cloud Compliance takes the data that it scanned and breaks it down by file type. [Learn more about types.](#)

## Name entity recognition

Cloud Compliance uses AI to extract natural persons' names from documents. [Learn about responding to Data Subject Access Requests.](#)

## Networking overview

Cloud Manager deploys the Cloud Compliance instance with a security group that enables inbound HTTP connections from the Connector instance.

When using Cloud Manager in SaaS mode, the connection to Cloud Manager is served over HTTPS, and the private data sent between your browser and the Cloud Compliance instance are secured with end-to-end encryption, which means NetApp and third parties can't read it.

If you need to use the local user interface instead of the SaaS user interface for any reason, you can still [access the local UI](#).

Outbound rules are completely open. Internet access is needed to install and upgrade the Cloud Compliance software and to send usage metrics.

If you have strict networking requirements, [learn about the endpoints that Cloud Compliance contacts](#).

## User access to compliance information

The role each user has been assigned provides different capabilities within Cloud Manager and within Cloud Compliance:

- An **Account Admin** can manage compliance settings and view compliance information for all working environments.
- A **Workspace Admin** can manage compliance settings and view compliance information only for systems that they have permissions to access. If a Workspace Admin can't access a working environment in Cloud Manager, then they can't see any compliance information for the working environment in the Compliance tab.
- Users with the **Compliance Viewer** role can only view compliance information and generate reports for systems that they have permission to access. These users cannot enable/disable scanning of volumes, buckets, or database schemas.

[Learn more about Cloud Manager roles](#) and how to [add users with specific roles](#).

## Get started

### Deploy Cloud Compliance

Complete a few steps to deploy the Cloud Compliance instance in your Cloud Manager workspace. You can deploy Cloud Compliance in the cloud or on an on-premises system.

The on-prem installation may be a good option if you prefer to scan on-premises ONTAP working environments using a Compliance instance that's also located on premises. But this is not a requirement. The Compliance software functions exactly the same way regardless of which installation method you choose.

## Quick start

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

### 1

#### Create a Connector

If you don't already have a Connector, create a Connector in Azure or AWS. See [creating a Connector in AWS](#) or [creating a Connector in Azure](#).

You can also [deploy the Connector on-premises](#) on an existing Linux host in your network or in the cloud.

### 2

#### Review prerequisites

Ensure that your environment can meet the prerequisites. This includes outbound internet access for the instance, connectivity between the Connector and Cloud Compliance over port 80, and more. [See the complete list](#).

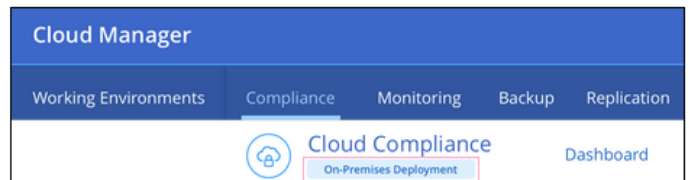
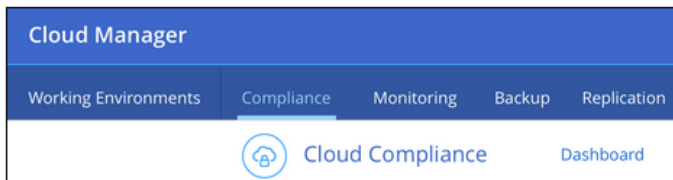
- When installed in the cloud, you need 16 vCPUs for the Cloud Compliance instance. See [more details about the instance type](#).
- When installed on premises, you need a Linux system that meets the [following requirements](#).

### 3

#### Deploy Cloud Compliance

Launch the installation wizard to deploy the Cloud Compliance instance.

You can deploy Cloud Compliance in the cloud or in an on-premises location. The only difference you'll notice in the UI is the words "On-Premises Deployment".



### 4

#### Subscribe to the Cloud Compliance service

The first 1 TB of data that Cloud Compliance scans in Cloud Manager is free. A subscription to the AWS or Azure Marketplace is required to continue scanning data after that point.

## Creating a Connector

If you don't already have a Connector, create a Connector in Azure or AWS. See [creating a Connector in AWS](#) or [creating a Connector in Azure](#). In most cases you will probably have a Connector set up before you attempt to activate Cloud Compliance because most [Cloud Manager features require a Connector](#), but there are cases where you'll need to set one up now.

There are some scenarios where you have to use a Connector that's deployed in AWS or Azure:



- When scanning data in Cloud Volumes ONTAP in AWS or in AWS S3 buckets, you use a connector in AWS.
- When scanning data in Cloud Volumes ONTAP in Azure or in Azure NetApp Files, you use a connector in Azure.

On-prem ONTAP systems, non-NetApp file shares, generic S3 Object storage, databases, and OneDrive folders can be scanned using either Connector.

Note that you can also [deploy the Connector on-premises](#) on an existing Linux host in your network or in the cloud. Some users planning to install Cloud Compliance on-prem may also choose to install the Connector on-prem.

As you can see, there may be some situations where you need to use [multiple Connectors](#).



If you're planning on scanning Azure NetApp Files, you need to make sure you're deploying in the same region as the volumes you wish to scan.

## Reviewing prerequisites

Review the following prerequisites to make sure that you have a supported configuration before you deploy Cloud Compliance.

## Enable outbound internet access from Cloud Compliance

Cloud Compliance requires outbound internet access. If your virtual or physical network uses a proxy server for internet access, ensure that the Cloud Compliance instance has outbound internet access to contact the following endpoints. When you deploy Cloud Compliance in the cloud, it's located in the same subnet as the Connector.

Review the appropriate table below depending on whether you are deploying Cloud Compliance in AWS, Azure, or on-premises.

### Required endpoints for AWS deployments:

Endpoints	Purpose
<a href="https://cloudmanager.cloud.netapp.com">https://cloudmanager.cloud.netapp.com</a>	Communication with the Cloud Manager service, which includes Cloud Central accounts.
<a href="https://netapp-cloud-account.auth0.com">https://netapp-cloud-account.auth0.com</a> <a href="https://auth0.com">https://auth0.com</a>	Communication with NetApp Cloud Central for centralized user authentication.
<a href="https://cloud-compliance-support-netapp.s3.us-west-2.amazonaws.com">https://cloud-compliance-support-netapp.s3.us-west-2.amazonaws.com</a> <a href="https://hub.docker.com">https://hub.docker.com</a> <a href="https://auth.docker.io">https://auth.docker.io</a> <a href="https://registry-1.docker.io">https://registry-1.docker.io</a> <a href="https://index.docker.io/">https://index.docker.io/</a> <a href="https://dseasb33srrn.cloudfront.net/">https://dseasb33srrn.cloudfront.net/</a> <a href="https://production.cloudflare.docker.com/">https://production.cloudflare.docker.com/</a>	Provides access to software images, manifests, and templates.
<a href="https://kinesis.us-east-1.amazonaws.com">https://kinesis.us-east-1.amazonaws.com</a>	Enables NetApp to stream data from audit records.

Endpoints	Purpose
<a href="https://cognito-idp.us-east-1.amazonaws.com">https://cognito-idp.us-east-1.amazonaws.com</a> <a href="https://cognito-identity.us-east-1.amazonaws.com">https://cognito-identity.us-east-1.amazonaws.com</a> <a href="https://user-feedback-store-prod.s3.us-west-2.amazonaws.com">https://user-feedback-store-prod.s3.us-west-2.amazonaws.com</a> <a href="https://customer-data-production.s3.us-west-2.amazonaws.com">https://customer-data-production.s3.us-west-2.amazonaws.com</a>	Enables Cloud Compliance to access and download manifests and templates, and to send logs and metrics.

#### Required endpoints for Azure and On-Prem deployments:

Endpoints	Purpose
<a href="https://cloudmanager.cloud.netapp.com">https://cloudmanager.cloud.netapp.com</a>	Communication with the Cloud Manager service, which includes Cloud Central accounts.
<a href="https://netapp-cloud-account.auth0.com">https://netapp-cloud-account.auth0.com</a> <a href="https://auth0.com">https://auth0.com</a>	Communication with NetApp Cloud Central for centralized user authentication.
<a href="https://support.compliance.cloudmanager.cloud.netapp.com/">https://support.compliance.cloudmanager.cloud.netapp.com/</a> <a href="https://hub.docker.com">https://hub.docker.com</a> <a href="https://auth.docker.io">https://auth.docker.io</a> <a href="https://registry-1.docker.io">https://registry-1.docker.io</a> <a href="https://index.docker.io/">https://index.docker.io/</a> <a href="https://dseasb33srrn.cloudfront.net/">https://dseasb33srrn.cloudfront.net/</a> <a href="https://production.cloudflare.docker.com/">https://production.cloudflare.docker.com/</a>	Provides access to software images, manifests, templates, and to send logs and metrics.
<a href="https://support.compliance.cloudmanager.cloud.netapp.com/">https://support.compliance.cloudmanager.cloud.netapp.com/</a>	Enables NetApp to stream data from audit records.
<b>On-premises installs only:</b> <a href="https://github.com/docker">https://github.com/docker</a> <a href="https://download.docker.com">https://download.docker.com</a> <a href="https://rhui3.us-west-2.aws.ce.redhat.com">https://rhui3.us-west-2.aws.ce.redhat.com</a> <a href="https://github-production-release-asset-2e65be.s3.amazonaws.com">https://github-production-release-asset-2e65be.s3.amazonaws.com</a> <a href="https://pypi.org">https://pypi.org</a> <a href="https://pypi.python.org">https://pypi.python.org</a> <a href="https://files.pythonhosted.org">https://files.pythonhosted.org</a> <a href="http://mirror.centos.org">http://mirror.centos.org</a> <a href="http://mirrorlist.centos.org">http://mirrorlist.centos.org</a> <a href="http://mirror.centos.org/centos/7/extras/x86_64/Packages/container-selinux-2.107-3.el7.noarch.rpm">http://mirror.centos.org/centos/7/extras/x86_64/Packages/container-selinux-2.107-3.el7.noarch.rpm</a>	Provides prerequisite packages for installation.

#### Ensure that Cloud Manager has the required permissions

Ensure that Cloud Manager has permissions to deploy resources and create security groups for the Cloud Compliance instance. You can find the latest Cloud Manager permissions in [the policies provided by NetApp](#).

## Check your vCPU limits

When installed in the cloud, ensure that your cloud provider's vCPU limit allows for the deployment of an instance with 16 cores. You'll need to verify the vCPU limit for the relevant instance family in the region where Cloud Manager is running.

In AWS, the instance family is *On-Demand Standard instances*. In Azure, the instance family is *Standard Dsv3 Family*.

See the following links for more details on vCPU limits:

- [AWS documentation: Amazon EC2 Service Limits](#)
- [Azure documentation: Virtual machine vCPU quotas](#)

## Ensure that Cloud Manager can access Cloud Compliance

Ensure connectivity between the Connector and the Cloud Compliance instance. The security group for the Connector must allow inbound and outbound traffic over port 80 to and from the Cloud Compliance instance.

This connection enables deployment of the Cloud Compliance instance and enables you to view information in the Compliance tab.

## Ensure that you can keep Cloud Compliance running

The Cloud Compliance instance needs to stay on to continuously scan your data.

## Ensure web browser connectivity to Cloud Compliance

After Cloud Compliance is enabled, ensure that users access the Cloud Manager interface from a host that has a connection to the Cloud Compliance instance.

The Cloud Compliance instance uses a private IP address to ensure that the indexed data isn't accessible to the internet. As a result, the web browser that you use to access Cloud Manager must have a connection to that private IP address. That connection can come from a direct connection to AWS or Azure (for example, a VPN), or from a host that's inside the same network as the Cloud Compliance instance.


## Deploying the Cloud Compliance instance in the cloud

Deploying an instance of Cloud Compliance in the cloud is the most common deployment model. But you have the option to [deploy the Compliance software on a Linux host](#) in your network or in the cloud.

The Compliance software functions exactly the same way regardless of which installation method you choose.

## Steps

1. In Cloud Manager, click **Compliance**.
2. Click **Activate Cloud Compliance**.



## Cloud Compliance


[How does it work? \(2\)](#)

### Always-on Privacy & Compliance Controls

Automated controls for data privacy regulations such as the GDPR, CCPA and more. Driven by powerful artificial intelligence algorithms, Cloud Compliance gets your business application data and cloud environments privacy ready.

**Activate Cloud Compliance**

#### Compliance Status



**Data Distribution**

- 75% Non-Sensitive
- 20% Personal
- 5% Sensitive Personal

**28,000** Personal Files [View All](#)


Email Address	2,700 Files
Credit Card	2,700 Files

**7,000** Sensitive Personal Files [View All](#)

Health	2,700 Files
Ethnicity	2,700 Files

- Click **Activate Compliance** to start the cloud deployment wizard.

### Select where to deploy Compliance




#### Deploy Compliance in the Cloud

**Recommended**

**Activate Compliance**

▶ We recommend deploying Compliance in the Cloud. Selecting this option will deploy the instance in the same location as the Cloud Manager Connector instance.






#### Deploy Compliance On-Premises

**Activate Compliance**

- The wizard displays progress as it goes through the deployment steps. It will stop and ask for input if it runs into any issues.

### Deploying Cloud Compliance

This may take up to 15 minutes. Check this page periodically to make sure the deployment continues successfully.



#### Deploying Cloud Compliance instance

Verifying connectivity to the Cloud Manager and to the Internet

Initializing Cloud Compliance

[Cancel deployment](#)

5. When the instance is deployed, click **Continue to configuration** to go to the *Scan Configuration* page.

## Result

Cloud Manager deploys the Cloud Compliance instance in your cloud provider.

## What's Next

From the Scan Configuration page you can select the data sources that you want to scan.

You can also [subscribe to the Cloud Compliance service](#) at this time. You will not be charged until the amount of data exceeds 1 TB.

## Deploying the Cloud Compliance instance on premises

You can download and install the Compliance software on a Linux host in your network if you do not want to [deploy it in the cloud](#).

The Compliance software functions exactly the same regardless of which installation method you choose.



Cloud Compliance is currently unable to scan S3 buckets and Azure NetApp Files when the Compliance instance is installed on premises. In these cases you'll need to deploy a separate Connector and instance of Compliance in the cloud and [switch between Connectors](#) for your different data sources.


## Host requirements

- Operating system: Red Hat Enterprise Linux or CentOS version 8.0 or 8.1
  - Version 7.8 can be used, but the Linux kernel version must be 4.14 or greater
  - The OS must be capable of installing the docker engine (for example, disable the *firewalld* service if needed)
- RAM: 64 GB (swap memory must be disabled on the host)
- CPU: 16 cores
- Disk: 500 GB SSD
- A Red Hat Enterprise Linux system must be registered with Red Hat Subscription Management. If it is not registered, the system cannot access repositories to update required 3rd party software during installation.
- Make sure port 8080 is open so you can see the installation progress in Cloud Manager.
- Root privileges are required to install Cloud Compliance.

See [Reviewing prerequisites](#) for the full list of requirements and endpoints that Cloud Compliance must be able to reach over the internet.

## Steps

1. Download the Cloud Compliance software from the [NetApp Support Site](#).
2. Copy the installer file to the Linux host you plan to use (using `scp` or some other method).
3. In Cloud Manager, click **Compliance**.
4. Click **Activate Cloud Compliance**.


Cloud Compliance


[How does it work? \(2\)](#)

## Always-on Privacy & Compliance Controls

Automated controls for data privacy regulations such as the GDPR, CCPA and more. Driven by powerful artificial intelligence algorithms, Cloud Compliance gets your business application data and cloud environments privacy ready.

Activate Cloud Compliance

Compliance Status



Data Distribution

75% Non-Sensitive

20% Personal

5% Sensitive Personal

28,000 Personal Files

View All

Email Address

2,700 Files

Credit Card

2,700 Files

7,000 Sensitive Personal Files

View All

Health


2,700 Files

Ethnicity

2,700 Files

- Click **Activate Compliance** to start the on-prem deployment wizard.

### Select where to deploy Compliance




Deploy Compliance in the Cloud

Recommended

Activate Compliance

▼



Deploy Compliance On-Premises

Activate Compliance

▲

> For special situations, for example, if you wish to scan on-premises Working Environments and you prefer Compliance accesses the data from an on-premises location.

- In the *Deploy Cloud Compliance On Premises* dialog, copy the provided command and paste it in a text file so you can use it later. For example:

```
sudo ./install.sh -a 12345 -c 27AG75 -t 2198qq
```

- Unzip the installer file on the host machine:

```
tar -xzf cc_onprem_installer.tar.gz
```

- When prompted by the installer, you can enter the required values in a series of prompts, or you can enter the complete command in the first prompt:

Enter parameters as prompted:	Enter the full command:
<ol style="list-style-type: none"> <li>1. Paste the information you copied from step 6:  <pre>sudo ./install.sh -a &lt;account_id&gt; -c &lt;agent_id&gt; -t &lt;token&gt;</pre> </li> <li>2. Enter the IP address or host name of the Compliance host machine so it can be accessed by the Connector instance.</li> <li>3. Enter the IP address or host name of the Cloud Manager Connector host machine so it can be accessed by the Cloud Compliance instance.</li> <li>4. Enter proxy details as prompted. If your Cloud Manager already uses a proxy, there is no need to enter this information again here since Cloud Compliance will automatically use the proxy used by Cloud Manager.</li> </ol>	<p>Alternatively, you can create the whole command in advance and enter it in the first prompt:</p> <pre>sudo ./install.sh -a &lt;account_id&gt; -c &lt;agent_id&gt; -t &lt;token&gt; --host &lt;cc_host&gt; --cm-host &lt;cm_host&gt; --proxy-host &lt;proxy_host&gt; --proxy-port &lt;proxy_port&gt; --proxy-scheme &lt;proxy_scheme&gt; --proxy -user &lt;proxy_user&gt; --proxy-password &lt;proxy_password&gt;</pre>

Variable values:

- *account\_id* = NetApp Account ID
- *agent\_id* = Connector ID
- *token* = jwt user token
- *cc\_host* = IP address or host name of the Cloud Compliance Linux system.
- *cm\_host* = IP address or host name of the Cloud Manager Connector system.
- *proxy\_host* = IP or host name of the proxy server if the host is behind a proxy server.
- *proxy\_port* = Port to connect to the proxy server (default 80).
- *proxy\_scheme* = Connection scheme: https or http (default http).
- *proxy\_user* = Authenticated user to connect to the proxy server, if basic authentication is required.
- *proxy\_password* = Password for the user name that you specified.

## Result

The Cloud Compliance installer installs packages, installs docker, registers the installation, and installs Cloud Compliance. Installation can take 10 to 20 minutes.

If there is connectivity over port 8080 between the host machine and the Connector instance, you will see the installation progress in the Compliance tab in Cloud Manager.

## What's Next

From the Scan Configuration page you can select the data sources that you want to scan.

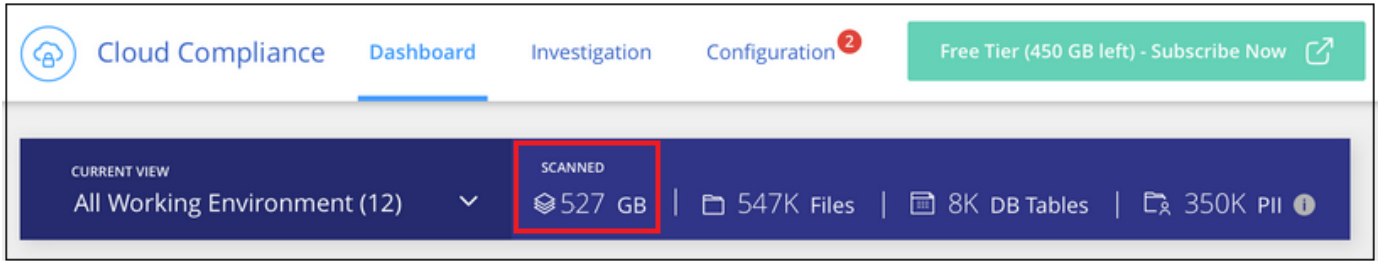
You can also [subscribe to the Cloud Compliance service](#) at this time. You will not be charged until the amount of data exceeds 1 TB. A subscription to either the AWS or Azure Marketplace can be used when you have deployed Cloud Compliance on an on-premises system.

## Subscribing to the Cloud Compliance service

The first 1 TB of data that Cloud Compliance scans in a Cloud Manager workspace is free. A subscription to

the AWS or Azure Marketplace is required to continue scanning data after that point.

You can subscribe at any time and you will not be charged until the amount of data exceeds 1 TB. You can always see the total amount of data that is being scanned from the Cloud Compliance Dashboard. And the *Subscribe Now* button makes it easy to subscribe when you are ready.



**Note:** If you are prompted by Cloud Compliance to subscribe, but you already have an Azure subscription, you're probably using the old **Cloud Manager** subscription and you need to change to the new **NetApp Cloud Manager** subscription. See [Changing to the new NetApp Cloud Manager plan in Azure](#) for details.

**Steps**

These steps must be completed by a user who has the *Account Admin* role.

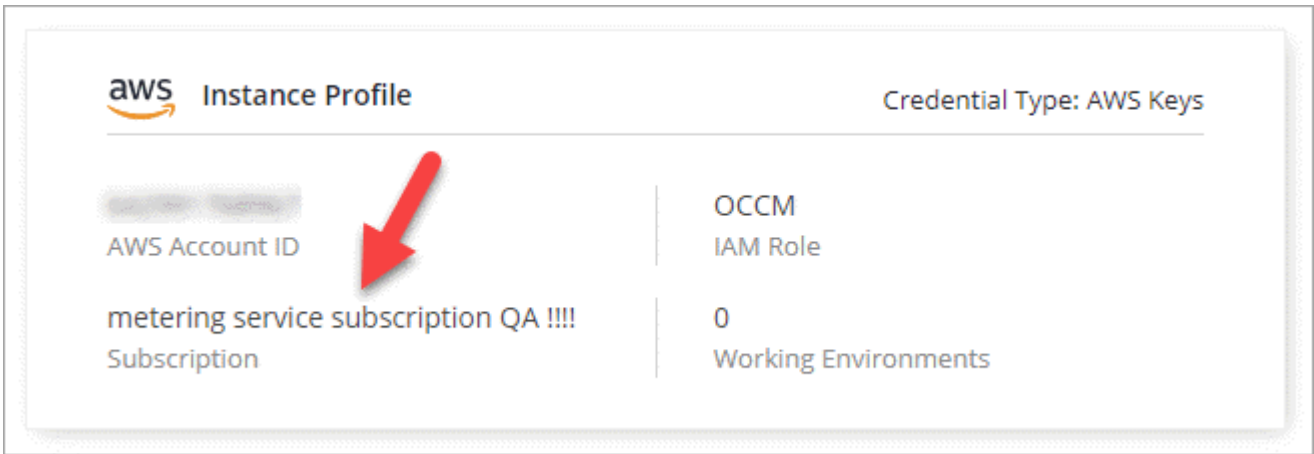
- 1. In the upper right of the Cloud Manager console, click the Settings icon, and select **Credentials**.



- 2. Find the credentials for the AWS Instance Profile or Azure Managed Service Identity.

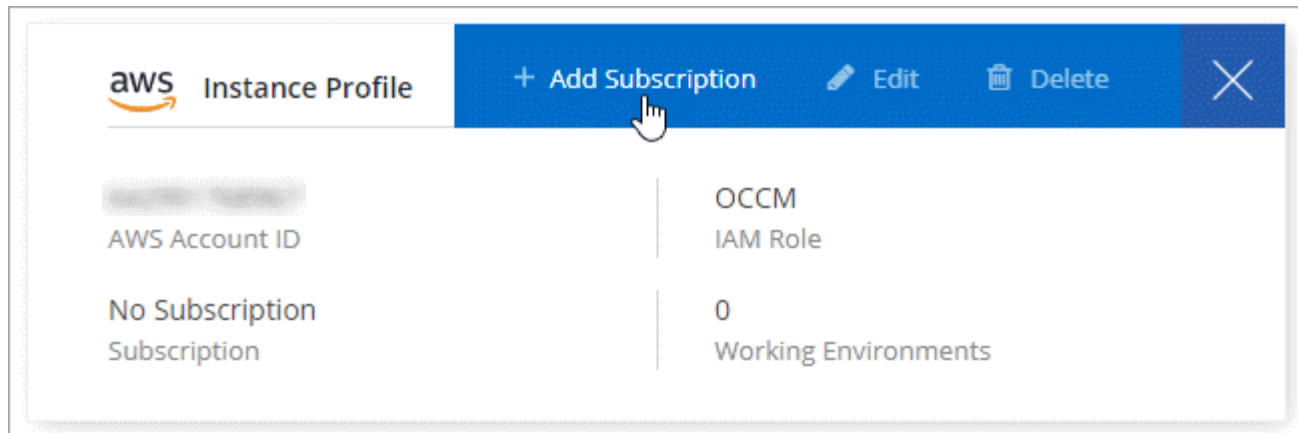
The subscription must be added to the Instance Profile or Managed Service Identity. Charging won't work otherwise.

If you already have a subscription, then you're all set—there's nothing else that you need to do.



- 3. If you don't have a subscription yet, hover over the credentials and click the action menu.
- 4. Click **Add Subscription**.





5. Click **Add Subscription**, click **Continue**, and follow the steps.

The following video shows how to associate a Marketplace subscription to an AWS subscription:

► [https://docs.netapp.com/us-en/occm/media/video\\_subscribing\\_aws.mp4](https://docs.netapp.com/us-en/occm/media/video_subscribing_aws.mp4) (video)

The following video shows how to associate a Marketplace subscription to an Azure subscription:

► [https://docs.netapp.com/us-en/occm/media/video\\_subscribing\\_azure.mp4](https://docs.netapp.com/us-en/occm/media/video_subscribing_azure.mp4) (video)

### Changing to the new Cloud Manager plan in Azure

Cloud Compliance was added to the Azure Marketplace subscription named **NetApp Cloud Manager** as of October 7, 2020. If you already have the original Azure **Cloud Manager** subscription it will not allow you to use Cloud Compliance.

You need to follow these steps to change to the new **NetApp Cloud Manager** subscription before you can start using Cloud Compliance.



If your existing Subscription was issued with a special private offer, you need to contact NetApp so that we can issue a new special private offer with Compliance included.

### Steps

1. In the upper right of the Cloud Manager console, click the Settings icon, and select **Credentials**.
2. Find the credentials for the Azure Managed Service Identity that you want to change the subscription for and hover over the credentials and click **Associate Subscription**.

The details for your current Marketplace Subscription are displayed.

3. Log in to the [Azure portal](#) and select **Software as a Service (SaaS)**.
4. Select the subscription for which you want to change the plan and click **Change Plan**.

The screenshot shows the Azure portal interface for a subscription named 'shiranSub3008'. On the left, a list of subscriptions includes 'shiranSub3008', 'shiran0510', and 'shiranDemoSub'. The main area displays the subscription details for 'shiranSub3008', including the offer 'Cloud Manager - Cloud Manager - Monthly' and the current plan 'Cloud Manager - Monthly'. The 'Change plan' button is highlighted in the 'Offer and plan details' section.

5. In the Change Plan page, select the **NetApp Cloud Manager** plan and click the **Change Plan** button.

The screenshot shows the 'Change plan' page in the Azure portal. It displays a table of available plans for the 'Cloud Manager' software. The 'NetApp Cloud Manager' plan is selected, and the 'Change plan' button is highlighted at the bottom of the page.

Software plan	Description	Price
<input checked="" type="radio"/> NetApp Cloud Manager	PLAN - INCLUDES COMPLIANCE	\$0.00 per month Plus: CVO Explore HA upto 2TB in HA pair \$0.49/node/hour: \$0.49 per node CVO Premium plan, up to 368TB (\$3.19/node/hour): \$3.19 per node CVO Standard plan, up to 10TB (\$1.98/node/hour): \$1.98 per node Cloud Compliance \$50/TB/Month: \$0.068 per tb/hour CVO Premium HA 368TB in HA pair \$2.56/node/hour: \$2.56 per node CVO Standar HA 10TB in HA pair \$1.77/node/hour: \$1.77 per node Cloud Tiering for On Prem ONTAP (\$0.07/TB/hour): \$0.07 per tb/hour Backup CVO to Blob \$50/TB per month (calc hourly): \$0.07 per tb/hour CVO Explore plan, up to 2TB (\$0.75/node/hour): \$0.75 per node
<input type="radio"/> Cloud Manager	OLD PLAN - DOES NOT INCLUDE COMPLIANCE	\$0.00 per month Plus: CVO Explore HA upto 2TB in HA pair \$0.49/node/hour: \$0.49 per node Backup CVO to Blob \$50/TB per month (calc hourly): \$0.07 per tb/hour CVO Standar HA 10TB in HA pair \$1.77/node/hour: \$1.77 per node Restore CVO to Blob \$50/TB per month (calc hourly): \$0.07 per tb/hour CVO Standard plan, up to 10TB (\$1.98/node/hour): \$1.98 per node CVO Premium HA 368TB in HA pair \$2.56/node/hour: \$2.56 per node CVO Premium plan, up to 368TB (\$3.19/node/hour): \$3.19 per node Cloud Tiering for On Prem ONTAP (\$0.07/TB/hour): \$0.07 per tb/hour CVO Explore plan, up to 2TB (\$0.75/node/hour): \$0.75 per node

6. Return to Cloud Manager, select the subscription, and hover over the "i" above subscription in the Credentials card to verify your subscription has changed.

# Activate scanning on your data sources

## Getting started with Cloud Compliance for Cloud Volumes ONTAP, on-premises ONTAP, or Azure NetApp Files

Complete a few steps to get started with Cloud Compliance for Cloud Volumes ONTAP, on-premises ONTAP systems, or 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 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](#)
- For Azure NetApp Files, [Cloud Manager must be set up to discover the configuration](#).



### Deploy the Cloud Compliance instance

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



### Enable Cloud Compliance 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 Compliance is enabled, ensure that it can access all volumes.

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

Click **Compliance** > **Scan 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 Compliance 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](#). And for Azure NetApp Files, [Cloud Manager must be set up to discover the configuration](#).

### Deploying the Cloud Compliance instance

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

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

Cloud Compliance 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.

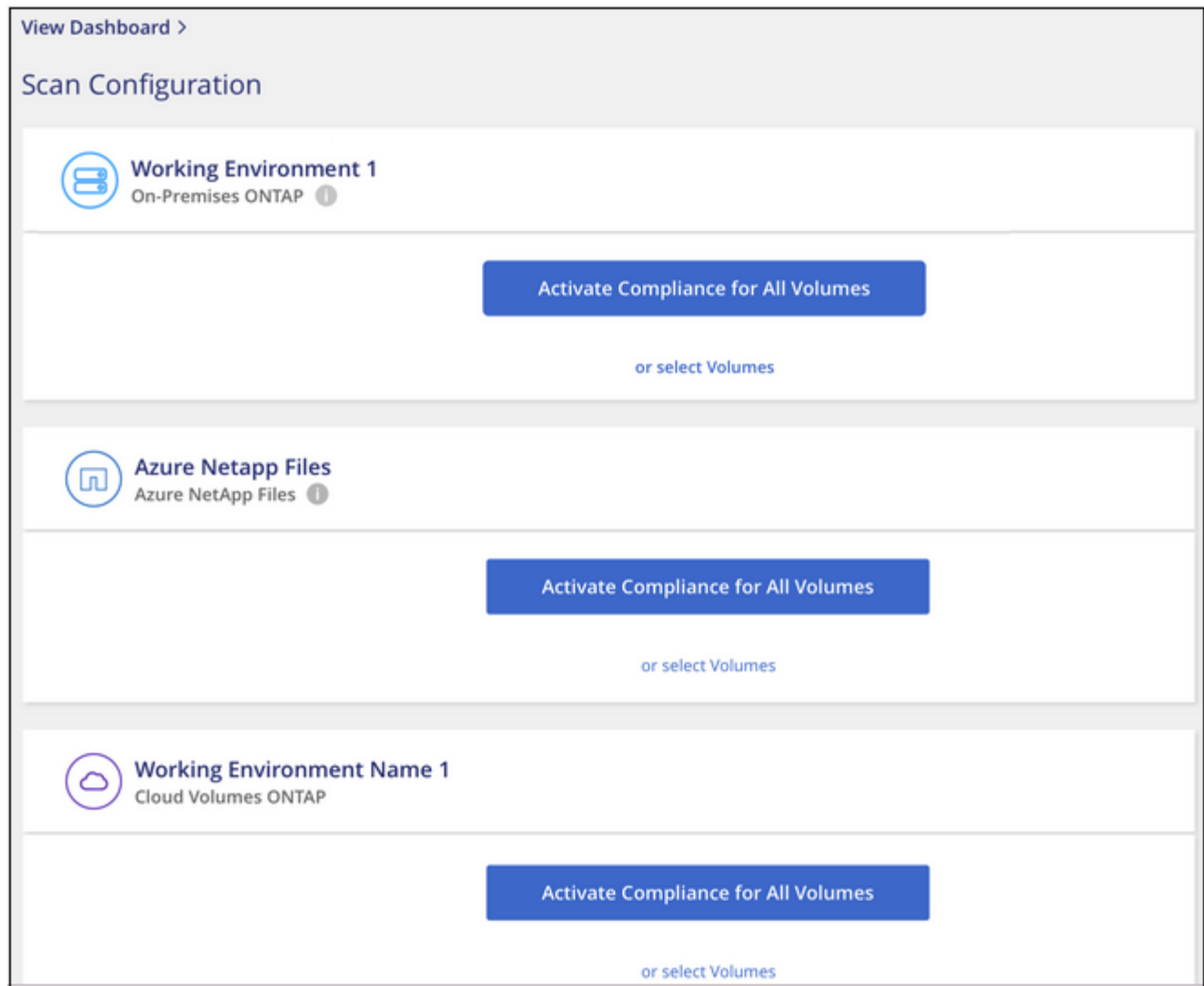
### Enabling Cloud Compliance in your working environments

You can enable Cloud Compliance on Cloud Volumes ONTAP systems (in AWS and Azure), on-premises ONTAP clusters, and Azure NetApp Files.



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 **Compliance** and then select the **Configuration** tab.



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

To scan only certain volumes in a working environment, 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 Compliance starts scanning the volumes you selected in the working environment. Results will be available in the Compliance dashboard as soon as Cloud Compliance 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 Compliance has access to volumes

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

## Steps

1. Make sure that there's a network connection between the Cloud Compliance instance and each network that includes volumes for Cloud Volumes ONTAP, Azure NetApp Files, or on-prem ONTAP clusters.

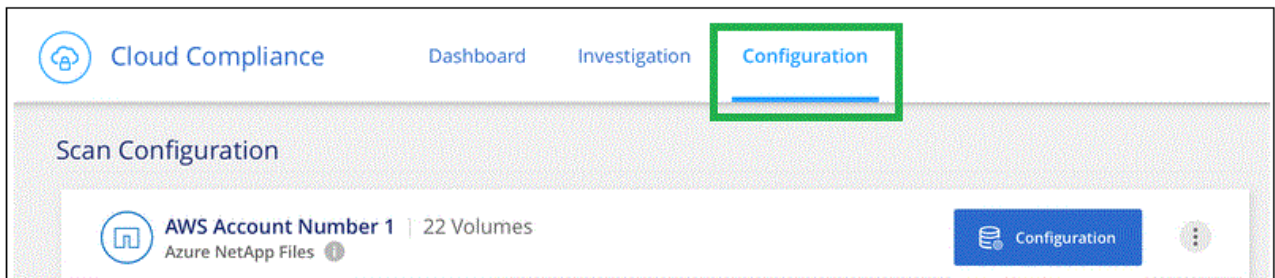


For Azure NetApp Files, Cloud Compliance can only scan volumes that are in the same region as Cloud Manager.

2. Ensure that the security group for Cloud Volumes ONTAP allows inbound traffic from the Cloud Compliance instance.

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

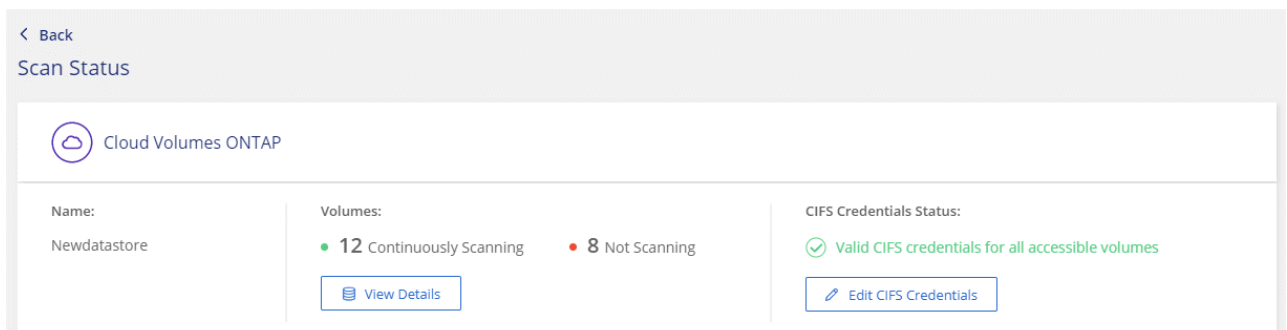
3. Ensure the following ports are open to the Cloud Compliance instance:
  - For NFS – ports 111 and 2049.
  - For CIFS – ports 139 and 445.
4. Ensure that NFS volume export policies include the IP address of the Cloud Compliance instance so it can access the data on each volume.
5. If you use CIFS,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



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

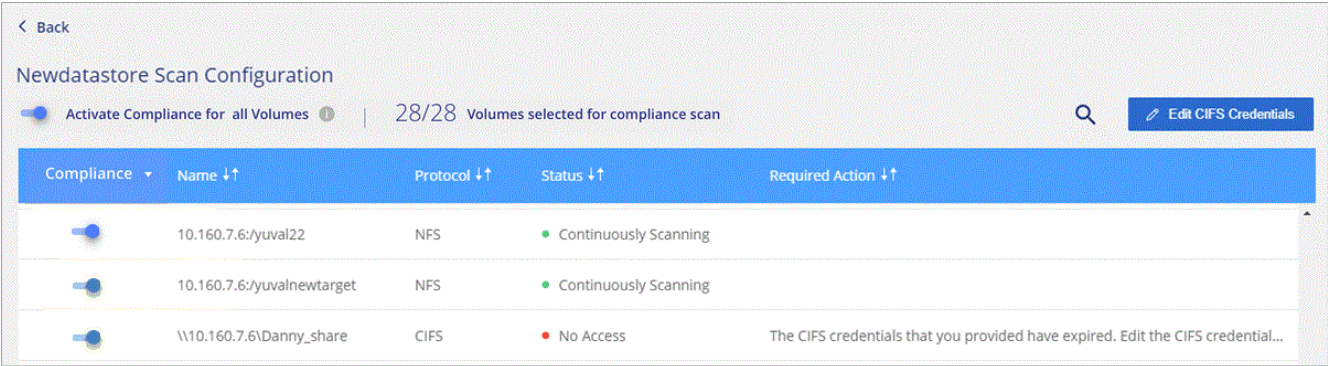
The credentials can be read-only, but providing admin credentials ensures that Cloud Compliance can read any data that requires elevated permissions. The credentials are stored on the Cloud Compliance instance.

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



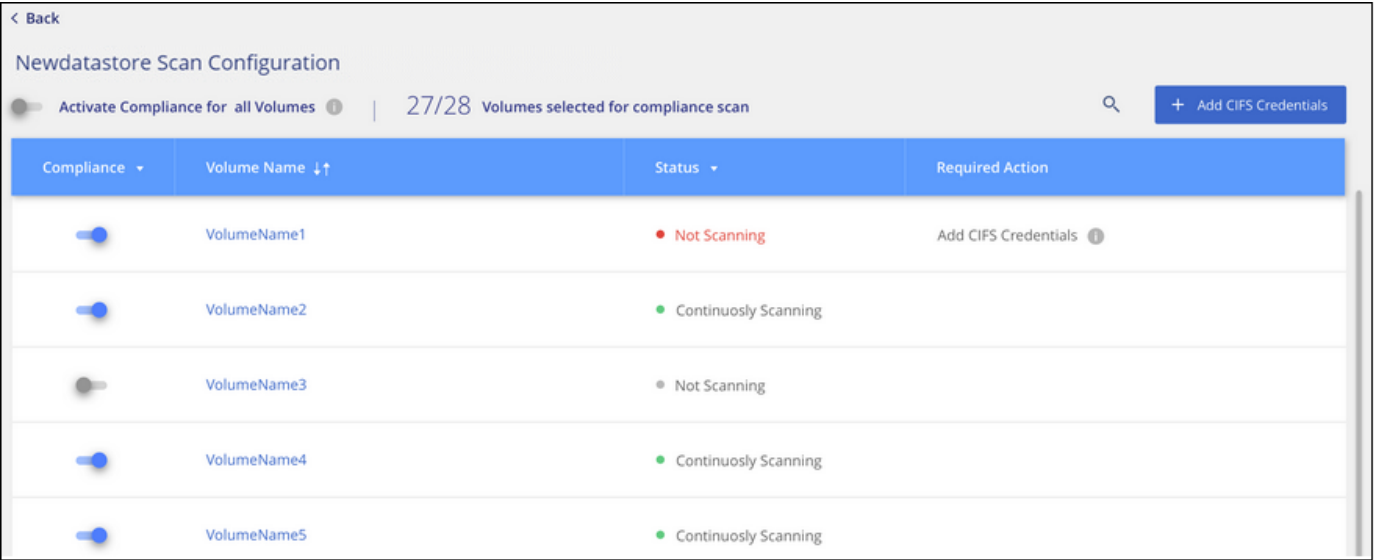
6. On the *Scan Configuration* page, click **View Details** to review the status for each CIFS and NFS volume and correct any errors.

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



Enabling and disabling compliance scans on volumes

You can stop or start scanning volumes in a working environment at any time from the Scan Configuration page. We recommend that you scan all volumes.



To:	Do this:
Disable scanning for a volume	Move the volume slider to the left
Disable scanning for all volumes	Move the <b>Activate Compliance for all Volumes</b> slider to the left
Enable scanning for a volume	Move the volume slider to the right
Enable scanning for all volumes	Move the <b>Activate Compliance for all Volumes</b> slider to the right



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.



## Scanning backup files from on-premises ONTAP systems

If you don't want Cloud Compliance to scan volumes directly on your on-prem ONTAP systems, a new 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 new feature to run compliance scans on those backup files.

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

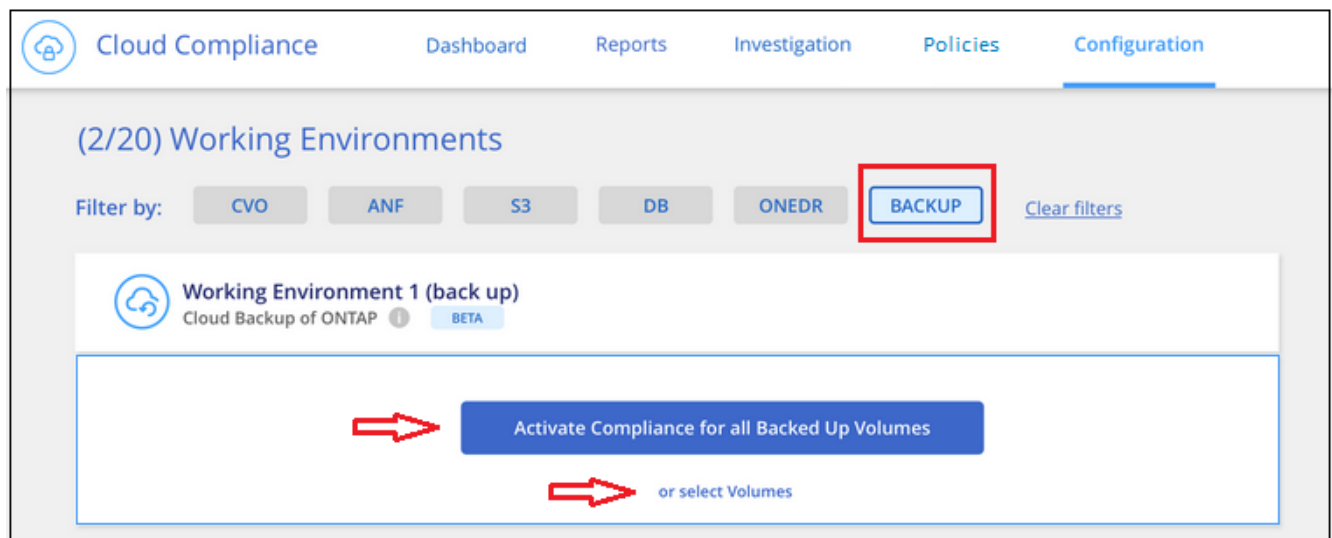
**Note:** When Compliance scans backup files it uses permissions granted through the Restore instance to access the backup files. Typically the Restore instance powers down when not actively restoring files, but it remains on when scanning backup files. See [more information about the Restore instance](#).

### Steps

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

1. At the top of Cloud Manager, click **Compliance** 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.




## Scan Configuration

**Working Environment 1**  
On-Premises ONTAP ⓘ

1

Activate Compliance for All Volumes

or select Volumes

**Working Environment 1 (back up)**  
Cloud Backup of ONTAP ⓘ BETA

2

Activate Compliance for all backed up Volumes

or select Volumes

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

The second item is the backup files 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 Compliance 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**.

### 'Working Environment Name' Scan Configuration

☐ Activate Compliance for all Volumes ⓘ | 22/28 Volumes selected for compliance scan

[Enable Access to DP Volumes](#) [Edit CIFS Credentials](#)

Compliance	Volume Name	Type	Status	Required Action
<input type="checkbox"/>	VolumeName1	DP	● Not Scanning	Enable access to DP Volumes ⓘ
<input checked="" type="checkbox"/>	VolumeName2	NFS	● Continuously Scanning	
<input type="checkbox"/>	VolumeName3	CIFS	● Not Scanning	

## 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 Compliance can scan CIFS volumes you can use those credentials, or you can specify a different set of Admin credentials.

**Provide Active Directory Credentials**

☒ Use existing CIFS Scanning Credentials (user1@domain2) ☐ Use Custom Credentials

Active Directory Domain <sup>1</sup> DNS IP Address <sup>1</sup>

DP Volumes, created from a SnapMirror relationship, do not allow external access by default. Continuing will create NFS shares from DP Volumes which have been activated for Compliance. The shares' export policies will allow access only from the Cloud Compliance instance. [Learn More](#)

**Enable Access to DP Volumes** Cancel

**Provide Active Directory Credentials**

☐ Use existing CIFS Scanning Credentials (user1@domain2) ☒ Use Custom Credentials

Username <sup>1</sup> Password

Active Directory Domain <sup>1</sup> DNS IP Address <sup>1</sup>

DP Volumes, created from a SnapMirror relationship, do not allow external access by default. Continuing will create NFS shares from DP Volumes which have been activated for Compliance. The shares' export policies will allow access only from the Cloud Compliance instance. [Learn More](#)

**Enable Access to DP Volumes** 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 Compliance creates an NFS share from each DP volume that was activated for Compliance so that it can be scanned. The share export policies only allow access from the Cloud Compliance 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 Scan Configuration page. Click this button and add CIFS credentials to enable access to these CIFS DP volumes.

## Getting started with Cloud Compliance for Amazon S3

Cloud Compliance can scan your Amazon S3 buckets to identify the personal and sensitive data that resides in S3 object storage. Cloud Compliance 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 Compliance, including preparing an IAM role and setting up connectivity from Cloud Compliance to S3. [See the complete list.](#)



### Deploy the Cloud Compliance instance

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

### 3

#### Activate Compliance on your S3 working environment

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

### 4

#### Select the buckets to scan

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

#### Reviewing S3 prerequisites

The following requirements are specific to scanning S3 buckets.

#### Set up an IAM role for the Cloud Compliance instance

Cloud Compliance 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 Cloud Compliance on the Amazon S3 working environment.

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

## Provide connectivity from Cloud Compliance to Amazon S3

Cloud Compliance 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 Compliance instance. You must also modify the security group to add an outbound HTTPS rule that enables traffic to the S3 endpoint. Otherwise, Cloud Compliance 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 Compliance instance

[Deploy Cloud Compliance 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 Compliance in an on-premises location is not currently supported when scanning S3 buckets.

## Activating Compliance on your S3 working environment

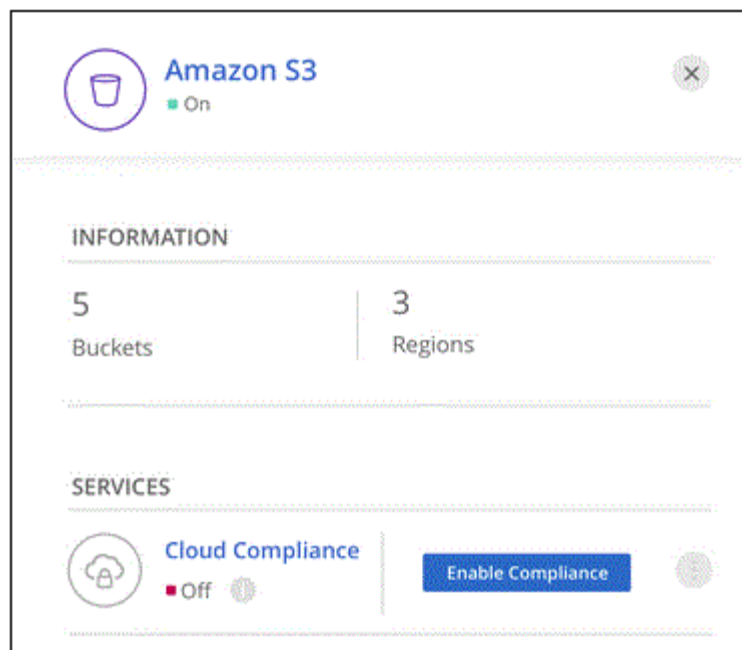
Enable Cloud Compliance on Amazon S3 after you verify the prerequisites.

### Steps

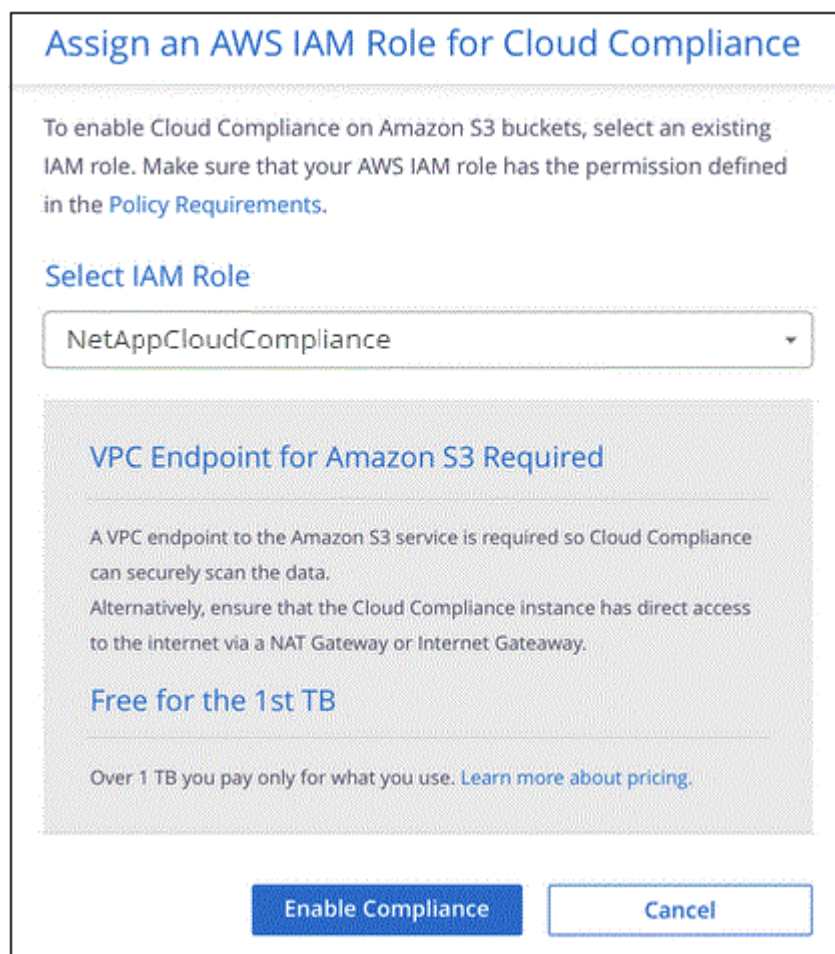
1. At the top of Cloud Manager, click **Canvas**.
2. Select the Amazon S3 working environment.



3. In the pane on the right, click **Enable Compliance**.




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



5. Click **Enable Compliance**.



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

## Result

Cloud Manager assigns the IAM role to the instance.

## Enabling and disabling compliance scans on S3 buckets

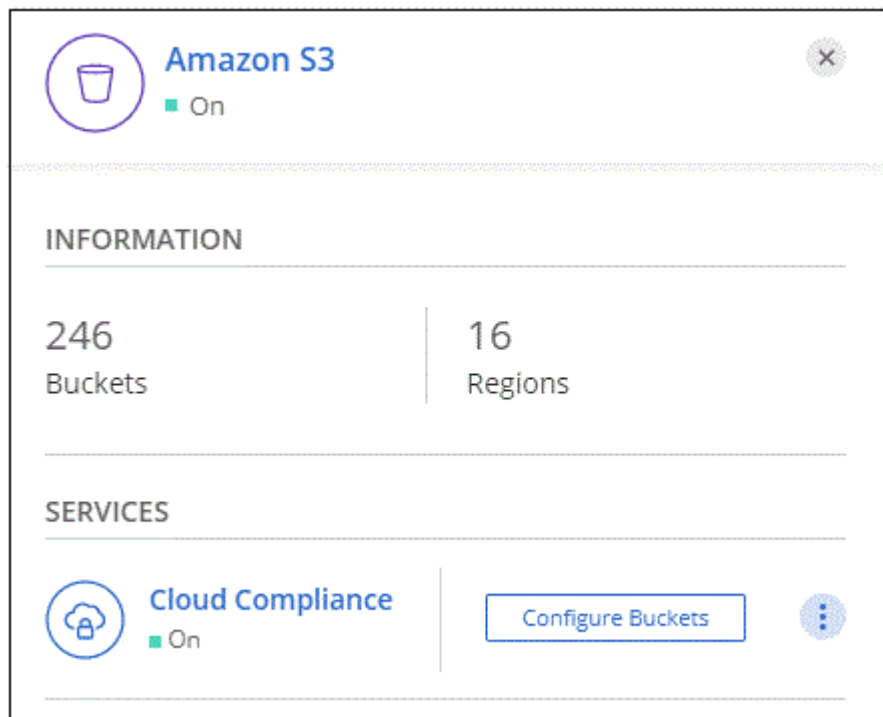
After Cloud Manager enables Cloud Compliance on Amazon S3, the next step is to configure the buckets that you want to scan.

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 Compliance 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 compliance on the buckets that you want to scan.

Cloud Compliance			
<a href="#">&lt; Back</a>			
Amazon S3 Scan Configuration <span style="float: right;">🔍</span>			
15/28 Buckets in Scan Scope. Toggle ON/OFF to enable Compliance per Bucket			
Compliance ▾	Bucket Name ↕	Status ▾	Required Action
<input checked="" type="checkbox"/>	BucketName1	● Not Scanning	Add Credentials
<input checked="" type="checkbox"/>	BucketName2	● Continuously Scanning	
<input type="checkbox"/>	BucketName3	● Not Scanning	

## Result

Cloud Compliance 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 Compliance instance.

## Steps

- 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



### 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 Compliance instance resides.
- Change the **Maximum CLI/API session duration** from 1 hour to 12 hours and save that change.
- Attach the Cloud Compliance IAM policy. Make sure it has the required permissions.

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

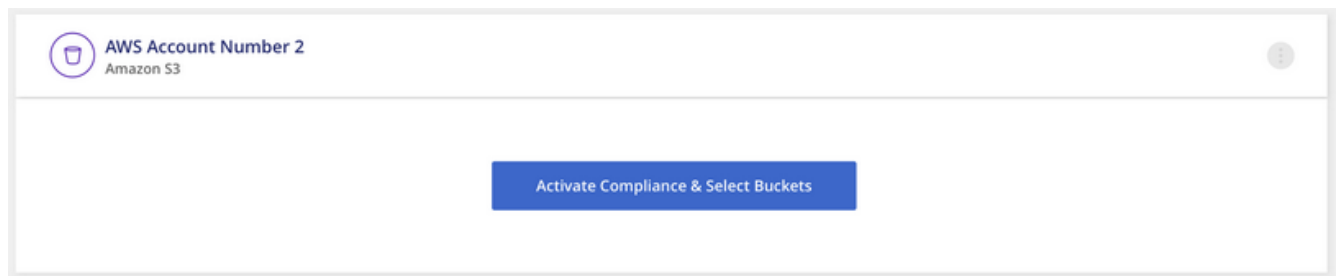
2. Go to the source AWS account where the Cloud Compliance instance resides and select the IAM role that is attached to the instance.
  - a. Change the **Maximum CLI/API session duration** from 1 hour to 12 hours and save that change.
  - b. Click **Attach policies** and then click **Create policy**.
  - 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 Compliance instance profile account now has access to the additional AWS account.

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



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

## Result

Cloud Compliance starts scanning the new S3 buckets that you enabled.

## Scanning database schemas

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

## 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 Compliance instance

[Deploy Cloud Compliance](#) 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 Compliance.

## Supported databases

Cloud Compliance can scan schemas from the following databases:

- 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 Compliance 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 Compliance system with all the required permissions.

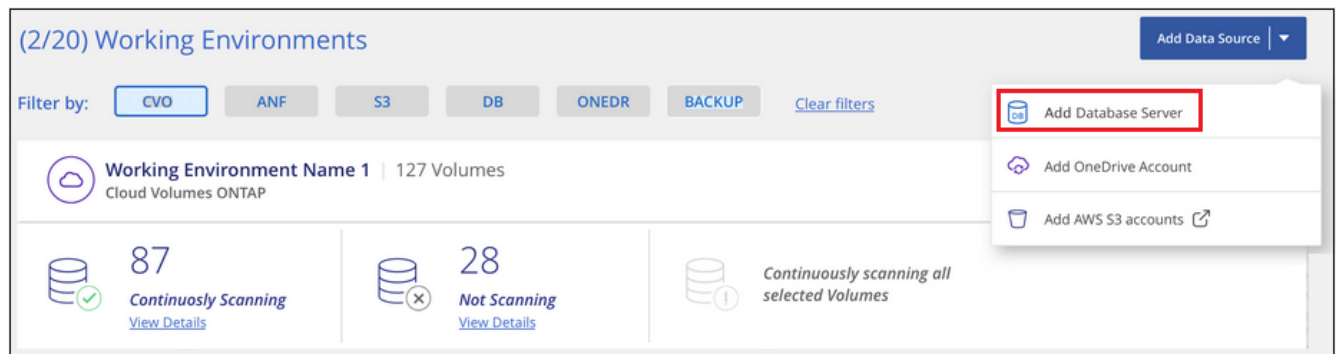
**Note:** For MongoDB, a read-only Admin role is required.

### Adding the database server

You must have [deployed an instance of Cloud Compliance 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 Compliance 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

[Add DB Server](#) [Cancel](#)

The database is added to the list of working environments.

### Enabling and disabling compliance scans on database schemas

You can stop or start scanning schemas at any time.

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

## Scan Configuration

Oracle DB 1 | 41 Schemas

[Configuration](#)

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

## Result

Cloud Compliance 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 Compliance.

### Quick start

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



### Review OneDrive prerequisites

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



### Deploy the Cloud Compliance instance

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



### 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.



### Add the users

Add the list of users from the OneDrive account that you want to scan. 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 Compliance.

- 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.

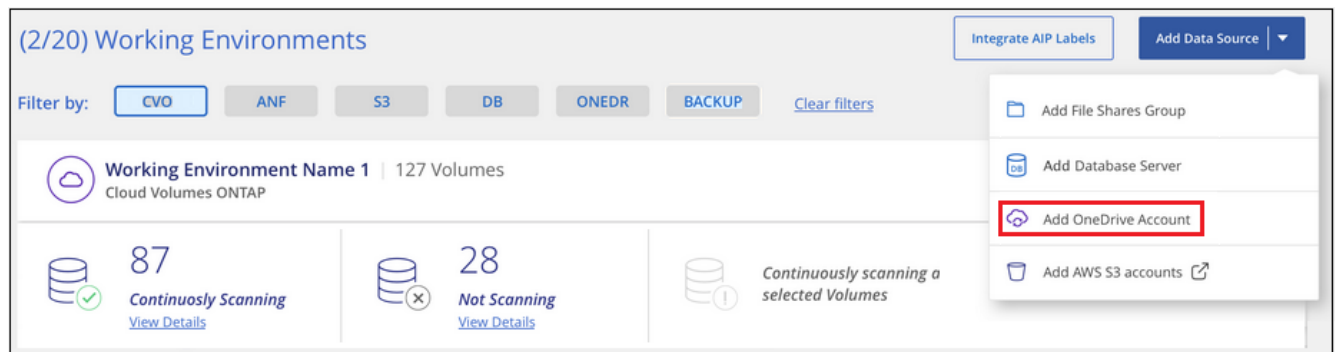
## Adding the OneDrive account

You must have [deployed an instance of Cloud Compliance 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 Compliance to read data from this account.

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

## Adding OneDrive users to compliance scans

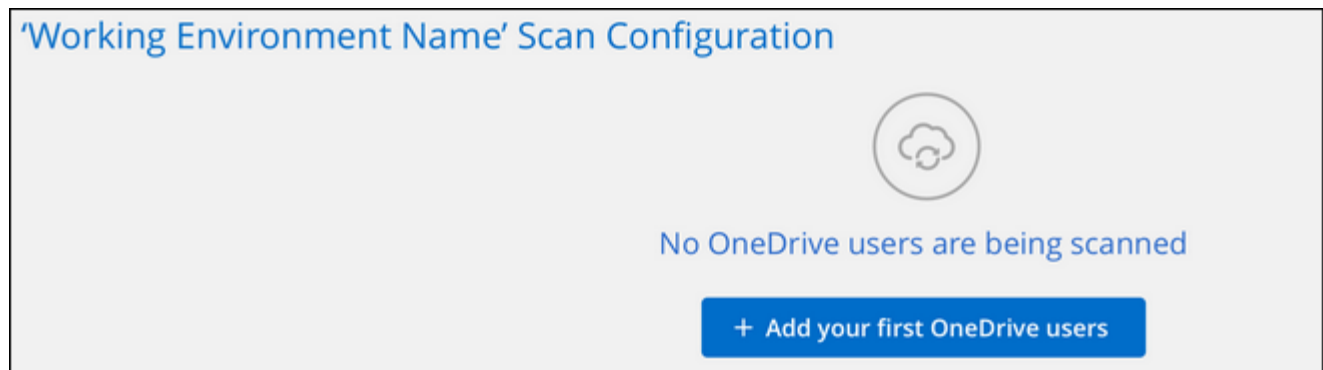
You can add individual OneDrive users, or all of your OneDrive users, so that their files will be scanned by Cloud Compliance.

### Steps

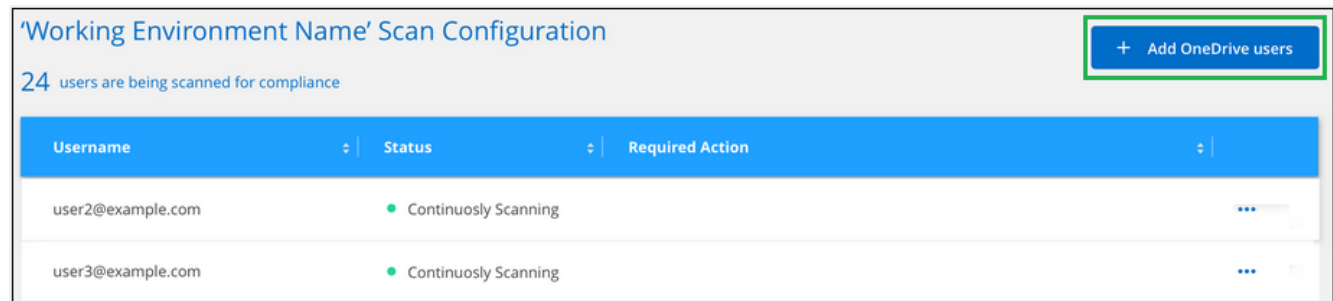
1. From the *Scan Configuration* page, click the **Configuration** button for the OneDrive account.



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 following text: "Provide a list of OneDrive users for Cloud Compliance to scan their data, line-separated. You can add up to 100 users at a time." Below this is a blue link that says "Type or paste below the OneDrive user accounts to add". Underneath the link is a section labeled "User Accounts" which contains a text area with seven lines of placeholder text: "user@example.com". At the bottom of the dialog box, there are two buttons: a blue "Add Users" button and a white "Cancel" button with a blue border.

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

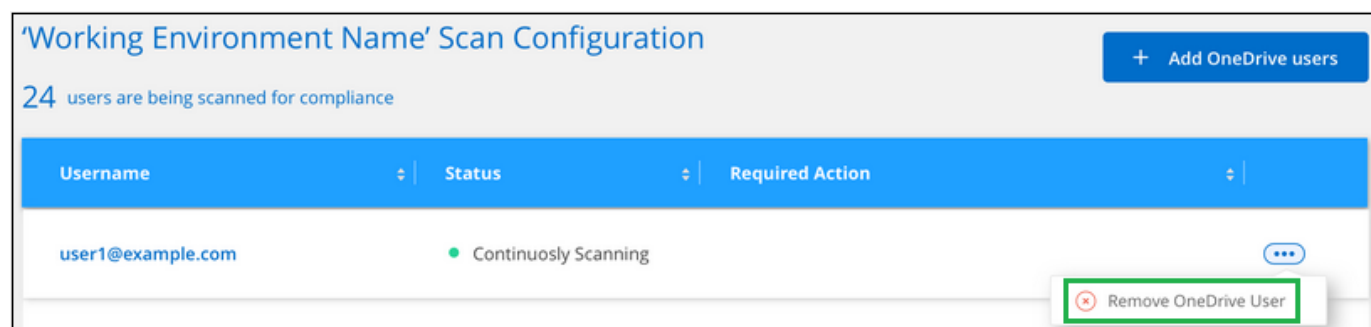
issue. In some cases you can re-add the user with a corrected email address.

## Result

Cloud Compliance 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 Compliance. 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.

#### 1 Review file share prerequisites

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

#### 2 Deploy the Cloud Compliance instance

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

#### 3 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.

#### 4 Add the file shares

Add the list of file shares that you want to scan. 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 Compliance.

- 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 Cloud Compliance instance and the shares.
- Make sure these ports are open to the Cloud Compliance 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 Compliance needs to scan any data that requires elevated permissions.

## Creating the group for the file shares

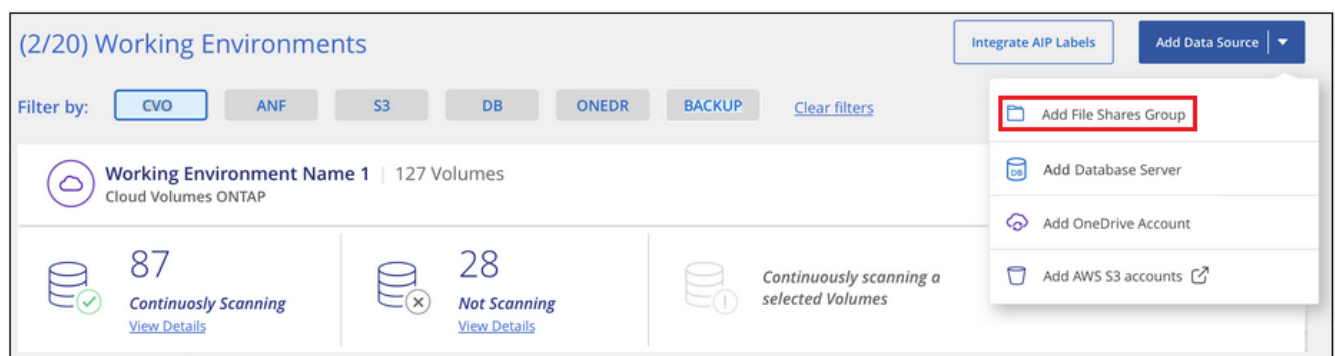
You must have [deployed an instance of Cloud Compliance in Cloud Manager already](#).

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 File 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

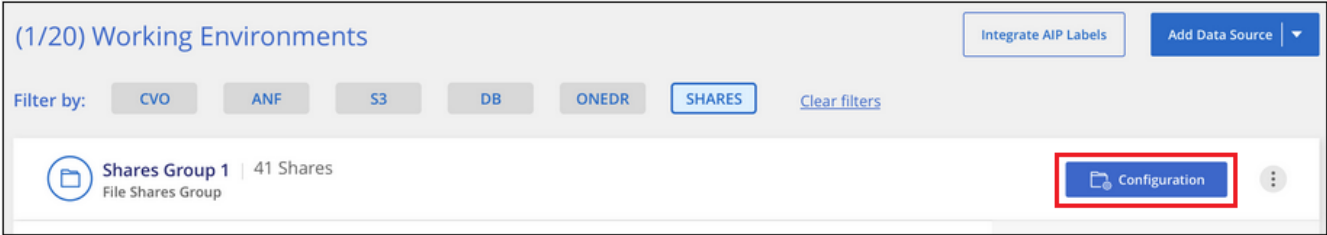
You add file shares to the File Shares Group so that the files in those shares will be scanned by Cloud Compliance. You add the shares in the format `<host_name>:/<share_path>`.

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.

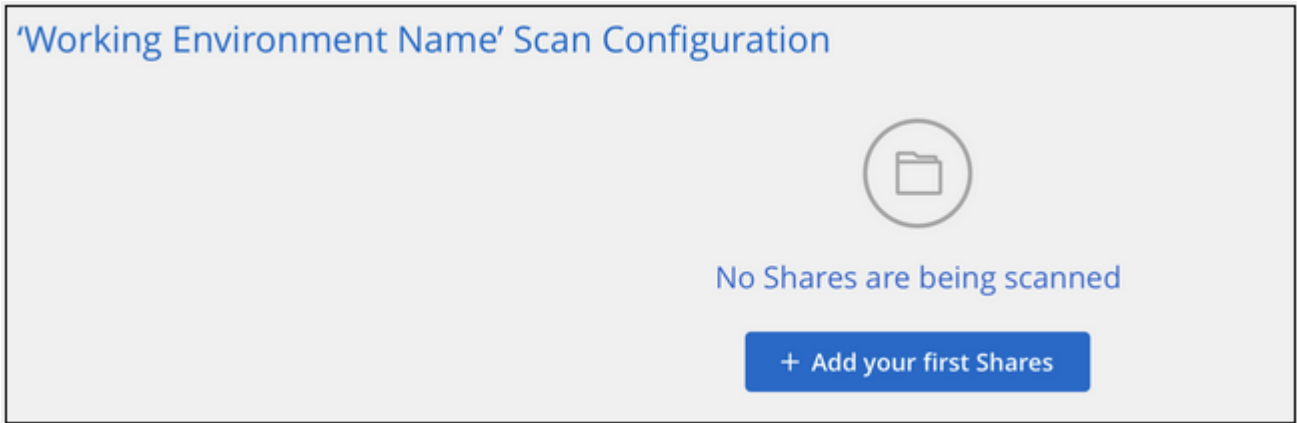
When adding both NFS and CIFS shares in a single group, you'll need to run through the process twice - once adding NFS shares, and then again adding the CIFS shares.

**Steps**

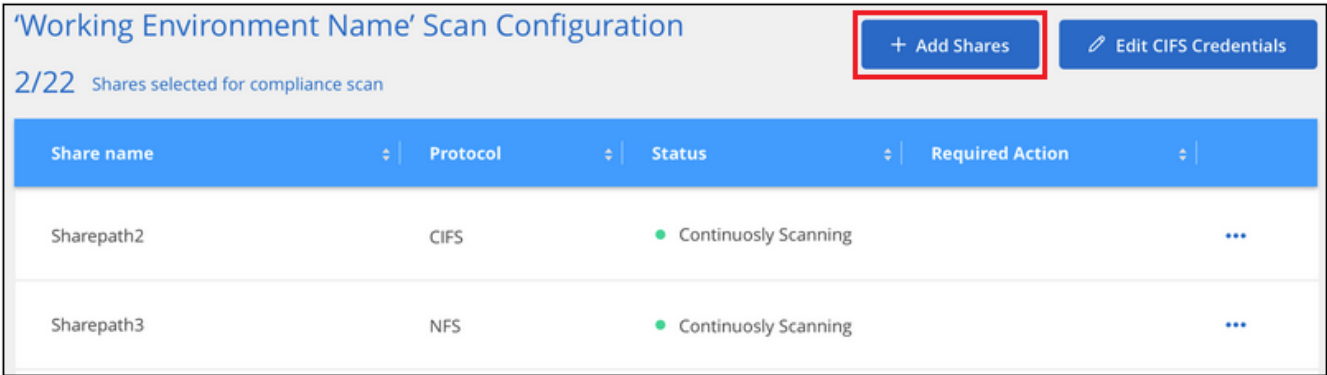
- 1. From the *Working Environments* page, click the **Configuration** button for the File Shares Group.



- 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**.



- 3. 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

☐ 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.

## Result

Cloud Compliance 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.

### 'Working Environment Name' Scan Configuration

2/22 Shares selected for compliance scan

Share name	Protocol	Status	Required Action	
Sharepath1	NFS	● Not Scanning	Add new credentials	<input type="button" value="⋮"/> <div> <input type="button" value="Remove Share"/> </div>

## Scanning object storage that uses S3 protocol

Complete a few steps to start scanning data within object storage directly with Cloud Compliance. Compliance can scan data from any Object Storage service which uses the Simple Storage Service (S3) protocol. This includes NetApp StorageGRID, IBM Cloud

Object Store, 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 Compliance can access the buckets.



#### Deploy the Cloud Compliance instance

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



#### Add the Object Storage Service

Add the object storage service to Cloud Compliance.



#### Select the buckets to scan

Select the buckets that you'd like to scan and Cloud Compliance 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 Compliance.

- 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 Compliance can access the buckets.

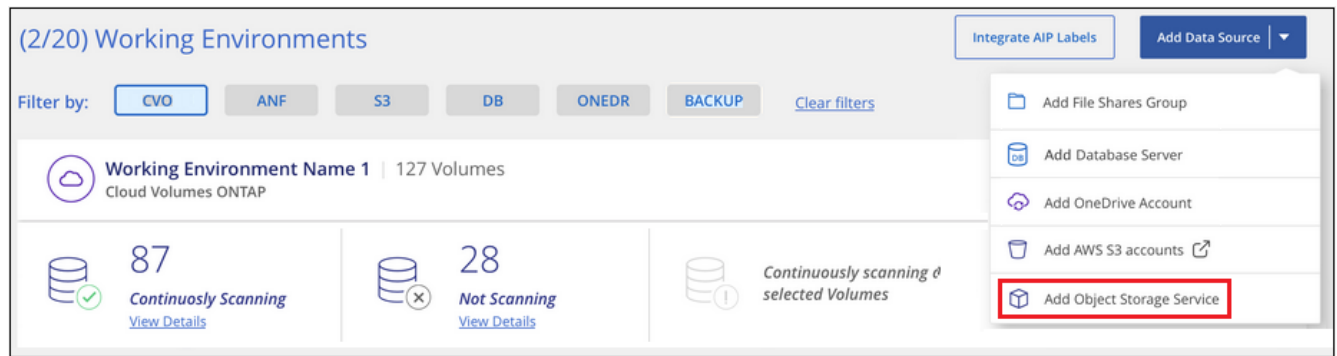
### Adding the object storage service to Cloud Compliance

You must have [deployed an instance of Cloud Compliance 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 Compliance can access the buckets in the object storage.

### Add Object Storage Service

Cloud Compliance can scan data from any Object Storage service. This includes NetApp StorageGRID, Azure Blob, IBM Object Store, MinIO, Linode, B2 Cloud Storage, and more. To continue, enter the following information. In the next step you can select which buckets to scan.

Name the Working Environment	Endpoint URL
<input style="width: 90%;" type="text"/>	<input style="width: 90%;" type="text"/>
Access Key	Secret Key
<input style="width: 90%;" type="text"/>	<input style="width: 90%;" type="text"/>

## Result

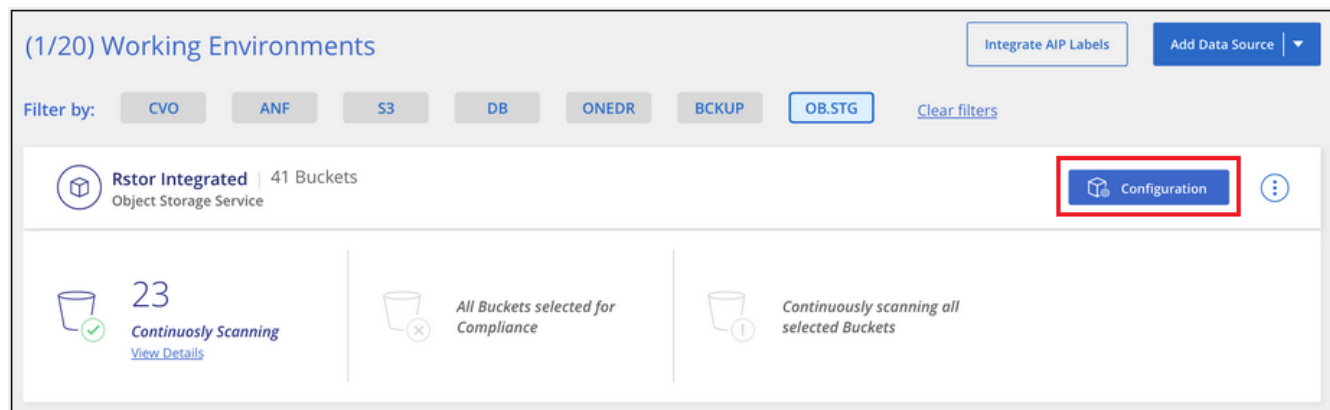
The new Object Storage Service is added to the list of working environments.

## Enabling and disabling compliance scans on object storage buckets

After you enable Cloud Compliance on your Object Storage Service, the next step is to configure the buckets that you want to scan. Cloud Compliance 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.



2. Enable compliance on the buckets that you want to scan.

Rstor Integrated Scan Configuration			
3/55 Buckets selected for Compliance scan			
Compliance ↓↑	Storage Repository (Bucket) ↓↑	Status ↓↑	Required Action ↓↑
<input type="checkbox"/>	logs-759995470648-us-east-1	• Not Scanning	
<input type="checkbox"/>	logs-759995470648-us-west-2	• Not Scanning	
<input checked="" type="checkbox"/>	carstock	• Continuously Scanning	

## Result

Cloud Compliance 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.

## Viewing governance details about the data stored in your organization

Gain control of the costs related to the data on your organizations' storage resources. Cloud Compliance identifies the amount of stale data, non-business data, duplicate files, and very large files in your systems so you can decide whether you want to remove or tier some files to less expensive object storage.

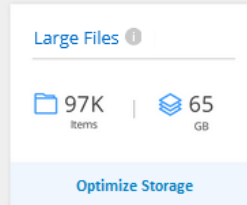
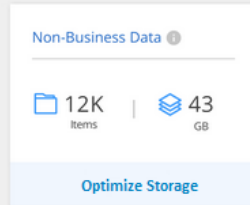
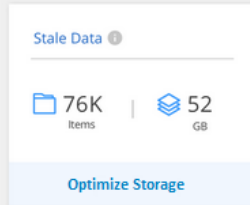
Additionally, if you are planning to migrate data from on-premises locations to the cloud, you can view the size of the data and whether any of the data contains sensitive information prior to moving it.

## The Governance dashboard

The Governance dashboard provides information so that you can increase the efficiency and control the costs related to the data stored on your storage resources.



## Saving Opportunities



## Policies

## Policies

[View All](#)

GDPR - Old Sensitive Data 768K Items

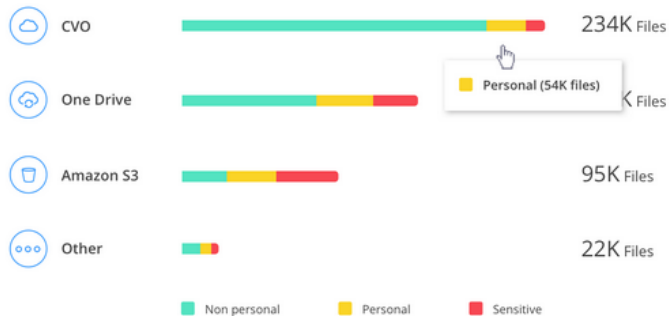
HIPAA - Patients Personal Data 566K Items

## Data Overview

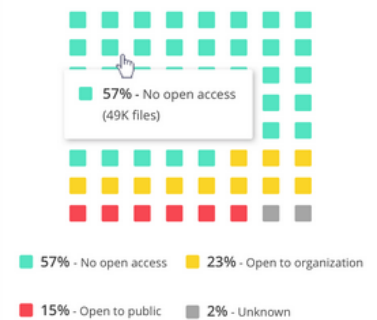
Scanned

27 GB | 547K Files | 8K Tables

## Top Data Repositories by Sensitivity Level



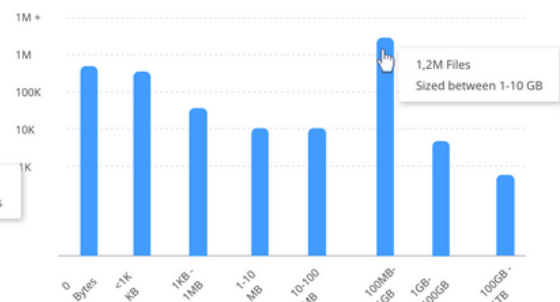
## Open Permissions



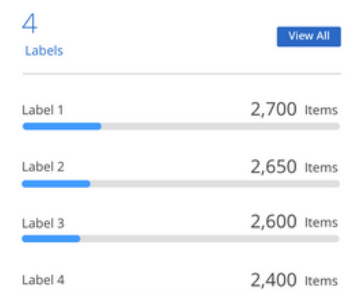
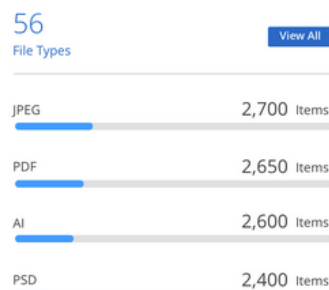
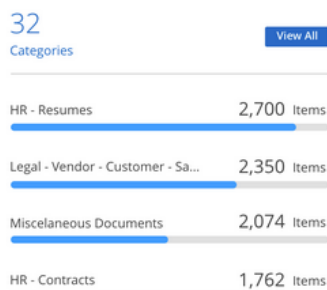
## Age of Data



## Size of Data



## Classification



## Saving Opportunities

You may want to investigate the items in the *Saving Opportunities* area to see if there is any data you should delete or tier to less expensive object storage. Click each item to view the filtered results in the Investigation

page.

- **Stale Data** - Data that was last modified over 3 years ago.
- **Non-Business Data** - Data considered not to be business related, based on their Category or File Type. This includes:
  - Application Data
  - Audio
  - Executables
  - Images
  - Logs
  - Videos
  - Miscellaneous (general "other" category)
- **Large Files** - Files larger than 100 MB.

### Policies with the largest number of results

Click the name of a Policy in the *Policy* area to display the results in the Investigation page. Click **View All** to view the list of all available Policies.

Click [here](#) to learn more about Policies.

### Top data repositories listed by data sensitivity

The *Top Data Repositories by Sensitivity Level* area lists up to the top four data repositories (working environments and data sources) that contain the most sensitive items. The bar chart for each working environment is divided into:

- Non-Personal data
- Personal data
- Sensitive Personal data

You can hover over each section to see the total number of items in each category.

### Data listed by types of Open Permissions

The *Open Permissions* area shows the percentage for each type of permissions that exist for all files that are being scanned. The chart shows the following types of permissions:

- No Open Access
- Open to Organization
- Open to Public
- Unknown Access

You can hover over each section to see the total number of files in each category. Click each area to view the filtered results in the Investigation page so that you can investigate further.



Files in OneDrive accounts and in databases are not represented in this chart.



## Age of Data and Size of Data graphs

You may want to investigate the items in the *Age* and *Size* graphs to see if there is any data you should delete or tier to less expensive object storage.

You can hover over a point in the charts to see details about the age or size of the data in that category. Click to view all the files filtered by that age or size range.

- **Age of Data graph** - Categorizes data based on the last time it was modified.
- **Size of Data graph** - Categorizes data based on size.

## Most identified data Classifications

The *Classification* area provides a list of the most identified [Categories](#), [File types](#), and [AIP Labels](#) in your scanned data.

### Categories

Categories can help you understand what's happening with your data by showing you the types of information that you have. For example, a category like "resumes" or "employee contracts" can include sensitive data. When you investigate the results, you might find that employee contracts are stored in an insecure location. You can then correct that issue.

See [Viewing files by categories](#) for more information.

### File types

Reviewing your file types can help you control your sensitive data because you might find that certain file types are not stored correctly.

See [Viewing file types](#) for more information.

### AIP labels

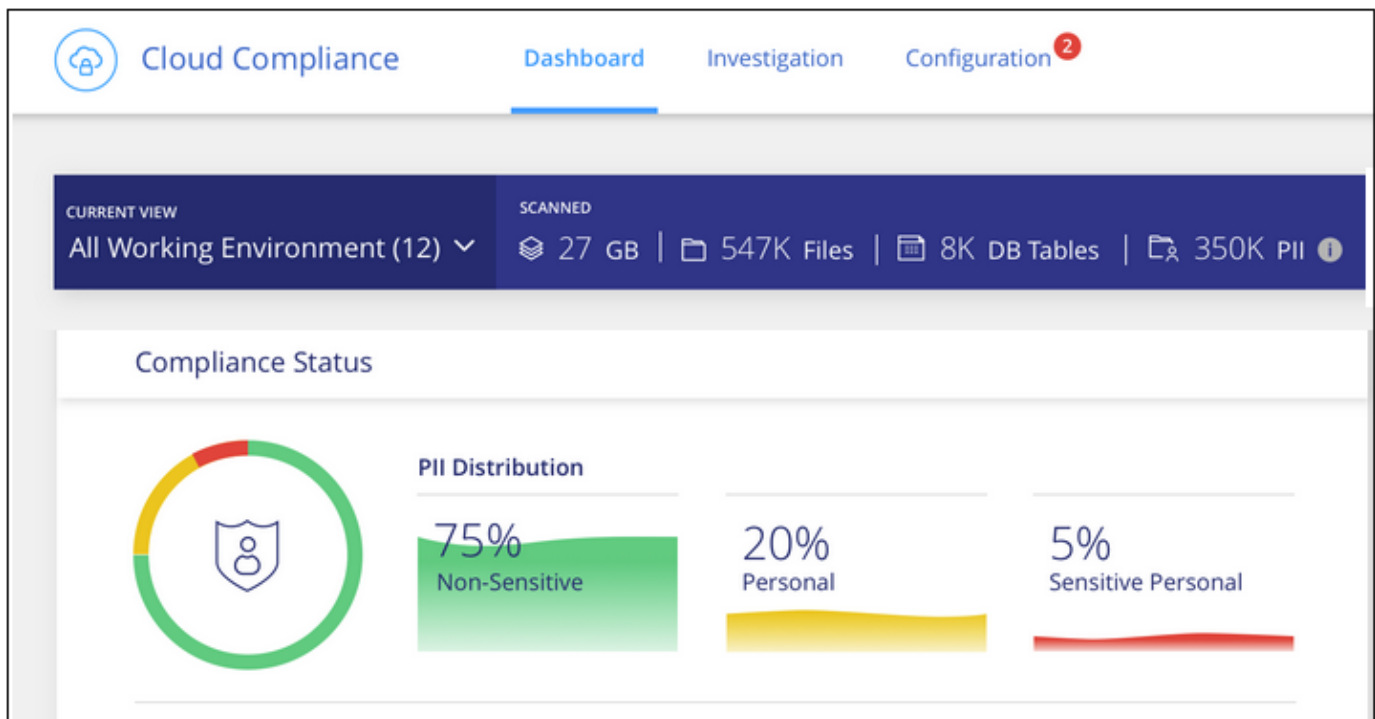
If you have subscribed to Azure Information Protection (AIP), you can classify and protect documents and files by applying labels to content. Reviewing the most used AIP labels that are assigned to files enables you to see which labels are most used in your files.

See [AIP Labels](#) for more information.

## Viewing compliance details about the data stored in your organization

Gain control of your private data by viewing details about the personal data and sensitive personal data in your organization. You can also gain visibility by reviewing the categories and file types that Cloud Compliance found in your data.

By default, the Cloud Compliance dashboard displays compliance data for all working environments and databases.



If you want to see data for only some of the working environments, [select those working environments](#).

You can also filter the results from the Data Investigation page and download a report of the results as a CSV file. See [Filtering data in the Data Investigation page](#) for details.

## Personal data

Cloud Compliance automatically identifies specific words, strings, and patterns (Regex) inside the data. For example, Personal Identification Information (PII), credit card numbers, social security numbers, bank account numbers, and more. [See the full list](#).

Additionally, if you have added a database server to be scanned, the *Data Fusion* feature allows you to scan your files to identify whether unique identifiers from your databases are found in those files or other databases. See [Adding personal data identifiers using Data Fusion](#) for details.

For some types of personal data, Cloud Compliance uses *proximity validation* to validate its findings. The validation occurs by looking for one or more predefined keywords in proximity to the personal data that was found. For example, Cloud Compliance identifies a U.S. social security number (SSN) as a SSN if it sees a proximity word next to it—for example, *SSN* or *social security*. [The table of personal data](#) shows when Cloud Compliance uses proximity validation.

## Viewing files that contain personal data

### Steps

1. At the top of Cloud Manager, click **Compliance** and click the **Dashboard** tab.
2. To investigate the details for all personal data, click the icon next to the personal data percentage.



## Cloud Compliance

### Compliance Status



#### Data Distribution

72%  
Non-Sensitive

18%  
Personal

10%  
Sensitive Personal

3. To investigate the details for a specific type of personal data, click **View All** and then click the **Investigate Results** icon for a specific type of personal data.



## Cloud Compliance

< Back

### Personal Files

7 Types | 41.1K Files

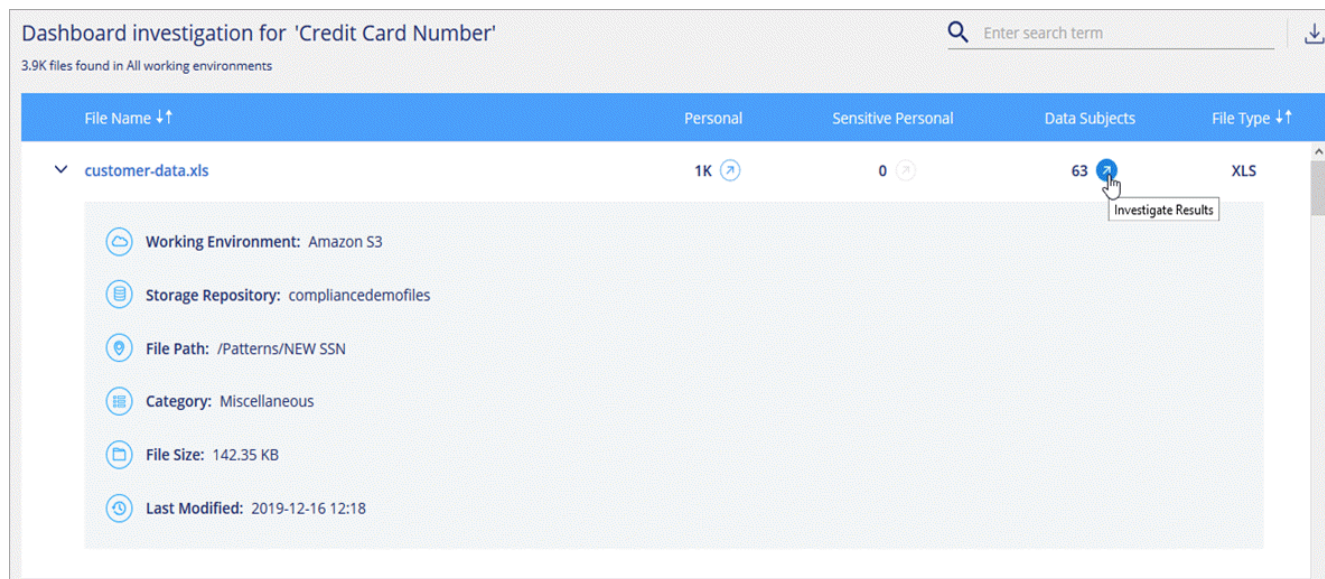
Email Address 40K Files

Credit Card Number 2.3K Files

U.K. National Insurance Number (NINO) 10 Files

Slovenian Tax Identification Number 2 Files

- Investigate the data by searching, sorting, expanding details for a specific file, clicking **Investigate Results** to see masked information, or by downloading the file list.



## Sensitive personal data

Cloud Compliance automatically identifies special types of sensitive personal information, as defined by privacy regulations such as [articles 9 and 10 of the GDPR](#). For example, information regarding a person's health, ethnic origin, or sexual orientation. [See the full list](#).

Cloud Compliance uses artificial intelligence (AI), natural language processing (NLP), machine learning (ML), and cognitive computing (CC) to understand the meaning of the content that it scans in order to extract entities and categorize it accordingly.

For example, one sensitive GDPR data category is ethnic origin. Because of its NLP abilities, Cloud Compliance can distinguish the difference between a sentence that reads "George is Mexican" (indicating sensitive data as specified in article 9 of the GDPR), versus "George is eating Mexican food."



Only English is supported when scanning for sensitive personal data. Support for more languages will be added later.

## Viewing files that contain sensitive personal data

### Steps

- At the top of Cloud Manager, click **Compliance**.
- To investigate the details for all sensitive personal data, click the icon next to the sensitive personal data percentage.



## Cloud Compliance

### Compliance Status



#### Data Distribution

72%  
Non-Sensitive

18%  
Personal

10%  
Sensitive Personal



3. To investigate the details for a specific type of sensitive personal data, click **View All** and then click the **Investigate Results** icon for a specific type of sensitive personal data.



## Cloud Compliance

< Back

### Sensitive Personal Files

7 Types | 23.4K Files

Religious Beliefs Reference

14K Files



Criminal Procedures Reference

12K Files



ICD-10-CM Medical Code

7.9K Files



Sex Life or Orientation Reference

3.9K Files



4. Investigate the data by searching, sorting, expanding details for a specific file, clicking **Investigate Results** to see masked information, or by downloading the file list.

## Categories

Cloud Compliance takes the data that it scanned and divides it into different types of categories. Categories are topics based on AI analysis of the content and metadata of each file. [See the list of categories.](#)

Categories can help you understand what's happening with your data by showing you the types of information that you have. For example, a category like resumes or employee contracts can include sensitive data. When you investigate the results, you might find that employee contracts are stored in an insecure location. You can then correct that issue.

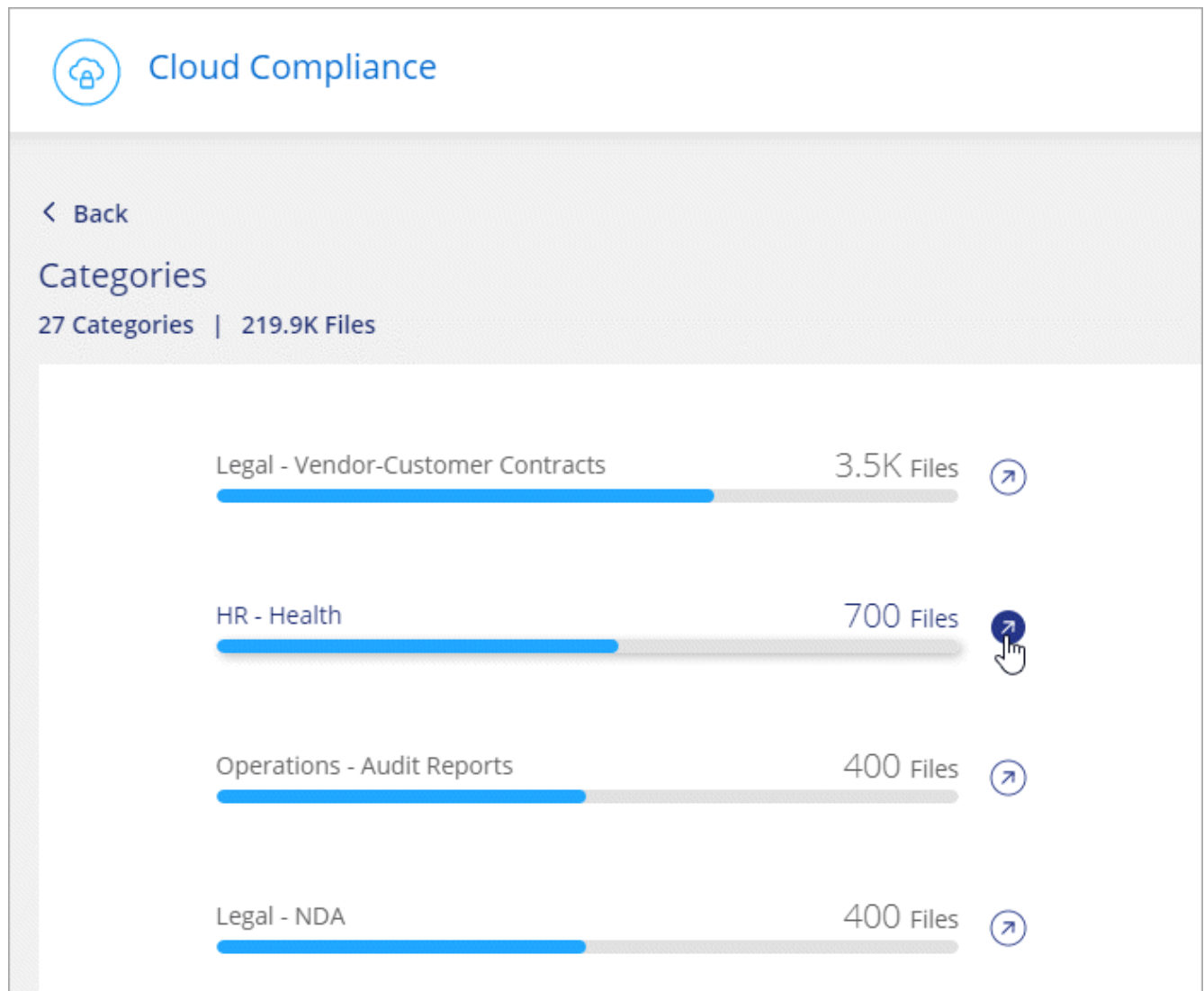


Only English is supported for categories. Support for more languages will be added later.

### Viewing files by categories

#### Steps

1. At the top of Cloud Manager, click **Compliance**.
2. Click the **Investigate Results** icon for one of the top 4 categories directly from the main screen, or click **View All** and then click the icon for any of the categories.



- Investigate the data by searching, sorting, expanding details for a specific file, clicking **Investigate Results** to see masked information, or by downloading the file list.

## File types

Cloud Compliance takes the data that it scanned and breaks it down by file type. Reviewing your file types can help you control your sensitive data because you might find that certain file types are not stored correctly. [See the list of file types](#).

For example, you might be storing CAD files that include very sensitive information about your organization. If they are unsecured, you can take control of the sensitive data by restricting permissions or moving the files to another location.

### Viewing file types

#### Steps

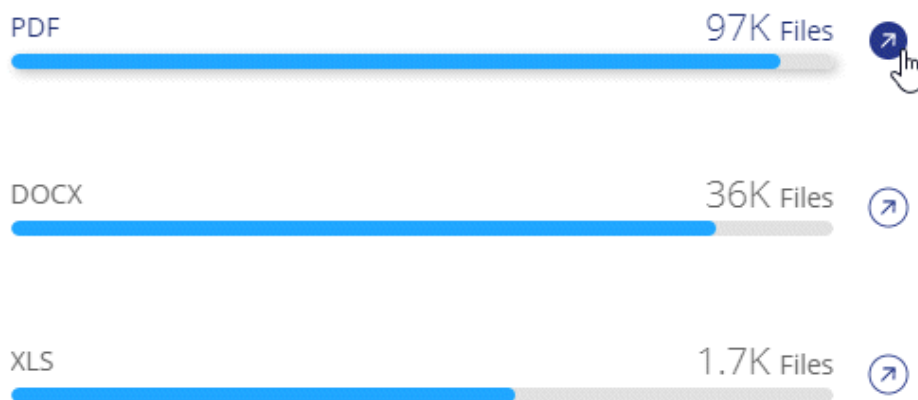
- At the top of Cloud Manager, click **Compliance**.
- Click the **Investigate Results** icon for one of the top 4 file types directly from the main screen, or click **View All** and then click the icon for any of the file types.



< Back

## File Types

36 File Types | 219.9K Files



- Investigate the data by searching, sorting, expanding details for a specific file, clicking **Investigate Results** to see masked information, or by downloading the file list.

### Viewing file metadata

In the Data Investigation results pane you can click  for any single file to view the file metadata.



Unstructured (32K Files)

Structured (323 DB Tables)

File Name	Personal	Sensitive Personal	Data Subjects	File Type	
> Expense Report EXP-TPO-106038887654	cvo	6	3	16	PDF
✓ Expense Report EXP-TPO-106038887654 <div> <div>cvo</div> <div>6</div> <div>3</div> <div>16</div> <div>PDF</div> </div>					

Working Environment: WorkingEnvironment1

Repository: Volume Name

File Path: /Prod/Expense Report EXP-TPO-1060388.pdf

Category: Legal

File Size: 22 MB

Last Modified: 2019-08-06 07:51

Open Permissions: NO OPEN PERMISSIONS [View all Permissions](#)

File Owner: Asaf Ley

Duplicates: 3 [View Details](#)

Assign a Label to this file

Delete this file

Give feedback on this result

In addition to showing you the working environment and volume where the file resides, the metadata shows much more information, including the file permissions, file owner, whether there are duplicates of this file, and assigned AIP label (if you have [integrated AIP in Cloud Compliance](#)). This information is useful if you're planning to [create Policies](#) because you can see all the information that you can use to filter your data.

Note that not all information is available for all data sources - just what is appropriate for that data source. For example, volume name, permissions, and AIP labels are not relevant for database files.

When viewing the details for a single file there are two actions you can take on the file:

- You can delete the file. See [Deleting source files](#) for details.
- If you have integrated AIP labels with Cloud Compliance, you can assign a label to this file, or change to a different label if one already exists. See [Assigning AIP labels manually](#) for details.

## Viewing permissions for files

To view a list of all users or groups who have access to a file, and the types of permissions they have, click **View all Permissions**.

File Name
Personal
Sensitive Personal
Data Subjects
File Type

Expense Report TPO-1060.pdf
cvo
6
3
16
PDF

Working Environment: WorkingEnvironment1
Repository: Volume Name
File Path: /Prod/labs-base/Expense Report TPO-1060.pdf
Category: Legal
File Size: 22 MB
Last Modified: 2019-08-06 07:51
Open Permissions: NO OPEN PERMISSIONS
File Owner: Avy

Assign a Label to this file
Delete this file

Permissions list for file Expense Report TPO-1060.pdf

Group or User	Read	Write
user1@company.com	✓	✗
user2@company.com	✓	✓
dist_list_IT@company.com	✓	✗
user4@company.com	✓	✓

Close

This button is available only for files in CIFS shares.

## Viewing whether files are duplicated in your storage systems

If there are duplicates of a certain file, this information appears next to the *Duplicates* field. This is useful if you want to identify areas where you can save storage space. It can also be helpful to make sure certain files that have specific permissions or sensitive information are not unnecessarily duplicated in your storage systems.

Only files that are in the formats described in [Types of files](#) can be identified as duplicate.

To view the list of duplicate files, click **View Details**. In the next page click **View Duplicates** to view the files in the Investigation page.

🕒 Last Modified: 2019-08-06 07:51

🔑 Open Permissions: NO OPEN PERMISSIONS [View all Permissions](#)

👤 File Owner: Asaf Ley

📄 Duplicates: 3 [View Details](#)

Duplicates of File 'Name 1'

📄 Duplicates: 3

📦 Total Size of all Duplicates: 1GB

🔍 File Hash: xxxxxxx

[View Duplicates](#)

[Close](#)

3 items

<input type="checkbox"/>	File Name		Personal	Sensitive Personal	Data Subjects	File Type	
<input type="checkbox"/>	Expense Report EXP-TPO-106038887654	cvo	6	3	16	PDF	▼
<input type="checkbox"/>	Expense Report EXP-TPO-106038887654	cvo	6	3	16	PDF	▼
<input type="checkbox"/>	Expense Report EXP-TPO-106038887654	cvo	6	3	16	PDF	▼

You can download the list of duplicate files and send it to your storage admin so they can decide which files, if any, can be deleted. Of you can [delete the file](#) yourself if you are confident that this version of the file is not needed.



You can use the "hash" value provided in this page and enter it directly in the Investigation page to search for a specific duplicate file at any time - or to be used in a Policy.

## Viewing Dashboard data for specific working environments

You can filter the contents of the Cloud Compliance dashboard to see compliance data for all working environments and databases, or for just specific working environments.

When you filter the dashboard, Cloud Compliance scopes the compliance data and reports to just those working environments that you selected.

### Steps

1. Click the filter drop-down, select the working environments that you'd like to view data for, and click **View**.



All Working Environments (12) ^

☒ Select all

☒ ANF - Azure NetApp Files

ANF

☒ Working Environment Name 1

CVO

☒ Working Environment Name 2

CVS

☒ Working Environment Name 3

CVS

☒ Working Environment Name 4

CVO

View

Cancel

Personal Files ⓘ

View All

Email Address 2,700 Files



Credit Card 2,700 Files



20%  
Personal



5%  
Sensitive Personal



7,000

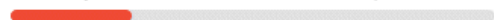
Sensitive Personal Files ⓘ

View All


Health 2,700 Files



Ethnicity 2,700 Files




## Filtering data in the Data Investigation page

You can filter the contents of the investigation page to display only the results you want to see. If you want to save a CSV version of the content as a report after you have refined it, click the  button.

Data Investigation		Unstructured (32K Files)		Structured (323 DB Tables)			
FILTERS		File Name	Personal	Sensitive Personal	Data Subjects	File Type	
Policies		> Expense Report EXP-TPO-10603888765435	cvo	6	3	16	PDF
Working Environment		> Expense Report EXP-TPO-10603888765435	cvo	6	3	16	PDF
Storage Repository		> Expense Report EXP-TPO-10603888765435	cvo	6	3	16	PDF
Category		> Expense Report EXP-TPO-10603888765435	cvo	6	3	16	PDF
Private Data		> Expense Report EXP-TPO-10603888765435	cvo	6	3	16	PDF
File Type		> Expense Report EXP-TPO-10603888765435	cvo	6	3	16	PDF

- The top-level tabs allow you to view data from files (unstructured data) or from databases (structured data).
- The controls at the top of each column allow you to sort the results in numerical or alphabetical order.
- The left-pane filters enable you to refine the results by working environment, storage repository, category, private data, file type, file size, last modified date, whether the S3 object's permissions are open to public access, etc...
- The *Policies* filter at the top of the Filters pane lists the custom filters that provide commonly requested combinations of filters; like a saved database query or Favorites list. Go [here](#) to view the list of predefined Policies and to see how you can create your own custom Policies.

## What's included in each file list report (CSV file)

From each Investigation page you can click the  button to download file lists (in CSV format) that include details about the identified files. If there are more than 10,000 results, only the top 10,000 appear in the list.

Each file list includes the following information:

- File name
- Location type
- Working environment
- Storage repository
- Protocol
- File path
- File type
- File size
- File owner
- Category
- Personal information
- Sensitive personal information
- Deletion detection date

A deletion detection date identifies the date that the file was deleted or moved. This enables you to identify when sensitive files have been moved. Deleted files aren't part of the file number count that appears in the dashboard or on the Investigation page. The files only appear in the CSV reports.

## Managing your private data

Cloud Compliance provides many ways for you to manage your private data. Some functionality just makes it easier to see the data that is most important to you, and other functionality allows you to make changes to the data.

- Using the "Policy" functionality you can create your own custom search queries so that you can easily see the results by clicking one button.
- You can send email alerts to Cloud Manager users when certain critical Policies return results.
- If you are subscribed to [Azure Information Protection \(AIP\)](#) to classify and protect your files, you can use Cloud Compliance to manage those AIP labels.
- You can delete files that seem insecure or too risky to leave in your storage system, or that you have identified as a duplicate.

See below for more functionality that is provided with both the Policies and AIP features.

## Controlling your data using Policies

Policies are like a favorites list of custom filters that provide search results in the Investigation page for commonly requested compliance queries. Cloud Compliance provides a set of predefined Policies based on common customer requests. You can create custom Policies that provide results for searches specific to your organization.

Policies provide the following functionality:

- [Predefined Policies](#) from NetApp based on user requests
- Ability to create your own custom Policies
- Launch the Investigation page with the results from your Policies in one click
- Send email alerts to Cloud Manager users when certain critical Policies return results so you can get notifications to protect your data
- Assign AIP (Azure Information Protection) labels automatically to all files that match the criteria defined in a Policy

The **Policies** tab in the Compliance Dashboard lists all the predefined and custom Policies available on this instance of Cloud Compliance.

Cloud Compliance | Dashboard | Reports | Investigation | **Policies** | Configuration<sup>2</sup>

### Policies List

**GDPR - Old Sensitive Data**  
Predefined Policy ⓘ

Email notifications: **ON** **Edit** ⓘ

Data with European IDs for GDPR received from XDR database, sharing with Legal area in charged of Jon Doe. Also, this lines can take up to 2 lines, we need to limit character input so the description does not take more than 2 lines here in the component.

**HIPAA - Patients Personal Data**  
Last modified: 17-10-20

Email notifications: **OFF** **Edit** ⓘ

The Health Insurance Portability and Accountability Act of 1996 (HIPAA) is a federal law that required the creation of national standards to protect sensitive patient health information from being disclosed without the patient's consent or knowledge.

In addition, Policies appear in the list of Filters in the Investigation page.

### Viewing Policy results in the Investigation page

To display the results for a Policy in the Investigation page, click the ⓘ button for a specific Policy, and then select **Investigate Results**.

Cloud Compliance | Dashboard | Reports | Investigation | **Policies** | Configuration<sup>2</sup>

### Policies List

**GDPR - Old Sensitive Data**  
Last modified: 30-09-20

Email notifications: **ON** **Edit** ⓘ

Data with European IDs for GDPR received from XDR database, sharing with Legal area in charged of Jon Doe. Also, this lines can take up to 2 lines, we character input so the description does not take more than 2 lines here in the component.

ⓘ

Investigate Results

Delete Policy

### Creating custom Policies

You can create your own custom Policies that provide results for searches specific to your organization.

#### Steps

1. From the Data Investigation page, define your search by selecting all the filters you want to use. See [Filtering data in the Data Investigation page](#) for details.
2. Once you have all the filter characteristics just the way you want them, click **Save this search as a Policy**.

**Data Investigation**

**FILTERS**

Clear All

Search filters

×

Policies

+

Working Environment

4

+

Storage Repository

+

Category

+

Private Data

6

+

File Type

+

Save this search as a Policy

3. Name the Policy and select other actions that can be performed by the Policy:
  - a. Enter a unique name and description.
  - b. Optionally, check the box if you want notification emails sent to Cloud Manager users, and choose the interval at which the email is sent.
  - c. Optionally, check the box to automatically assign AIP labels to files that match the Policy parameters, and select the label. (Learn more about [AIP labels](#).)
  - d. Click **Create Policy**.

Create Policy

Saving this filtered view will create a new Policy, you can view/edit it in the "Policy" tab

Name this Policy
 

HIPAA - Patient Personal Data

Give it a description to quickly identify it
 

Files containing patient health information that is more than 30 days old

☒ Send email updates about this Policy to Cloud Manager users on this account every 

Week

☐ Automatically label matches of this Policy with: 

select label

Create Policy

Cancel



## Result

The new Policy appears in the Policies tab.

## Editing Policies

You can modify certain parts of a Policy depending on the type of Policy:

- Custom Policies - You can modify the *Name*, the *Description*, whether email notifications are sent, and whether AIP labels are added.
- Predefined Policies - You can modify only whether email notifications are sent and whether AIP labels are added.




If you need to change the filter parameters for a custom Policy, you'll need to create a new Policy with the parameters you want, and then delete the old Policy.

To modify a Policy, click the **Edit** button, enter your changes on the *Edit Policy* page, and click **Save Policy**.

## Deleting Policies

You can delete any custom Policy that you created if you no longer need it. You can't delete any of the predefined Policies.

To delete a Policy, click the  button for a specific Policy, click **Delete Policy**, and then click **Delete Policy** again in the confirmation dialog.

## Categorizing your data using AIP labels

You can manage AIP labels in the files that Cloud Compliance is scanning if you have subscribed to [Azure Information Protection \(AIP\)](#). AIP enables you to classify and protect documents and files by applying labels to content. Cloud Compliance enables you to view the labels that are already assigned to files, add labels to files, and change labels when a label already exists.

Cloud Compliance support AIP labels within the following file types: .PDF, .DOCX, .DOC, .PPTX, .XLS, .XLSX.

Note that you can't currently change labels in files larger than 30 MB. For OneDrive accounts the maximum file size is 4 MB.



If a file has a label which doesn't exist anymore in AIP, Cloud Compliance considers it as a file without a label.

## Integrating AIP labels in your workspace

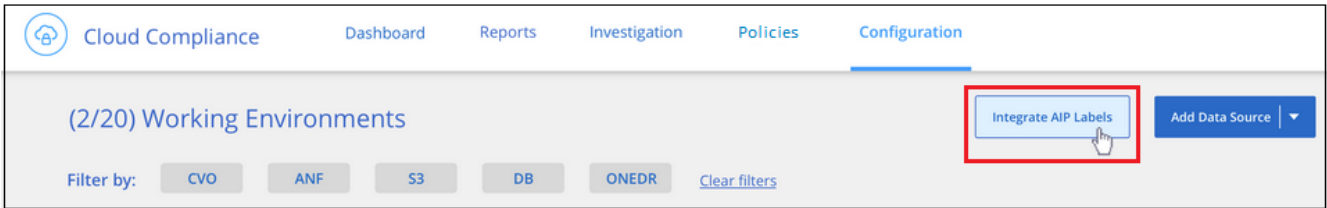
Before you can manage AIP labels, you need to integrate the AIP label functionality into Cloud Compliance by signing into your existing Azure account. Once enabled, you can manage AIP labels within files for all [working environments and data sources](#) in your Cloud Manager workspace.

## Requirements

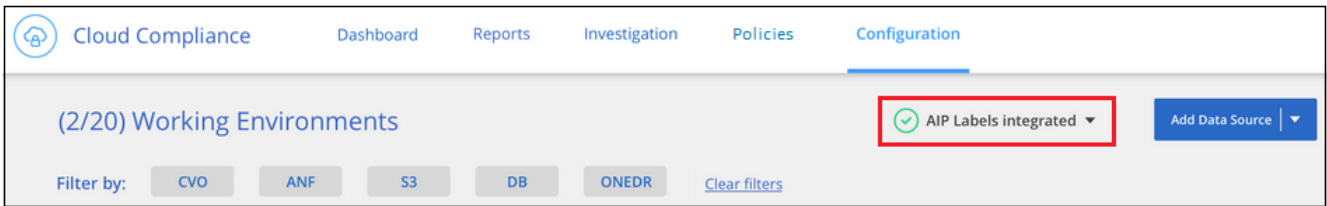
- You must have an account and an Azure Information Protection license.
- You must have the login credentials for the Azure account.
- If you plan to change labels in files that reside in Amazon S3 buckets, ensure that the permission `s3:PutObject` is included in the IAM role. See [setting up the IAM role](#).

Steps

- 1. From the Cloud Compliance Configuration page, click **Integrate AIP Labels**.



- 2. In the Integrate AIP Labels dialog, click **Sign in to Azure**.
- 3. In the Microsoft page that appears, select the account and enter the required credentials.
- 4. Return to the Cloud Compliance tab and you'll see the message "AIP Labels were integrated successfully with the account <account\_name>".
- 5. Click **Close** and you'll see the text *AIP Labels integrated* at the top of the page.



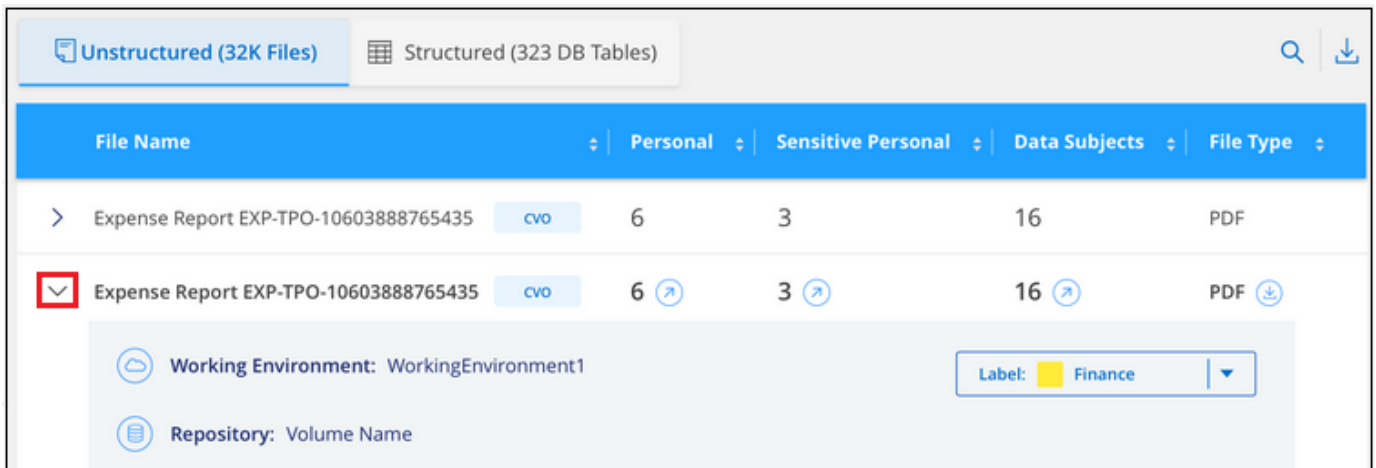
Result

You can view and assign AIP labels from the results pane of the Investigation page. You can also assign AIP labels to files using Policies.

Viewing AIP labels in your files

You can view the current AIP label that is assigned to a file.

In the Data Investigation results pane, click  for the file to expand the file metadata details.



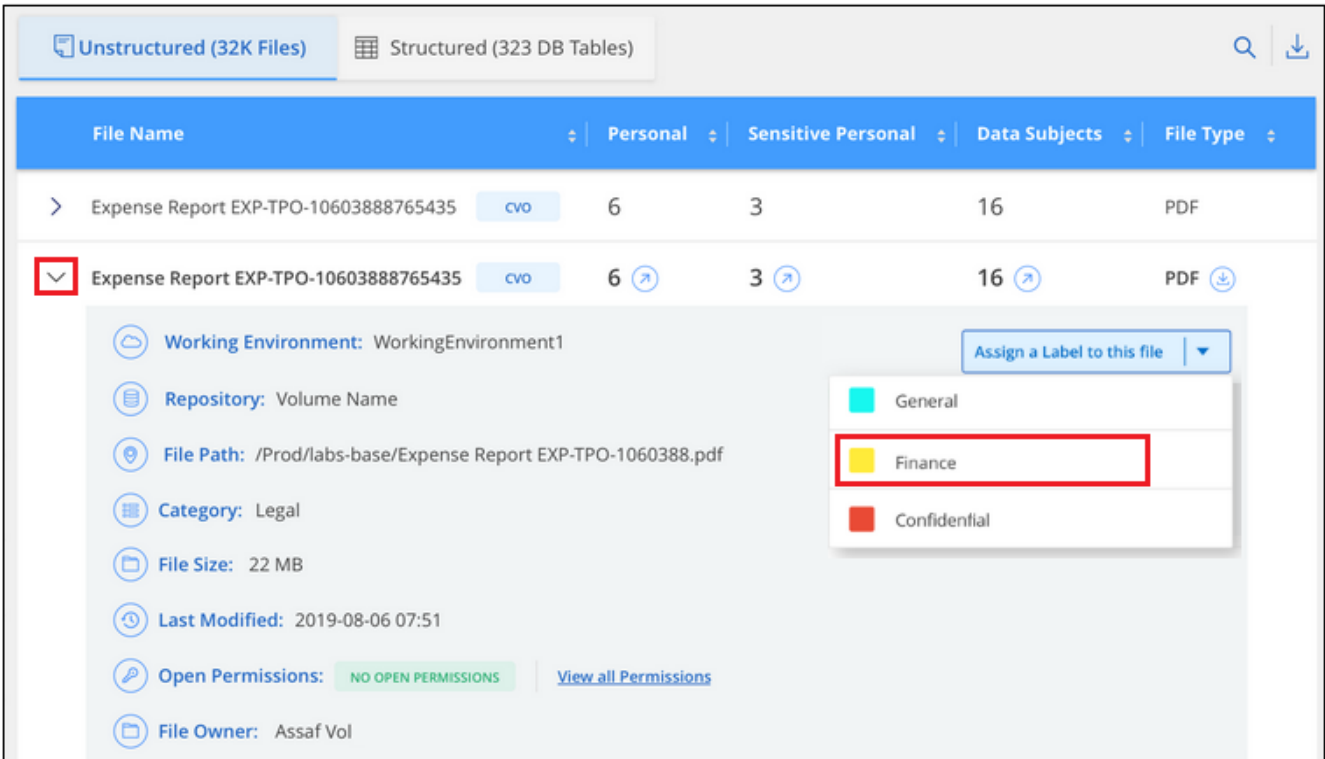
Assigning AIP labels manually

You can add, change, and remove AIP labels from your files using Cloud Compliance.

Follow these steps to assign an AIP label to a single file.

Steps

- 1. In the Data Investigation results pane, click  for the file to expand the file metadata details.



- 2. Click **Assign a Label to this file** and then select the label.

The label appears in the file metadata.

Assigning AIP labels automatically with Policies

You can assign an AIP label to all the files that meet the criteria of the Policy. You can specify the AIP label when creating the Policy, or you can add the label when editing any Policy.

Labels are added or updated in files continuously as Cloud Compliance scans your files.

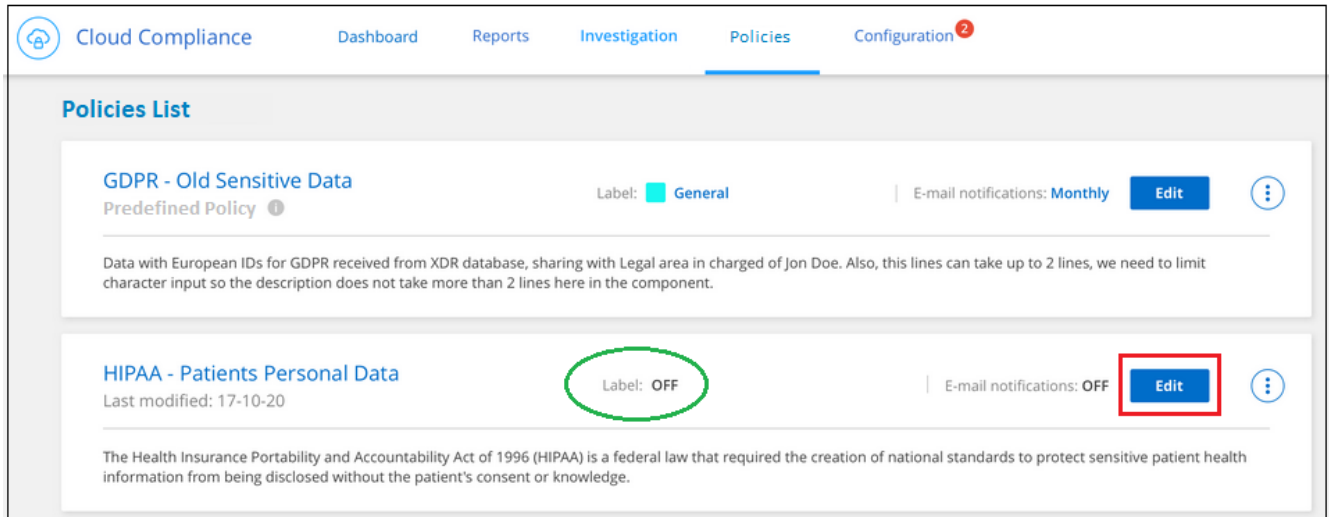
Depending on whether a label is already applied to a file, and the classification level of the label, the following actions are taken when changing a label:

If the file...	Then...
Has no label	The label is added
Has an existing label of a lower level of classification	The higher level label is added
Has an existing label of a higher level of classification	The higher level label is retained
Is assigned a label both manually and by a Policy	The higher level label is added
Is assigned two different labels by two Policies	The higher level label is added

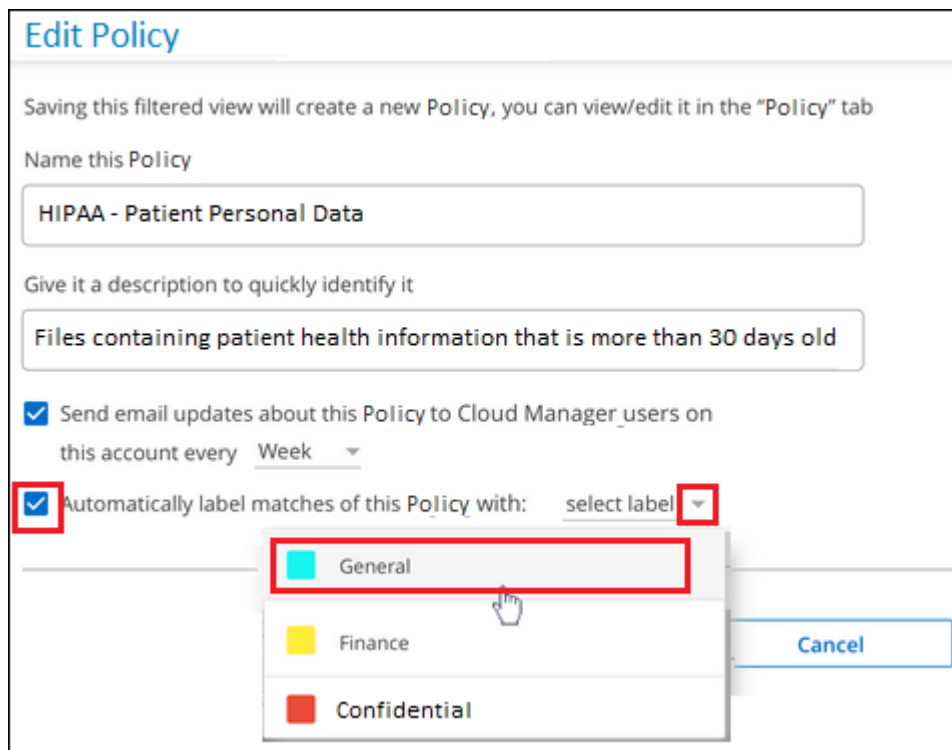
Follow these steps to add an AIP label to an existing Policy.

## Steps

1. From the Policies List page, click **Edit** for the Policy where you want to add (or change) the AIP label.



2. In the Edit Policy page, check the box to enable automatic labels for files that match the Policy parameters, and select the label (for example, **General**).



3. Click **Save Policy** and the label appears in the Policy description.



If a Policy was configured with a label, but the label has since been removed from AIP, the label name is turned to OFF and the label is not assigned anymore.

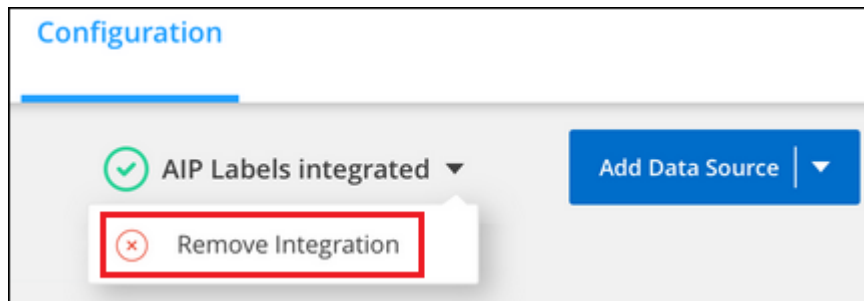
## Removing the AIP integration

If you no longer want the ability to manage AIP labels in files, you can remove the AIP account from the Cloud Compliance interface.

Note that no changes are made to the labels you have added using Cloud Compliance. The labels that exist in files will stay as they currently exist.

### Steps

1. From the *Scan Configuration* page, click **AIP Labels integrated > Remove Integration**.



2. Click **Remove Integration** from the confirmation dialog.

## Sending email alerts when non-compliant data is found

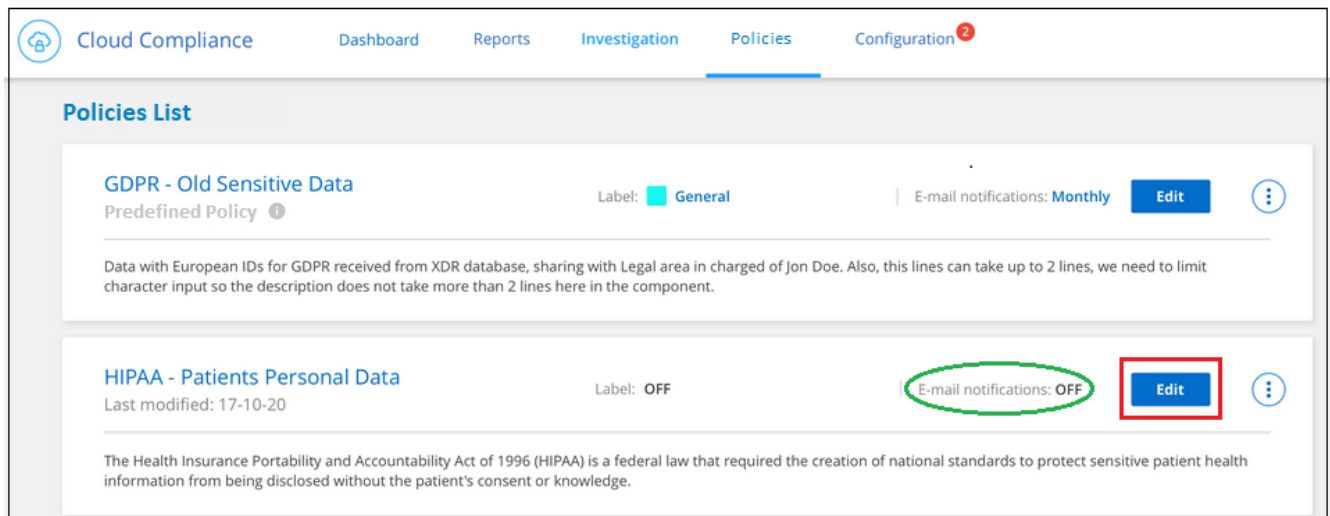
Cloud Compliance can send email alerts to Cloud Manager users when certain critical Policies return results so you can get notifications to protect your data. You can choose to send the email notifications on a daily, weekly, or monthly basis.

You can configure this setting when creating the Policy or when editing any Policy.

Follow these steps to add email updates to an existing Policy.

### Steps

1. From the Policies List page, click **Edit** for the Policy where you want to add (or change) the email setting.



2. In the Edit Policy page, check the box if you want notification emails sent to Cloud Manager users, and choose the interval at which the email is sent (for example, every **Week**).

3. Click **Save Policy** and the interval at which the email is sent appears in the Policy description.

### Result

The first email is sent now if there are any results from the Policy - but only if any files meet the Policy criteria. No personal information is sent in the notification emails. The email indicates that there are files that match the Policy criteria, and it provides a link to the Policy results.

## Deleting source files

You can permanently remove source files that seem insecure or too risky to leave in your storage system, or that you have identified as a duplicate. This action is permanent and there is no undo.



You can't delete files that reside in databases or files that reside in volume Backups.

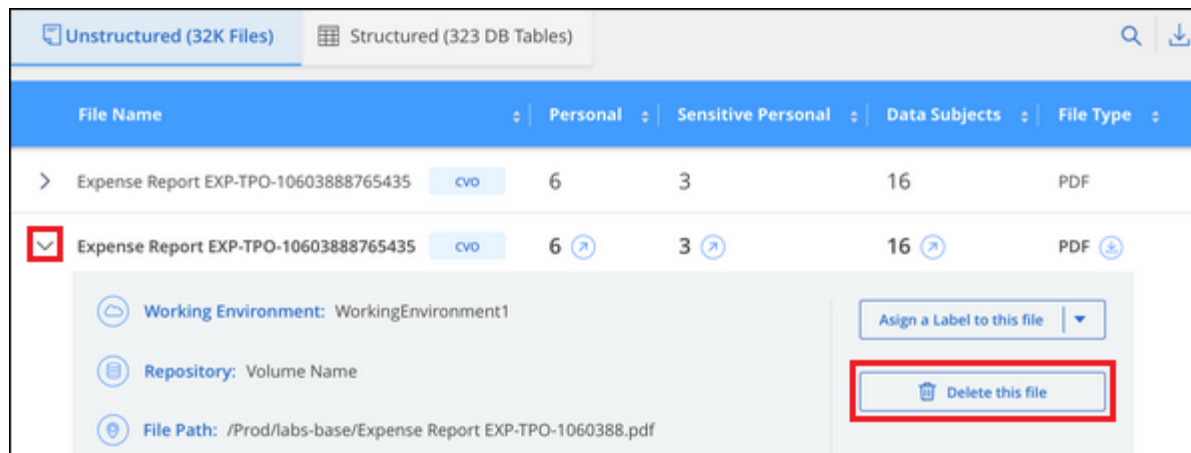
### Requirements

Deleting files requires the following permissions:

- For NFS data – the export policy needs to be defined with write permissions.
- For CIFS data – the CIFS credentials need to have write permissions.
- For S3 data - the IAM role must include the following permission: `s3:DeleteObject`

### Steps

1. In the Data Investigation results pane, click  for the file to expand the file metadata details.



2. Click **Delete this file**.
3. Because the delete operation is permanent, you must type "**permanently delete**" in the subsequent *Delete File* dialog and click **Delete File**.

## List of predefined Policies

Cloud Compliance provides the following system-defined Policies:

Name	Description	Logic
S3 publicly-exposed private data	S3 Objects containing personal or sensitive personal information, with open Public read access.	(S3 Public) AND contains personal OR sensitive personal info)
PCI DSS – Stale data over 30 days	Files containing Credit Card information, last modified over 30 days ago.	Contains credit card AND last modified over 30 days
HIPAA – Stale data over 30 days	Files containing Health information, last modified over 30 days ago.	Contains health data (defined same way as in HIPAA report) AND last modified over 30 days
Private data – Stale over 7 years	Files containing personal or sensitive personal information, last modified over 7 years ago.	Files containing personal or sensitive personal information, last modified over 7 years ago
GDPR – European citizens	Files containing more than 5 identifiers of an EU country's citizens or DB Tables containing identifiers of an EU country's citizens.	Files containing over 5 identifiers of an (one) EU citizens or DB Tables containing rows with over 15% of columns with one country's EU identifiers. (any one of the national identifiers of the European countries. Does not include Brazil, California, USA SSN, Israel, South Africa)
CCPA – California residents	Files containing over 10 California Driver's License identifiers or DB Tables with this identifier.	Files containing over 10 California Driver's License identifiers OR DB Tables containing California Driver's license
Data Subject names – High risk	Files with over 50 Data Subject names.	Files with over 50 Data Subject names

Name	Description	Logic
Email Addresses – High risk	Files with over 50 Email Addresses, or DB Columns with over 50% of their rows containing Email Addresses	Files with over 50 Email Addresses, or DB Columns with over 50% of their rows containing Email Addresses
Personal data – High risk	Files with over 20 Personal data identifiers, or DB Columns with over 50% of their rows containing Personal data identifiers.	Files with over 20 personal, or DB Columns with over 50% of their rows containing personal
Sensitive Personal data – High risk	Files with over 20 Sensitive Personal data identifiers, or DB Columns with over 50% of their rows containing Sensitive Personal data.	Files with over 20 sensitive personal, or DB Columns with over 50% of their rows containing sensitive personal

## Adding personal data identifiers using Data Fusion

A feature we call *Data Fusion* allows you to scan your organizations' data to identify whether unique identifiers from your databases are found in files or other databases - basically making your own list of "personal data" that is identified in Cloud Compliance scans. This gives you the full picture about where potentially sensitive data resides in *all* your files.

### Creating custom personal data identifiers from your databases

You can choose the additional identifiers that Cloud Compliance will look for in its' scans by selecting a specific column, or columns, in a database table. For example, the diagram below shows how data fusion is used to scan your volumes, buckets, and databases for occurrences of all your Customer IDs from your Oracle database.



## Databases -- Structured Data

Database: Oracle  
Schema: Accounts  
Table: Customers  
Column: Customer ID

Account	Name	Customer ID	Address
1234	ABC Co	135876	125 Main St
1235	XYZ Co	213536	35A Brick R
1236	Cat Co	359264	55 Wind Av
1237	Dog Co	472637	11025 Cor
1238	Zebra Co	582455	36 Sahara
...	...	...	...

*Scan your volumes and buckets for occurrences of the Customer IDs in your Oracle database*

## Files -- Unstructured Data

File in Volume 1

```
XXXXXXXXXXXXX
xx213536xxx
XXXXXXXXXXXXX
xx472637xxx
XXXXXXXXXXXXX
XXXXXXXXXXXXX
```

File in Volume 2

```
XXXXXXXXXXXXX
XXXXXXXXXXXXX
XXXXXXXXXXXXX
XXXXXXXXXXXXX
XXXXXXXXXXXXX
xxx472637xx
```

File in Bucket 1

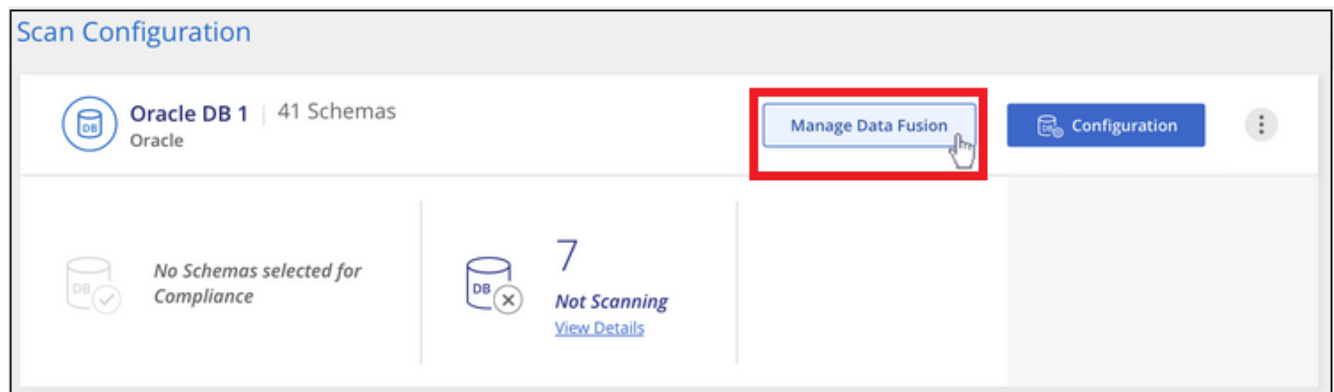
```
XXXXXXXXXXXXX
XXXXXXXXXXXXX
xx213536xxx
XXXXXXXXXXXXX
XXXXXXXXXXXXX
XXXXXXXXXXXXX
```

As you can see, two unique Customer IDs have been found in two volumes and in one S3 bucket. Any matches in database tables will also be identified.

### Steps

You must have [added at least one database server](#) to Cloud Compliance before you can add data fusion sources.

1. In the Scan Configuration page, click **Manage Data Fusion** in the database where the source data resides.



2. Click **Add Data Fusion source** on the next page.
3. In the *Add Data Fusion Source* page:
  - a. Select the Database Schema from the drop-down menu.
  - b. Enter the Table name in that schema.
  - c. Enter the Column, or Columns, that contain the unique identifiers you want to use.

When adding multiple columns, enter each column name, or table view name, on a separate line.

#### 4. Click **Add Data Fusion Source**.

The Data Fusion inventory page displays the database source columns that you have configured for Cloud Compliance to scan.

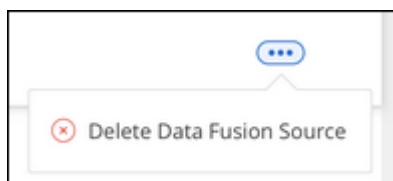
'DB Name 1' Data Fusion			<a href="#">+ Add Data Fusion source</a>
With Data Fusion, Cloud Compliance can identify occurrences of your organization's unique identifiers found in your unstructured data stores, using structured data indexes containing those unique identifiers as a source reference. <a href="#">Learn More</a>			
Database Schema	Table	Data Fusion Source Columns	
SchemaName1	Table 1	Column 12, Column 14, Column 18	...
SchemaName2	Table 2	Column 12, Column 14, Column 18	...

### Results

After the next scan, the results will include this new information in the Dashboard under the "Personal" results section, and in the Investigation page in the "Personal Data" filter. Each source column you added appears in the filter list as "Table.Column", for example `Customers.Customer ID`.

### Deleting a Data Fusion source

If at some point you decide not to scan your files using a certain Data Fusion source, you can select the source row from the Data Fusion inventory page and click **Delete Data Fusion Source**.



## Viewing compliance reports

Cloud Compliance provides reports that you can use to better understand the status of your organization's data privacy program.

By default, the Cloud Compliance dashboard displays compliance data for all working environments and databases. If you want to view reports that contain data for only some of the working environments, [select those working environments](#).



NetApp can't guarantee 100% accuracy of the personal data and sensitive personal data that Cloud Compliance identifies. You should always validate the information by reviewing the data.

## Privacy Risk Assessment Report

The Privacy Risk Assessment Report provides an overview of your organization's privacy risk status, as required by privacy regulations such as GDPR and CCPA. The report includes the following information:

### Compliance status

A [severity score](#) and the distribution of data, whether it's non-sensitive, personal, or sensitive personal.

### Assessment overview

A breakdown of the types of personal data found, as well as the categories of data.

### Data subjects in this assessment

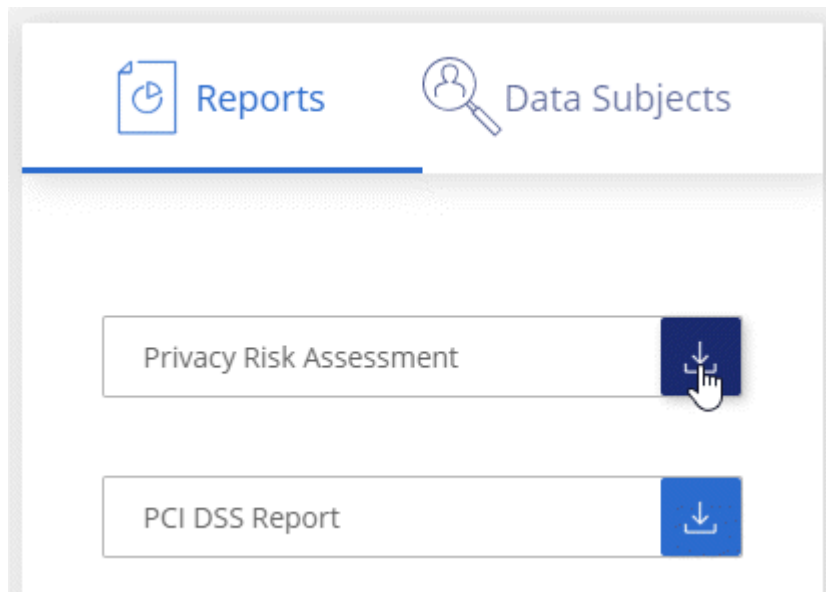
The number of people, by location, for which national identifiers were found.

## Generating the Privacy Risk Assessment Report

Go to the Compliance tab to generate the report.

### Steps

1. At the top of Cloud Manager, click **Compliance**.
2. Under **Reports**, click the download icon next to **Privacy Risk Assessment**.



### Result

Cloud Compliance generates a PDF report that you can review and send to other groups as needed.

### Severity score

Cloud Compliance calculates the severity score for the Privacy Risk Assessment Report on the basis of three variables:

- The percentage of personal data out of all data.
- The percentage of sensitive personal data out of all data.
- The percentage of files that include data subjects, determined by national identifiers such as national IDs,

Social Security numbers, and tax ID numbers.

The logic used to determine the score is as follows:

Severity score	Logic
0	All three variables are exactly 0%
1	One of the variables are larger than 0%
2	One of the variables are larger than 3%
3	Two of the variables are larger than 3%
4	Three of the variables are larger than 3%
5	One of the variables are larger than 6%
6	Two of the variables are larger than 6%
7	Three of the variables are larger than 6%
8	One of the variables are larger than 15%
9	Two of the variables are larger than 15%
10	Three of the variables are larger than 15%

## PCI DSS Report

The Payment Card Industry Data Security Standard (PCI DSS) Report can help you identify the distribution of credit card information across your files. The report includes the following information:

### Overview

How many files contain credit card information and in which working environments.

### Encryption

The percentage of files containing credit card information that are on encrypted or unencrypted working environments. This information is specific to Cloud Volumes ONTAP.

### Ransomware Protection

The percentage of files containing credit card information that are on working environments that do or don't have ransomware protection enabled. This information is specific to Cloud Volumes ONTAP.

### Retention

The timeframe in which the files were last modified. This is helpful because you shouldn't keep credit card information for longer than you need to process it.

### Distribution of Credit Card Information

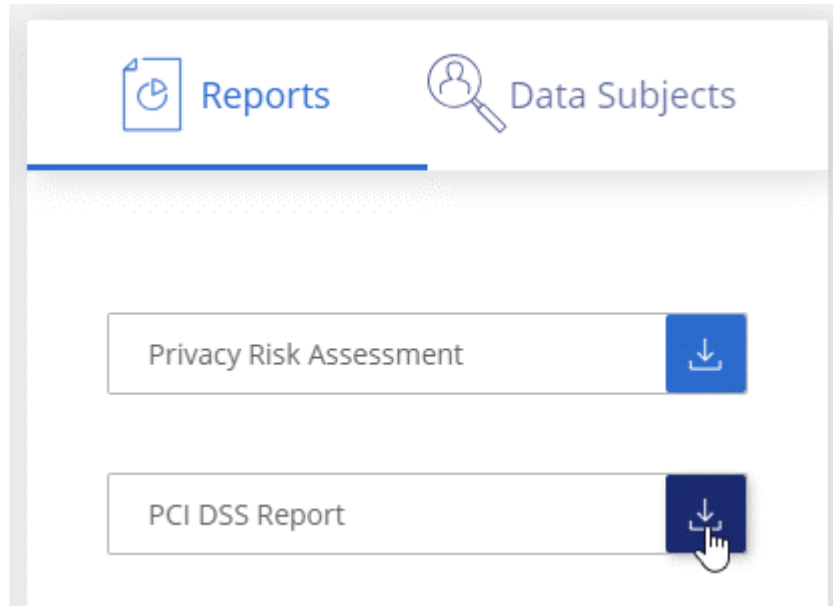
The working environments where the credit card information was found and whether encryption and ransomware protection are enabled.

## Generating the PCI DSS Report

Go to the Compliance tab to generate the report.

### Steps

1. At the top of Cloud Manager, click **Compliance**.
2. Under **Reports**, click the download icon next to **PCI DSS Report**.



### Result

Cloud Compliance generates a PDF report that you can review and send to other groups as needed.

## HIPAA Report

The Health Insurance Portability and Accountability Act (HIPAA) Report can help you identify files containing health information. It is designed to aid in your organization's requirement to comply with HIPAA data privacy laws. The information Cloud Compliance looks for includes:

- Health reference pattern
- ICD-10-CM Medical code
- ICD-9-CM Medical code
- HR – Health category
- Health Application Data category

The report includes the following information:

### Overview

How many files contain health information and in which working environments.

### Encryption

The percentage of files containing health information that are on encrypted or unencrypted working environments. This information is specific to Cloud Volumes ONTAP.

### Ransomware Protection

The percentage of files containing health information that are on working environments that do or don't have ransomware protection enabled. This information is specific to Cloud Volumes ONTAP.

## Retention

The timeframe in which the files were last modified. This is helpful because you shouldn't keep health information for longer than you need to process it.

## Distribution of Health Information

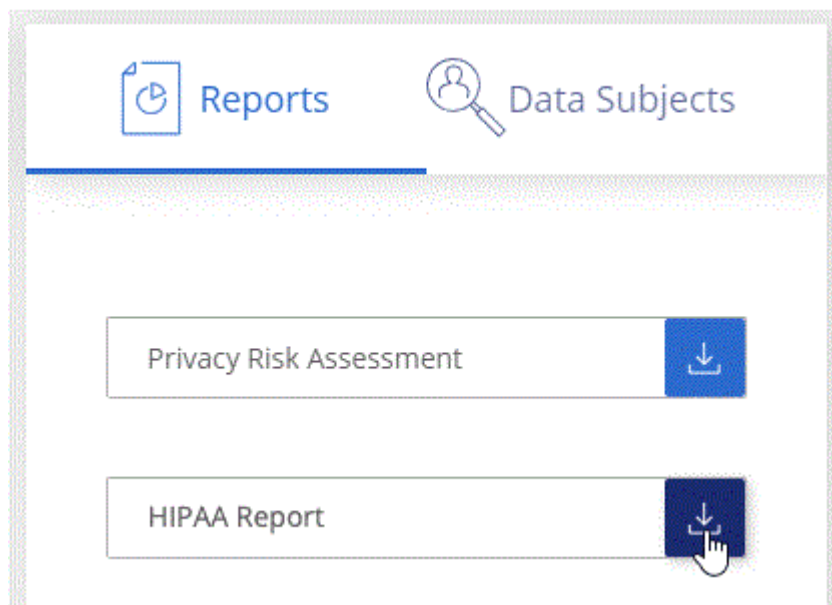
The working environments where the health information was found and whether encryption and ransomware protection are enabled.

## Generating the HIPAA Report

Go to the Compliance tab to generate the report.

### Steps

1. At the top of Cloud Manager, click **Compliance**.
2. Under **Reports**, click the download icon next to **HIPAA Report**.



### Result

Cloud Compliance generates a PDF report that you can review and send to other groups as needed.

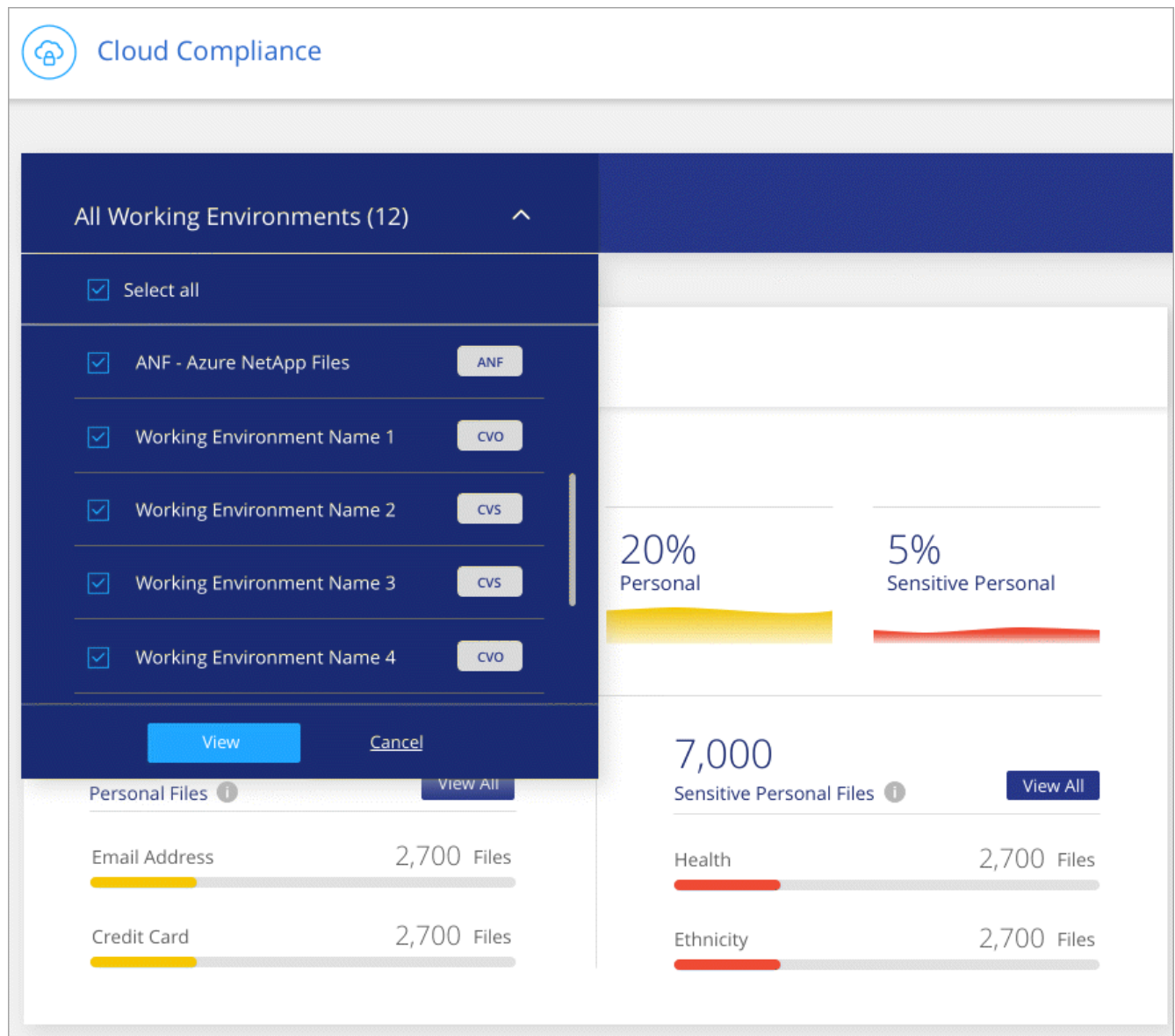
## Selecting the working environments for reports

You can filter the contents of the Cloud Compliance dashboard to see compliance data for all working environments and databases, or for just specific working environments.

When you filter the dashboard, Cloud Compliance scopes the compliance data and reports to just those working environments that you selected.

### Steps

1. Click the filter drop-down, select the working environments that you'd like to view data for, and click **View**.



## Responding to a Data Subject Access Request

Respond to a Data Subject Access Request (DSAR) by searching for a subject's full name or known identifier (such as an email address) and then downloading a report. The report is designed to aid in your organization's requirement to comply with GDPR or similar data privacy laws.



NetApp can't guarantee 100% accuracy of the personal data and sensitive personal data that Cloud Compliance identifies. You should always validate the information by reviewing the data.

### What is a Data Subject Access Request?

Privacy regulations such as the European GDPR grant data subjects (such as customers or employees) the right to access their personal data. When a data subject requests this information, this is known as a DSAR (data subject access request). Organizations are required to respond to these requests "without undue delay,"

and at the latest within one month of receipt.

## How can Cloud Compliance help you respond to a DSAR?

When you perform a data subject search, Cloud Compliance finds all of the files, buckets, and databases that have that person's name or identifier in it. Cloud Compliance checks the latest pre-indexed data for the name or identifier. It doesn't initiate a new scan.

After the search is complete, you can then download the list of files for a Data Subject Access Request report. The report aggregates insights from the data and puts it into legal terms that you can send back to the person.

## Searching for data subjects and downloading reports

Search for the data subject's full name or known identifier and then download a file list report or DSAR report. You can search by [any personal information type](#).

Only English is supported when searching for the names of data subjects. Support for more languages will be added later.



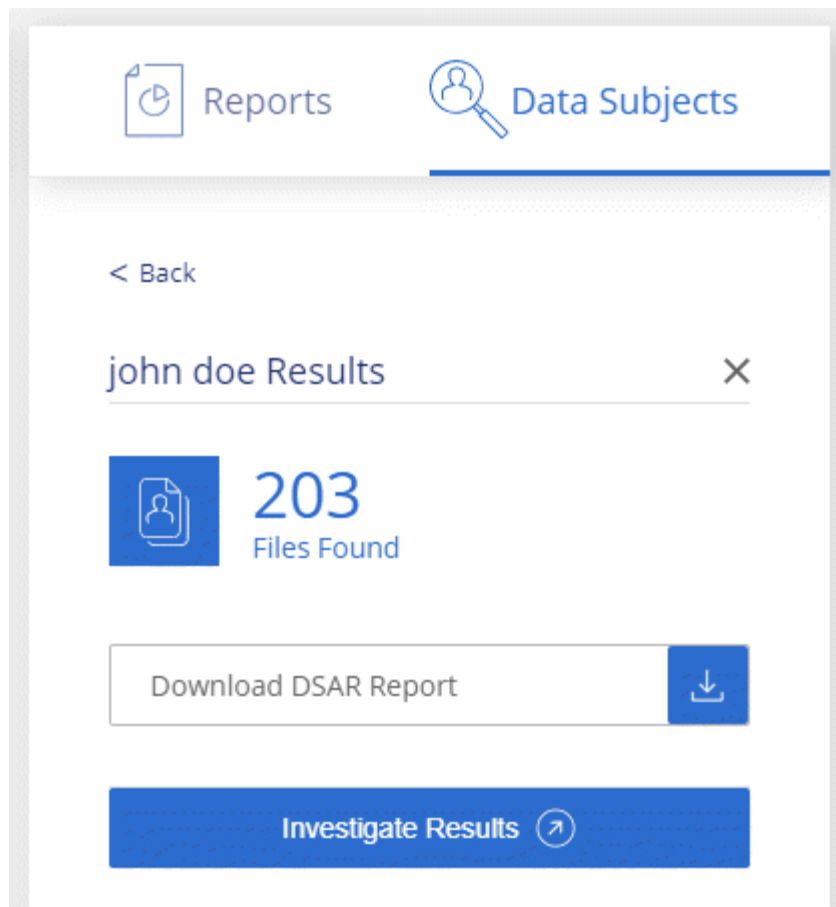
Data subject search is not supported within databases at this time.

### Steps

1. At the top of Cloud Manager, click **Compliance**.
2. Click **Data Subjects**.
3. Search for the data subject's full name or known identifier.

Here's an example that shows a search for the name *john doe*:





4. Choose one of the available options:

- **Download DSAR Report:** A formal response to the access request that you can send to the data subject. This report contains automatically-generated information based on data that Cloud Compliance found on the data subject and is designed to be used as a template. You should complete the form and review it internally before sending it to the data subject.
- **Investigate Results:** A page that enables you to investigate the data by searching, sorting, expanding details for a specific file, and by downloading the file list.



If there are more than 10,000 results, only the top 10,000 appear in the file list.

## Categories of private data

There are many types of private data that Cloud Compliance can identify in your volumes, Amazon S3 buckets, databases, and OneDrive folders. See the categories below.



If you need Cloud Compliance to identify other private data types, such as additional national ID numbers or healthcare identifiers, email [ng-contact-cloud-compliance@netapp.com](mailto:ng-contact-cloud-compliance@netapp.com) with your request.

## Types of personal data

The personal data found in files can be general personal data or national identifiers. The third column identifies whether Cloud Compliance uses [proximity validation](#) to validate its findings for the identifier.

Note that you can add to the list of personal data that is found in your files if you are scanning a database server. The *Data Fusion* feature allows you to choose the additional identifiers that Cloud Compliance will look for in its' scans by selecting columns in a database table. See [Adding personal data identifiers using Data Fusion](#) for details.

Type	Identifier	Proximity validation?
General	Email address	No
	Credit card number	No
	IBAN number (International Bank Account Number)	No
	IP address	No

Type	Identifier	Proximity validation?
National Identifiers	Belgian ID (Numero National)	Yes
	Brazilian ID (CPF)	Yes
	Bulgarian ID (UCN)	Yes
	California Driver's License	Yes
	Croatian ID (OIB)	Yes
	Cyprus Tax Identification Number (TIC)	Yes
	Czech/Slovak ID	Yes
	Danish ID (CPR)	Yes
	Dutch ID (BSN)	Yes
	Estonian ID	Yes
	Finnish ID (HETU)	Yes
	French Tax Identification Number (SPI)	Yes
	German Tax Identification Number (Steuerliche Identifikationsnummer)	Yes
	Greek ID	Yes
	Hungarian Tax Identification Number	Yes
	Irish ID (PPS)	Yes
	Israeli ID	Yes
	Italian Tax Identification Number	Yes
	Latvian ID	Yes
	Lithuanian ID	Yes
	Luxembourg ID	Yes
	Maltese ID	Yes
	Polish ID (PESEL)	Yes
	Portuguese Tax Identification Number (NIF)	Yes
	Romanian ID (CNP)	Yes
	Slovenian ID (EMSO)	Yes
	South African ID	Yes
	Spanish Tax Identification Number	Yes
	Swedish ID	Yes
	U.K. ID (NINO)	Yes
	USA Social Security Number (SSN)	Yes

## **Types of sensitive personal data**

The sensitive personal data that Cloud Compliance can find in files includes the following:

### **Criminal Procedures Reference**

Data concerning a natural person's criminal convictions and offenses.

### **Ethnicity Reference**

Data concerning a natural person's racial or ethnic origin.

### **Health Reference**

Data concerning a natural person's health.

### **ICD-9-CM Medical Codes**

Codes used in the medical and health industry.

### **ICD-10-CM Medical Codes**

Codes used in the medical and health industry.

### **Philosophical Beliefs Reference**

Data concerning a natural person's philosophical beliefs.

### **Political Opinions Reference**

Data concerning a natural person's political opinions.

### **Religious Beliefs Reference**

Data concerning a natural person's religious beliefs.

### **Sex Life or Orientation Reference**

Data concerning a natural person's sex life or sexual orientation.

## **Types of categories**

Cloud Compliance categorizes your data as follows:

### **Finance**

- Balance Sheets
- Purchase Orders
- Invoices
- Quarterly Reports

### **HR**

- Background Checks
- Compensation Plans
- Employee Contracts
- Employee Reviews
- Health
- Resumes

**Legal**

- NDAs
- Vendor-Customer contracts

**Marketing**

- Campaigns
- Conferences

**Operations**

- Audit Reports

**Sales**

- Sales Orders

**Services**

- RFI
- RFP
- SOW
- Training

**Support**

- Complaints and Tickets

**Metadata categories**

- Application Data
- Archive Files
- Audio
- Business Application Data
- CAD Files
- Code
- Database and index files
- Design Files
- Email Application Data
- Executables
- Financial Application Data
- Health Application Data
- Images
- Logs
- Miscellaneous Documents
- Miscellaneous Presentations
- Miscellaneous Spreadsheets
- Videos

## Types of files

Cloud Compliance scans all files for category and metadata insights and displays all file types in the file types section of the dashboard.

But when Cloud Compliance detects Personal Identifiable Information (PII), or when it performs a DSAR search, only the following file formats are supported:

.PDF, .DOCX, .DOC, .PPTX, .XLS, .XLSX, .CSV, .TXT, .RTF, and .JSON.

## Accuracy of information found

NetApp can't guarantee 100% accuracy of the personal data and sensitive personal data that Cloud Compliance identifies. You should always validate the information by reviewing the data.

Based on our testing, the table below shows the accuracy of the information that Cloud Compliance finds. We break it down by *precision* and *recall*:

### Precision

The probability that what Cloud Compliance finds has been identified correctly. For example, a precision rate of 90% for personal data means that 9 out of 10 files identified as containing personal information, actually contain personal information. 1 out of 10 files would be a false positive.

### Recall

The probability for Cloud Compliance to find what it should. For example, a recall rate of 70% for personal data means that Cloud Compliance can identify 7 out of 10 files that actually contain personal information in your organization. Cloud Compliance would miss 30% of the data and it won't appear in the dashboard.

Cloud Compliance is in a Controlled Availability release and we are constantly improving the accuracy of our results. Those improvements will be automatically available in future Cloud Compliance releases.


Type	Precision	Recall
Personal data - General	90%-95%	60%-80%
Personal data - Country identifiers	30%-60%	40%-60%
Sensitive personal data	80%-95%	20%-30%
Categories	90%-97%	60%-80%

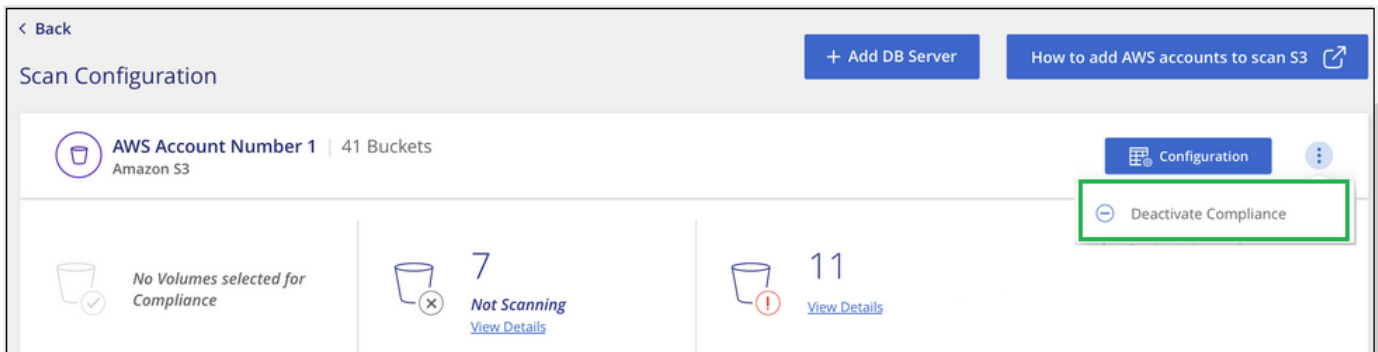
## Removing data sources from Cloud Compliance

If you need to, you can stop Cloud Compliance from scanning one or more working environments, databases, file share groups, or OneDrive accounts. You can also delete the Cloud Compliance instance if you no longer want to use Cloud Compliance with your working environments.

### Deactivating compliance scans for a working environment

When you deactivate scans, Cloud Compliance no longer scans the data on the working environment and it removes the indexed compliance insights from the Cloud Compliance instance (the data from the working environment itself isn't deleted).


1. From the *Scan Configuration* page, click the  button in the row for the working environment, and then click **Deactivate Compliance**.

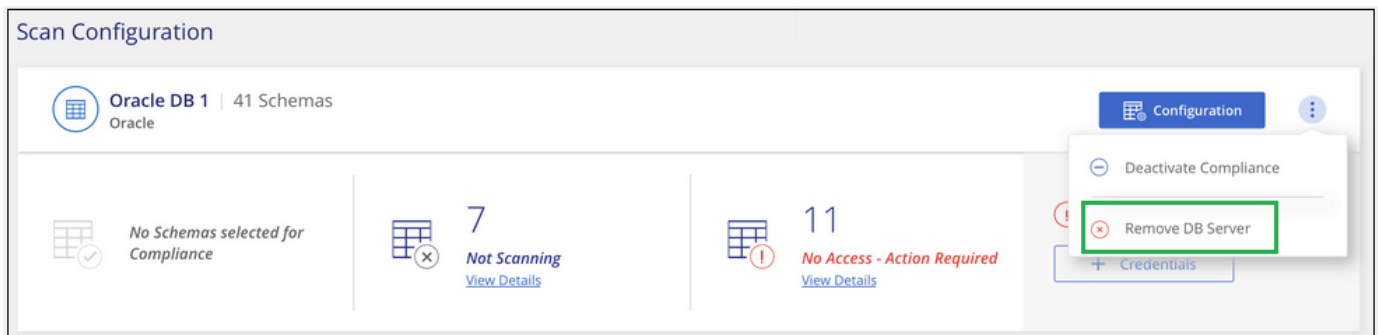


You can also disable compliance scans for a working environment from the Services panel when you select the working environment.

## Removing a database from Cloud Compliance

If you no longer want to scan a certain database, you can delete it from the Cloud Compliance interface and stop all scans.

1. From the *Scan Configuration* page, click the  button in the row for the database, and then click **Remove DB Server**.



## Removing a OneDrive account from Cloud Compliance

If you no longer want to scan user files from a certain OneDrive account, you can delete the account from the Cloud Compliance interface and stop all scans.

### Steps

1. From the *Scan Configuration* page, click the  button in the row for the OneDrive account, and then click **Remove OneDrive Account**.




2. Click **Delete Account** from the confirmation dialog.

## Removing a group of file shares from Cloud Compliance

If you no longer want to scan user files from a file shares group, you can delete the File Shares Group from the Cloud Compliance interface and stop all scans.

### Steps

1. From the *Scan Configuration* page, click the  button in the row for the File Shares Group, and then click **Remove File Shares Group**.



2. Click **Delete Group of Shares** from the confirmation dialog.

## Deleting the Cloud Compliance instance

You can delete the Cloud Compliance instance if you no longer want to use Cloud Compliance. Deleting the instance also deletes the associated disks where the indexed data resides.

1. Go to your cloud provider's console and delete the Cloud Compliance instance.

The instance is named *CloudCompliance* with a generated hash (UUID) concatenated to it. For example: *CloudCompliance-16bb6564-38ad-4080-9a92-36f5fd2f71c7*

## Frequently asked questions about Cloud Compliance

This FAQ can help if you're just looking for a quick answer to a question.

### What is Cloud Compliance?

Cloud Compliance is a cloud offering that uses Artificial Intelligence (AI) driven technology to help organizations understand data context and identify sensitive data across your storage systems. The systems can be Azure NetApp Files configurations, Cloud Volumes ONTAP systems hosted in AWS or Azure, Amazon S3 buckets, on-prem ONTAP systems, non-NetApp file shares, generic S3 object storage, databases, and OneDrive accounts.



Cloud Compliance provides pre-defined parameters (such as sensitive information types and categories) to address new data compliance regulations for data privacy and sensitivity, such as GDPR, CCPA, HIPAA, and more.

## Why should I use Cloud Compliance?

Cloud Compliance can empower you with data to help you:

- Comply with data compliance and privacy regulations.
- Comply with data retention policies.
- Easily locate and report on specific data in response to data subjects, as required by GDPR, CCPA, HIPAA, and other data privacy regulations.

## What are the common use cases for Cloud Compliance?

- Identify Personal Identifiable Information (PII).
- Identify a wide scope of sensitive information as required by GDPR and CCPA privacy regulations.
- Comply with new and upcoming data privacy regulations.

[Learn more about the use cases for Cloud Compliance.](#)

## What types of data can be scanned with Cloud Compliance?

Cloud Compliance supports scanning of unstructured data over NFS and CIFS protocols that are managed by Cloud Volumes ONTAP, Azure NetApp Files, and on-prem ONTAP systems. Cloud Compliance can also scan data stored on Amazon S3 buckets, in generic S3 object storage, and non-NetApp file shares.

Additionally, Cloud Compliance can scan databases that are located anywhere, and user files from OneDrive accounts.

[Learn how scans work.](#)

## Which cloud providers are supported?

Cloud Compliance operates as part of Cloud Manager and currently supports AWS and Azure. This provides your organization with unified privacy visibility across different cloud providers.

## How do I access Cloud Compliance?

Cloud Compliance is operated and managed through Cloud Manager. You can access Cloud Compliance features from the **Compliance** tab in Cloud Manager.

## How does Cloud Compliance work?

Cloud Compliance deploys another layer of Artificial Intelligence alongside your Cloud Manager system and storage systems. It then scans the data on volumes, buckets, databases, and OneDrive accounts and indexes the data insights that are found.

[Learn more about how Cloud Compliance works.](#)

## How much does Cloud Compliance cost?

The cost to use Cloud Compliance depends on the amount of data that you're scanning. The first 1 TB of data that Cloud Compliance scans in a Cloud Manager workspace is free. A subscription to the AWS or Azure Marketplace is required to continue scanning data after that point. See [pricing](#) for details.

## What type of instance or VM is required for Cloud Compliance?

- In Azure, Cloud Compliance runs on a Standard\_D16s\_v3 VM with a 512 GB disk.
- In AWS, Cloud Compliance runs on an m5.4xlarge instance with a 500 GB GP2 disk.

In regions where m5.4xlarge isn't available, Cloud Compliance runs on an m4.4xlarge instance instead.

You can also download and install Compliance software on a Linux host in your network or in the cloud. Everything works the same and you continue to manage your scan configuration and results through Cloud Manager. See [Deploying Cloud Compliance on premises](#) for system requirements and installation details.



Cloud Compliance is currently unable to scan S3 buckets and ANF files when it is installed on premises.

[Learn more about how Cloud Compliance works.](#)

## How often does Cloud Compliance scan my data?

Data changes frequently, so Cloud Compliance scans your data continuously with no impact to your data. While the initial scan of your data might take longer, subsequent scans only scan the incremental changes, which reduces system scan times.

[Learn how scans work.](#)

## Does Cloud Compliance offer reports?

Yes. The information offered by Cloud Compliance can be relevant to other stakeholders in your organizations, so we enable you to generate reports to share the insights.

The following reports are available for Cloud Compliance:

### Privacy Risk Assessment report

Provides privacy insights from your data and a privacy risk score. [Learn more.](#)

### Data Subject Access Request report

Enables you to extract a report of all files that contain information regarding a data subject's specific name or personal identifier. [Learn more.](#)

### PCI DSS report

Helps you identify the distribution of credit card information across your files. [Learn more.](#)

### HIPAA report

Helps you identify the distribution of health information across your files. [Learn more.](#)

## Reports on a specific information type

Reports are available that include details about the identified files that contain personal data and sensitive personal data. You can also see files broken down by category and file type. [Learn more](#).

## Does scan performance vary?

Scan performance can vary based on the network bandwidth and the average file size in your cloud environment.

## Which file types are supported?

Cloud Compliance scans all files for category and metadata insights and displays all file types in the file types section of the dashboard.

When Cloud Compliance detects Personal Identifiable Information (PII), or when it performs a DSAR search, only the following file formats are supported:

.PDF, .DOCX, .DOC, .PPTX, .XLS, .XLSX, .CSV, .TXT, .RTF, and .JSON.

## How do I enable Cloud Compliance?

First you need to deploy an instance of Cloud Compliance in Cloud Manager. Once the instance is running, you can enable it on existing working environments and databases from the **Compliance** tab or by selecting a specific working environment.

[Learn how to get started](#).



Activating Cloud Compliance results in an immediate initial scan. Compliance results display shortly after.

## How do I disable Cloud Compliance?

You can disable Cloud Compliance from the Canvas page after you select an individual working environment, database, file share group, or OneDrive account.

[Learn more](#).



To completely remove the Cloud Compliance instance, you can manually remove the Cloud Compliance instance from your cloud provider's portal.

## What happens if data tiering is enabled on Cloud Volumes ONTAP?

You might want to enable Cloud Compliance on a Cloud Volumes ONTAP system that tiers cold data to object storage. If data tiering is enabled, Cloud Compliance scans all of the data—data that's on disks and cold data tiered to object storage.

The compliance scan doesn't heat up the cold data—it stays cold and tiered to object storage.

## Can I use Cloud Compliance to scan on-premises ONTAP storage?

Yes. As long as you have discovered the on-prem ONTAP cluster as a working environment in Cloud Manager, you can scan any of the volume data.

Alternatively, you can run compliance scans on backup files created from your on-prem ONTAP volumes. So if you're already creating backup files from your on-prem systems using [Cloud Backup](#), you can run compliance scans on those backup files.

[Learn more.](#)

## Can Cloud Compliance send notifications to my organization?

Yes. In conjunction with the Policies feature, you can send email alerts to Cloud Manager users (daily, weekly, or monthly) when a Policy returns results so you can get notifications to protect your data. [Learn more about Policies.](#)

You can also download status reports from the Investigation page in .CSV format that you can share internally in your organization.

## Can I customize the service to my organization's needs?

Cloud Compliance provides out-of-the-box insights to your data. These insights can be extracted and used for your organization's needs.

Additionally, you can use the **Data Fusion** capability to have Cloud Compliance scan all your data based on criteria found in specific columns in databases you are scanning — essentially allowing you to make your own custom personal data types.

[Learn more.](#)

## Can Cloud Compliance work with the AIP labels I have embedded in my files?

Yes. You can manage AIP labels in the files that Cloud Compliance is scanning if you have subscribed to [Azure Information Protection \(AIP\)](#). You can view the labels that are already assigned to files, add labels to files, and change existing labels.

[Learn more.](#)

## Can I limit Cloud Compliance information to specific users?

Yes, Cloud Compliance is fully integrated with Cloud Manager. Cloud Manager users can only see information for the working environments they are eligible to view according to their workspace privileges.

Additionally, if you want to allow certain users to just view Cloud Compliance scan results without having the ability to manage Cloud Compliance settings, you can assign those users the *Cloud Compliance Viewer* role.

[Learn more.](#)

## 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.