Scan data deprecations

BlueXP classification

NetApp
June 20, 2024

Table of Contents

Scan data deprecations

Scan Amazon S3 buckets

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

NOTE This information is relevant only for BlueXP classification legacy versions 1.30 and earlier.

Quick start

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

1. **Set up the S3 requirements in your cloud environment**

Ensure that your cloud environment can meet the requirements for BlueXP classification, including preparing an IAM role and setting up connectivity from BlueXP classification to S3. See the complete list.

2. **Deploy the BlueXP classification instance**

Deploy BlueXP classification if there isn't already an instance deployed.

3. **Activate BlueXP classification on your S3 working environment**

Select the Amazon S3 working environment, click **Enable**, 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 BlueXP classification will start scanning them.

Reviewing S3 prerequisites

The following requirements are specific to scanning S3 buckets.

Set up an IAM role for the BlueXP classification instance

BlueXP classification 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. BlueXP prompts you to select an IAM role when you enable BlueXP classification on the Amazon S3 working environment.

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

Provide connectivity from BlueXP classification to Amazon S3

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

Deploy BlueXP classification in BlueXP if there isn't already an instance deployed.

You need to deploy the instance using a Connector deployed in AWS so that BlueXP automatically discovers the S3 buckets in this AWS account and displays them in an Amazon S3 working environment.

Note: Deploying BlueXP classification in an on-premises location is not currently supported when scanning S3 buckets.

Upgrades to BlueXP classification software is automated as long as the instance has internet connectivity.

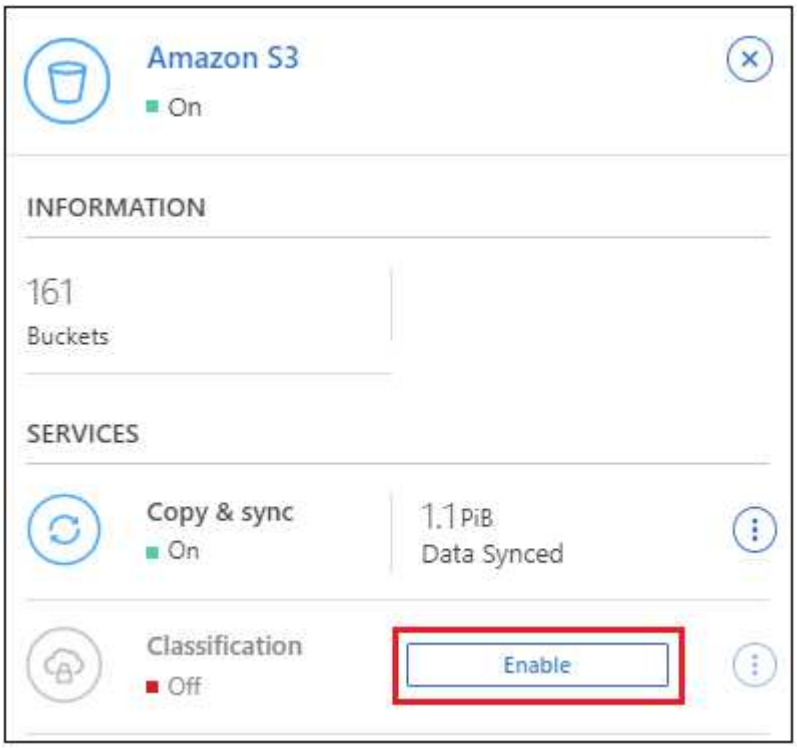
Activating BlueXP classification on your S3 working environment

Enable BlueXP classification on Amazon S3 after you verify the prerequisites.

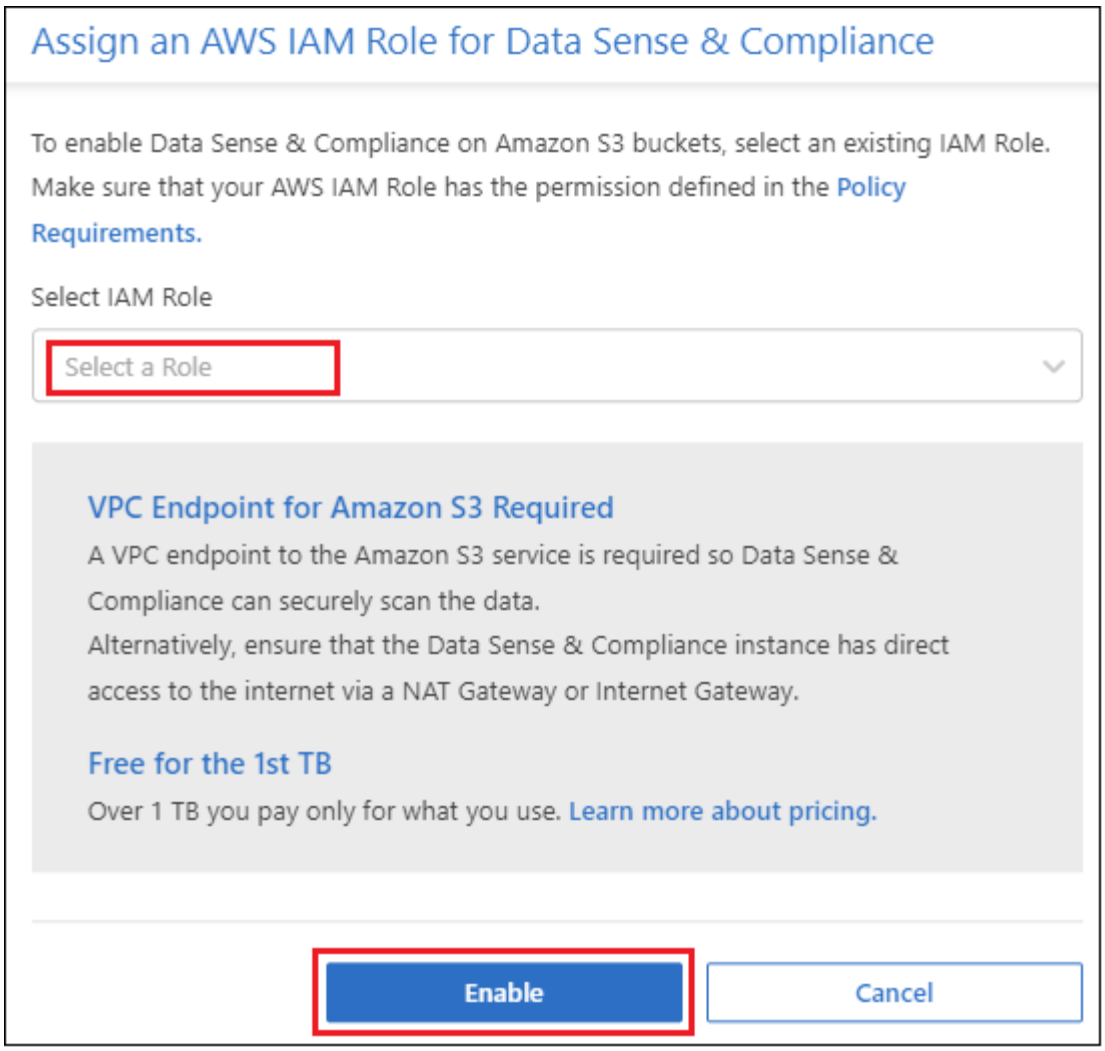
Steps

1. From the BlueXP left navigation menu, click **Storage > Canvas**.
2. Select the Amazon S3 working environment.

3. In the Services pane on the right, click **Enable** next to **Classification**.



4. When prompted, assign an IAM role to the BlueXP classification instance that has the required permissions.



5. Click **Enable**.

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

Result

BlueXP assigns the IAM role to the instance.

Enabling and disabling compliance scans on S3 buckets

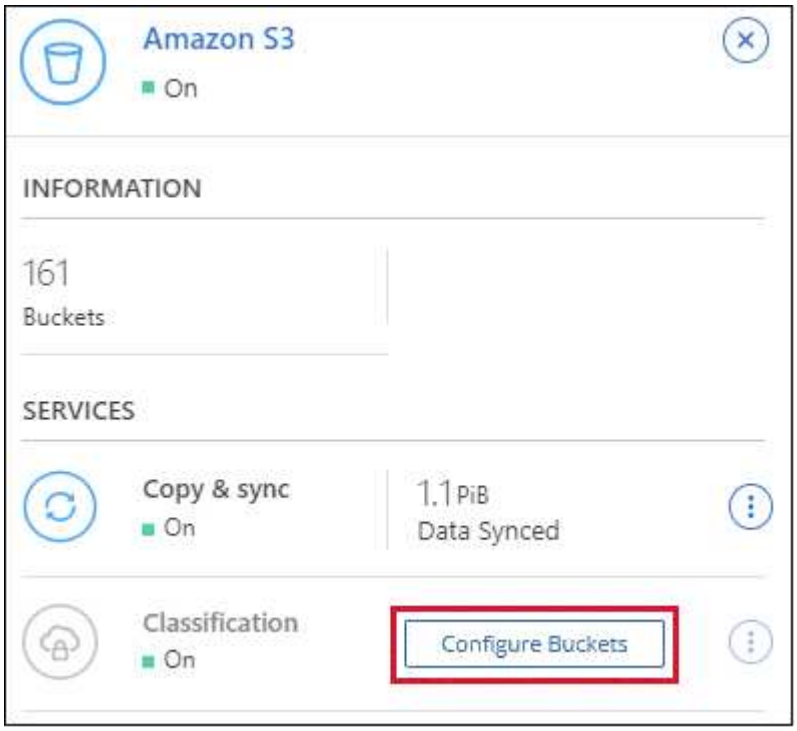
After BlueXP enables BlueXP classification on Amazon S3, the next step is to configure the buckets that you want to scan.

When BlueXP 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.

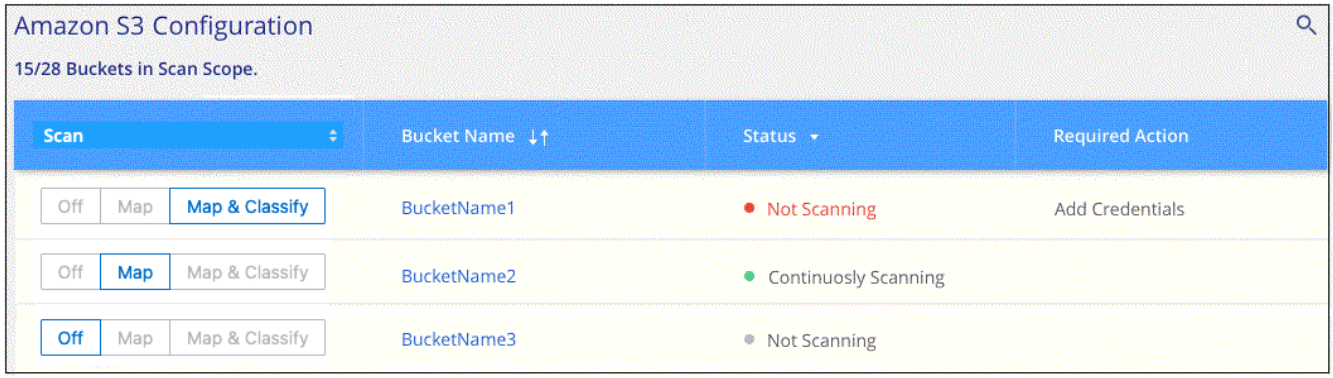
BlueXP classification can also scan S3 buckets that are in different AWS accounts.

Steps

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



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



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

Result

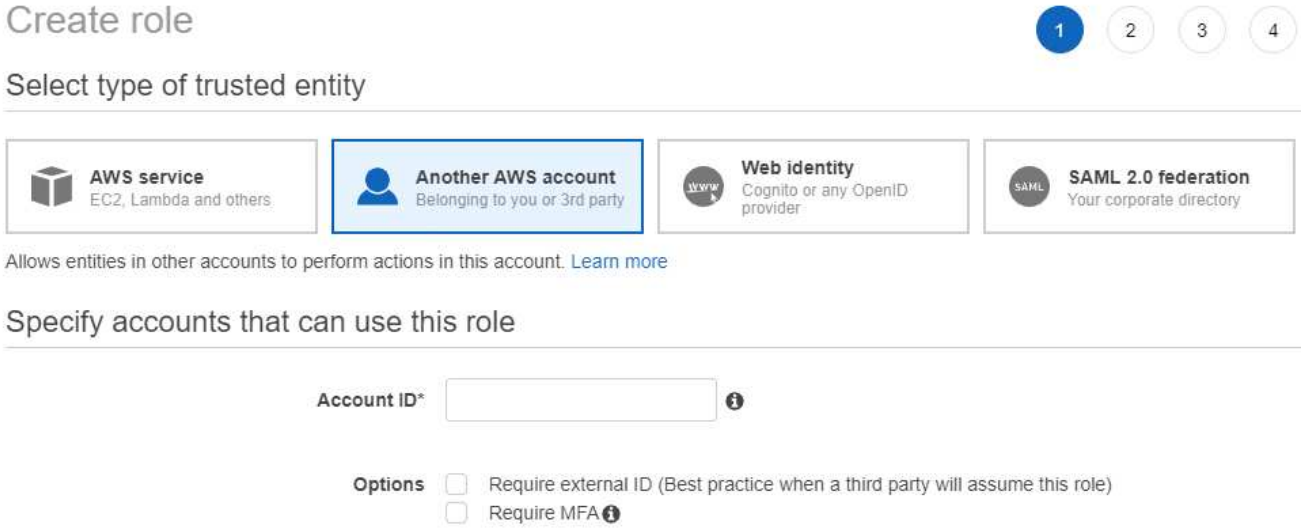
BlueXP classification starts scanning the S3 buckets that you enabled. If there are any errors, they'll appear in the Status column, alongside the required action to fix the error.

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 BlueXP classification instance.

Steps

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



   Be sure to do the following:

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

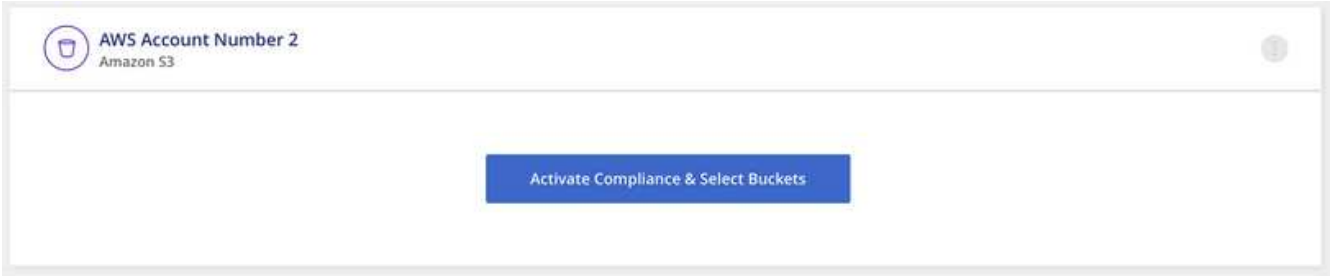
```
{
  "Version": "2012-10-17",
  "Statement": [
      {
          "Effect": "Allow",
          "Action": [
              "s3:Get*",
              "s3:List*",
              "s3:PutObject"
          ],
          "Resource": "*"
      },
  ]
}
```

2. Go to the source AWS account where the BlueXP classification 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 BlueXP classification instance profile account now has access to the additional AWS account.

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



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

Result

BlueXP classification starts scanning the new S3 buckets that you enabled.

Scan OneDrive accounts

Complete a few steps to start scanning files in your user's OneDrive folders with BlueXP

classification.

NOTE This information is relevant only for BlueXP classification legacy versions 1.30 and earlier.

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 BlueXP classification instance

Deploy BlueXP classification 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 and select the type of scanning

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

Reviewing OneDrive requirements

Review the following prerequisites to make sure that you have a supported configuration before you enable BlueXP classification.

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

Deploying the BlueXP classification instance

Deploy BlueXP classification if there isn't already an instance deployed.

BlueXP classification can be deployed in the cloud or in an on-premises location that has internet access.

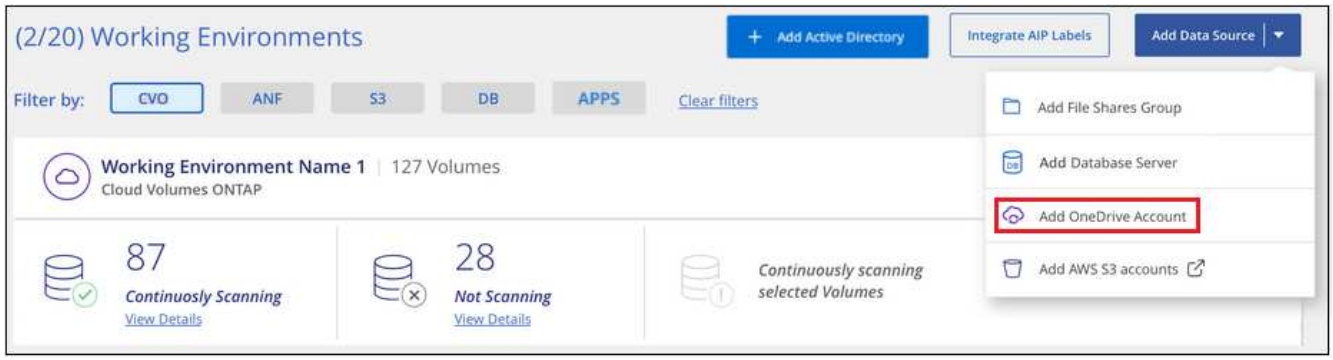
Upgrades to BlueXP classification software is automated as long as the instance has internet connectivity.

Adding the OneDrive account

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 BlueXP classification to read data from this account.

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

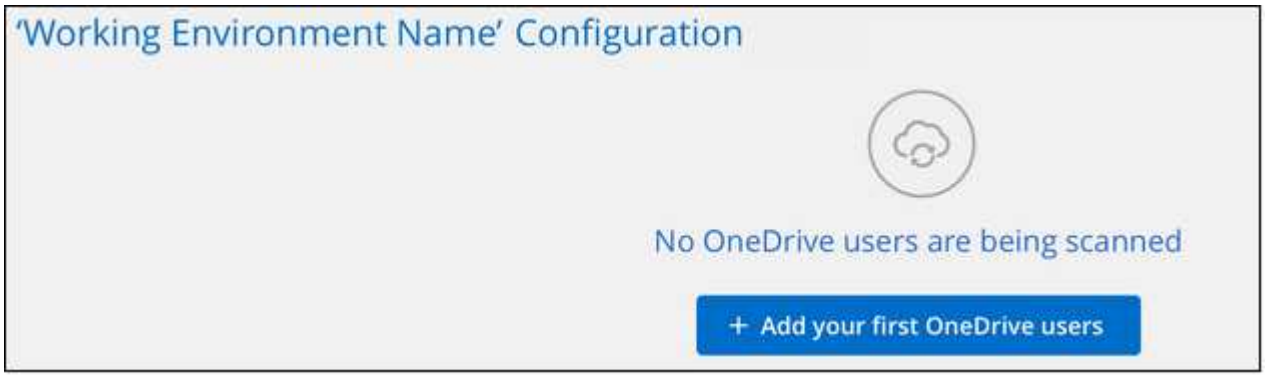
Adding OneDrive users to compliance scans

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

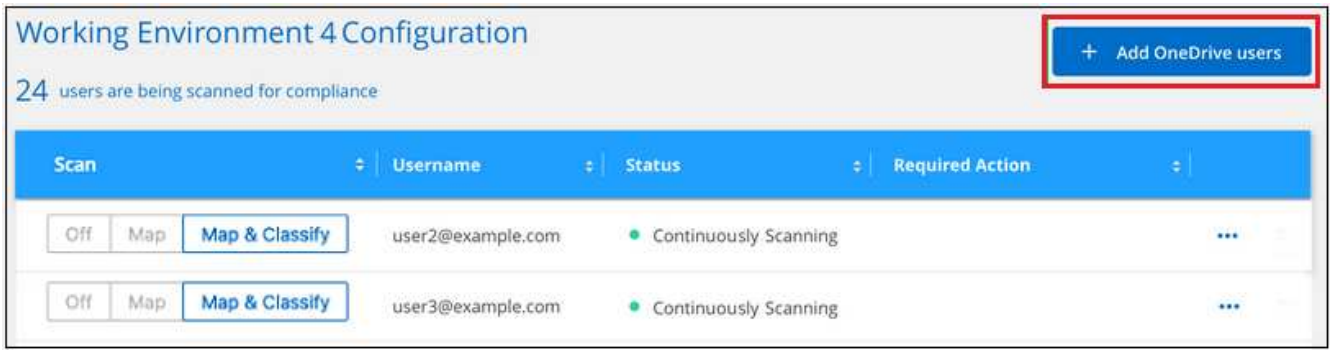
Steps

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

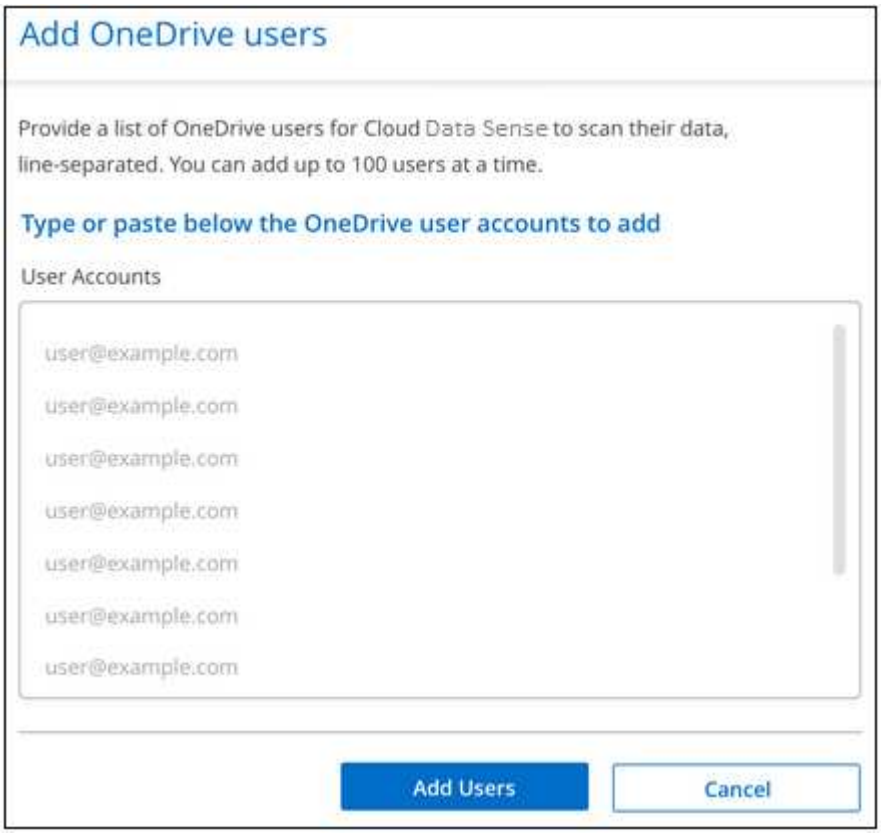
2. If this is the first time adding users for this OneDrive account, click **Add your first OneDrive users**.



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



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



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

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

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

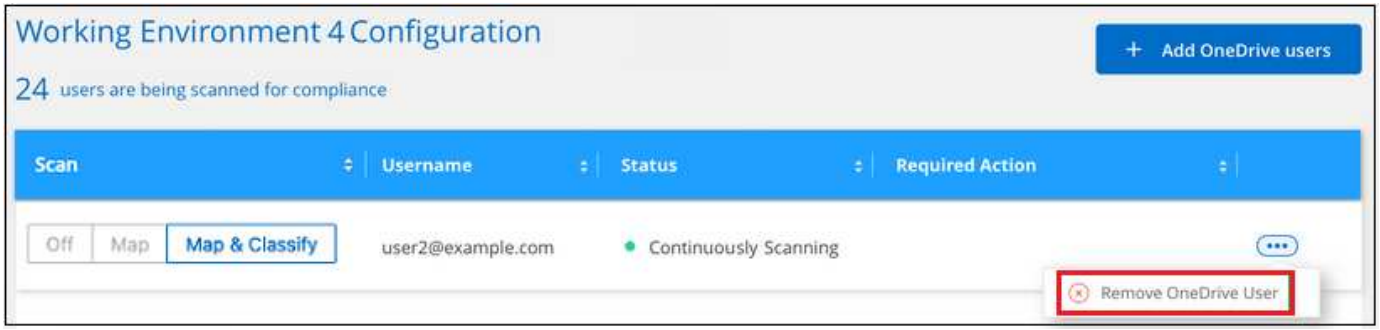
To:	Do this:
Enable mapping-only scans on user files	Click **Map**
Enable full scans on user files	Click **Map & Classify**
Disable scanning on user files	Click **Off**

Result

BlueXP classification 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.



Scan SharePoint accounts

Complete a few steps to start scanning files in your SharePoint Online and SharePoint On-Premise accounts with BlueXP classification.

NOTE This information is relevant only for BlueXP classification legacy versions 1.30 and earlier.

Quick start

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

1 **Review SharePoint prerequisites**

Ensure that you have qualified credentials to log into the SharePoint account, and that you have the URLs for the SharePoint sites that you want to scan.

2 **Deploy the BlueXP classification instance**

Deploy BlueXP classification if there isn't already an instance deployed.

3 **Log into the SharePoint account**

Using qualified user credentials, log into the SharePoint account that you want to access so that it is added as a new data source/working environment.

4 **Add the SharePoint site URLs to scan**

Add the list of SharePoint site URLs that you want to scan in the SharePoint account, and select the type of scanning. You can add up to 100 URLs at time - and up to 1,000 sites total for each account.

Review SharePoint requirements

Review the following prerequisites to make sure you are ready to activate BlueXP classification on a SharePoint account.

- You must have the Admin user login credentials for the SharePoint account that provides read access to all SharePoint sites.
  - For SharePoint Online you can use a non-Admin account, but that user must have permission to access all the SharePoint sites that you want to scan.
- For SharePoint On-Premise, you'll also need the URL of the SharePoint Server.
- You will need a line-separated list of the SharePoint site URLs for all the data you want to scan.

Deploy the BlueXP classification instance

Deploy BlueXP classification if there isn't already an instance deployed.

- For SharePoint Online, BlueXP classification can be deployed in the cloud.
- For SharePoint On-Premises, BlueXP classification can be installed in an on-premises location that has internet access or in an on-premises location that does not have internet access.

When BlueXP classification is installed in a site without internet access, the BlueXP Connector also must be installed in that same site without internet access. Learn more.

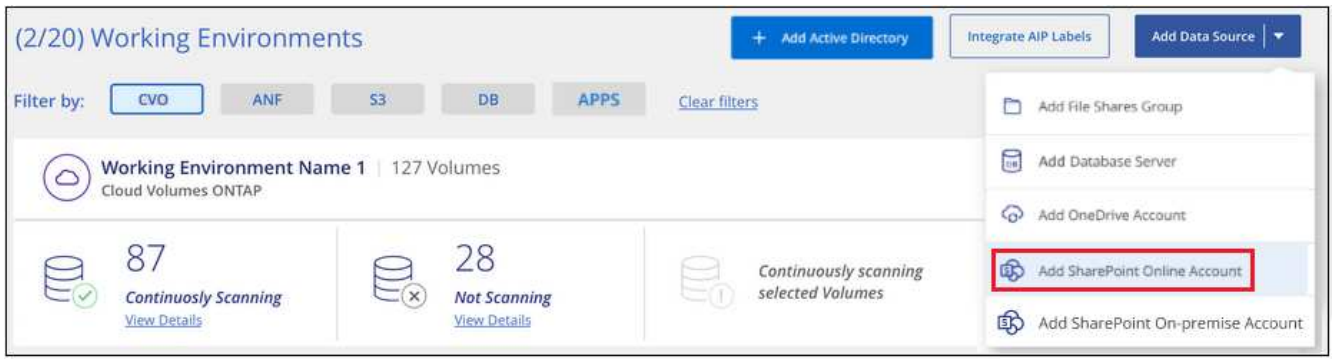
Upgrades to BlueXP classification software is automated as long as the instance has internet connectivity.

Add a SharePoint Online account

Add the SharePoint Online account where the user files reside.

Steps

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



2. In the Add a SharePoint Online Account dialog, click **Sign in to SharePoint**.
3. In the Microsoft page that appears, select the SharePoint account and enter the user and password (Admin user or other user with access to the SharePoint sites), then click **Accept** to allow BlueXP classification to read data from this account.

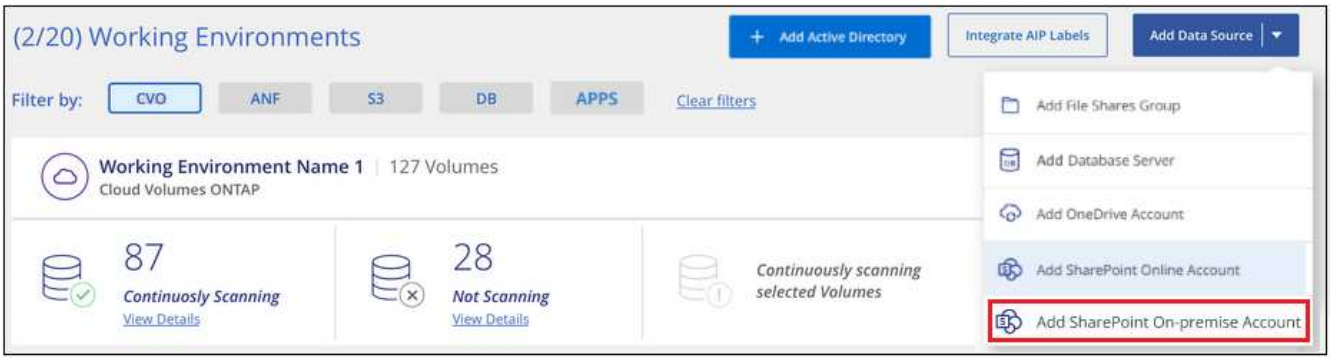
The SharePoint Online account is added to the list of working environments.

Add a SharePoint On-premise account

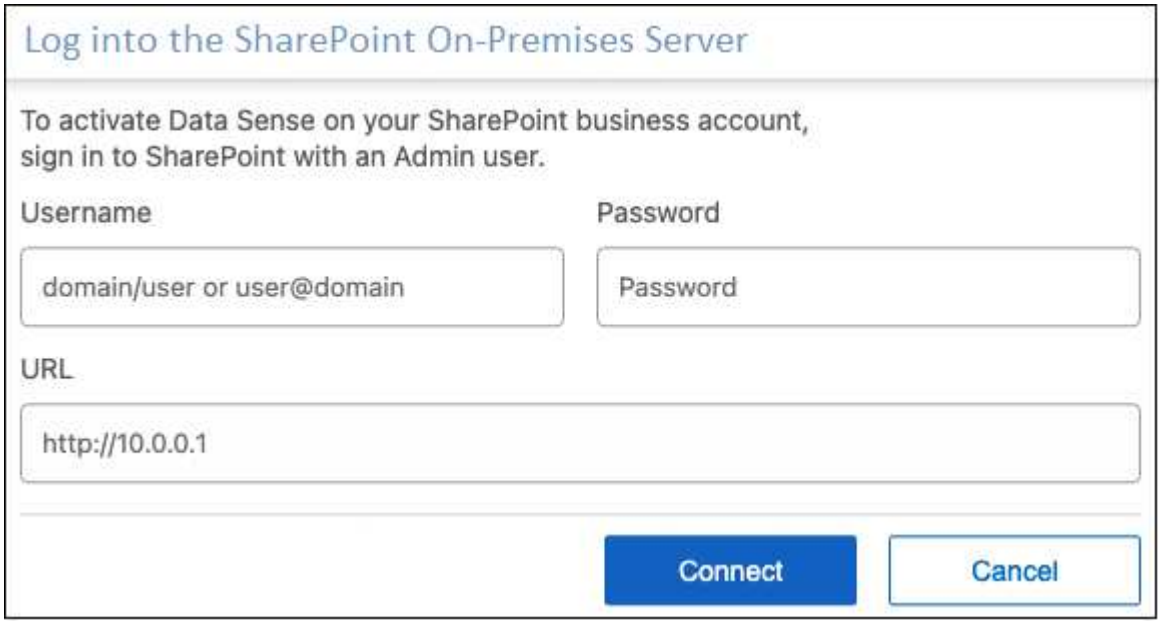
Add the SharePoint On-premise account where the user files reside.

Steps

1. From the Working Environments Configuration page, click **Add Data Source** > **Add SharePoint On-premise Account**.



2. In the Log into the SharePoint On-Premise Server dialog, enter the following information:
   - Admin user in the format "domain/user" or "user@domain", and admin password
   - URL of the SharePoint Server



3. Click **Connect**.

The SharePoint On-premise account is added to the list of working environments.

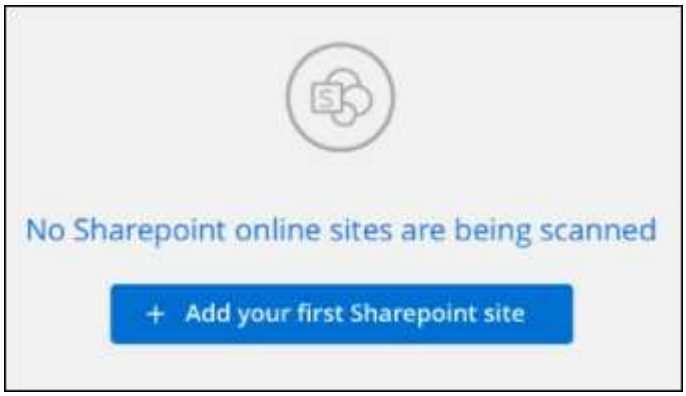
Add SharePoint sites to compliance scans

You can add individual SharePoint sites, or up to 1,000 SharePoint sites in the account, so that the associated files will be scanned by BlueXP classification. The steps are the same whether you are adding SharePoint Online or SharePoint On-premise sites.

Steps

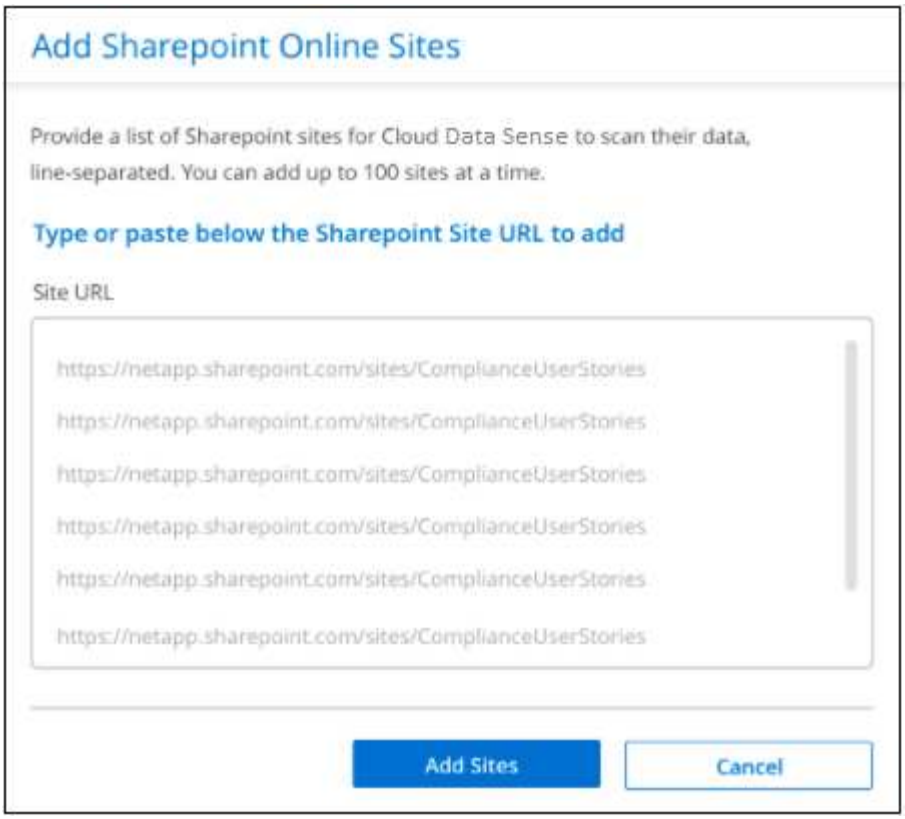
1. From the *Configuration* page, click the **Configuration** button for the SharePoint account.

2. If this is the first time adding sites for this SharePoint account, click **Add your first SharePoint site**.



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

3. Add the URLs for the sites whose files you want to scan - one URL per line (up to 100 maximum per session) - and click **Add Sites**.



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

If the dialog lists any sites that could not be added, capture this information so that you can resolve the issue. In some cases you can re-add the site with a corrected URL.

4. If you need to add more than 100 sites for this account, just click **Add SharePoint Sites** again until you have added all your sites for this account (up to 1,000 sites total for each account).
5. Enable mapping-only scans, or mapping and classification scans, on the files in the SharePoint sites.

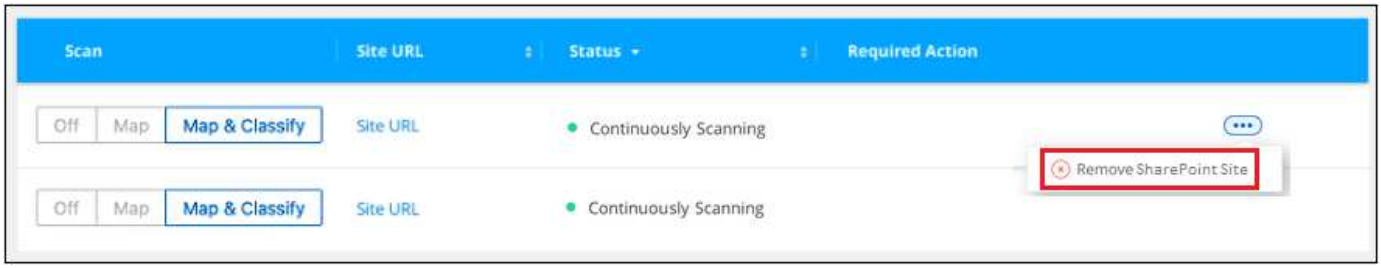
To:	Do this:
Enable mapping-only scans on files	Click **Map**
Enable full scans on files	Click **Map & Classify**
Disable scanning on files	Click **Off**

Result

BlueXP classification starts scanning the files in the SharePoint sites you added, and the results are displayed in the Dashboard and in other locations.

Remove a SharePoint site from compliance scans

If you remove a SharePoint site in the future, or decide not to scan files in a SharePoint site, you can remove individual SharePoint sites from having their files scanned at any time. Just click **Remove SharePoint Site** from the Configuration page.



Note that you can delete the entire SharePoint account from BlueXP classification if you no longer want to scan any user data from the SharePoint account.

Scan Google Drive accounts

Complete a few steps to start scanning user files in your Google Drive accounts with BlueXP classification.

NOTE This information is relevant only for BlueXP classification legacy versions 1.30 and earlier.

Quick start

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

1 **Review Google Drive prerequisites**

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

2 **Deploy BlueXP classification**

Deploy BlueXP classification if there isn't already an instance deployed.

3 **Log into the Google Drive account**

Using Admin user credentials, log into the Google Drive account that you want to access so that it is added as a new data source.

4 **Select the type of scanning for the user files**

Select the type of scanning you want to perform on the user files; mapping or mapping and classifying.

Review Google Drive requirements

Review the following prerequisites to make sure you are ready to enable BlueXP classification on a Google Drive account.

- You must have the Admin login credentials for the Google Drive account that provides read access to the user's files

Current restrictions

The following BlueXP classification features are not currently supported with Google Drive files:

- When viewing files in the Data Investigation page, the actions in the button bar aren't active. You can't copy, move, delete, etc. any files.
- Permissions can't be identified within files in Google Drive, so no permission information is displayed in the Investigation page.

Deploy BlueXP classification

Deploy BlueXP classification if there isn't already an instance deployed.

BlueXP classification can be deployed in the cloud or in an on-premises location that has internet access.

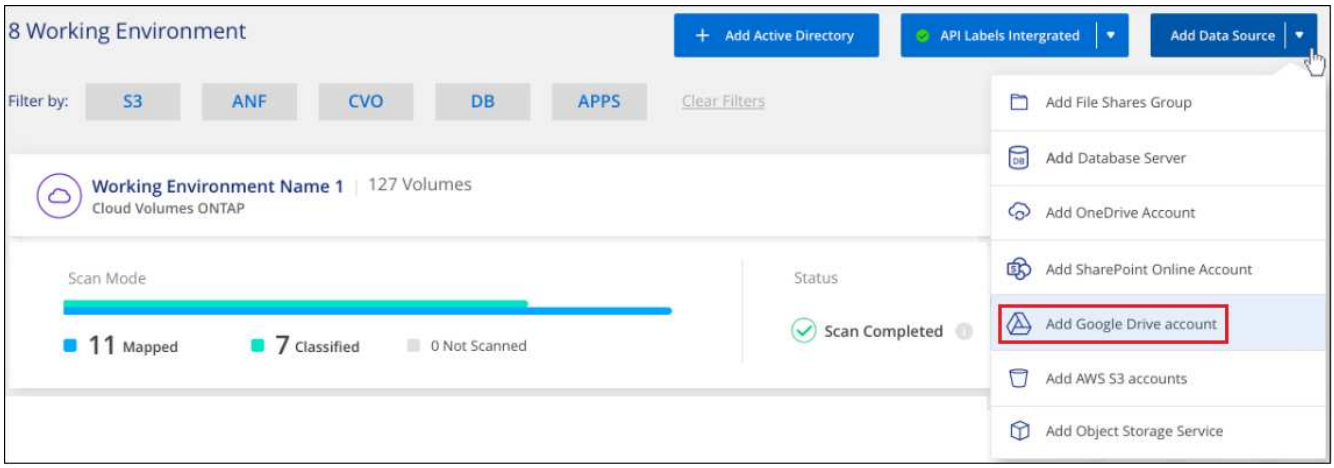
Upgrades to BlueXP classification software is automated as long as the instance has internet connectivity.

Add the Google Drive account

Add the Google Drive account where the user files reside. If you want to scan files from multiple users, you'll need to run through this step for each user.

Steps

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



2. In the Add a Google Drive Account dialog, click **Sign in to Google Drive**.
3. In the Google page that appears, select the Google Drive account and enter the required Admin user and password, then click **Accept** to allow BlueXP classification to read data from this account.

The Google Drive account is added to the list of working environments.

Select the type of scanning for user data

Select the type of scanning that BlueXP classification will perform on the user's data.

Steps

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

2. Enable mapping-only scans, or mapping and classification scans, on the files in the Google Drive account.

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

Result

BlueXP classification starts scanning the files in the Google Drive account you added, and the results are displayed in the Dashboard and in other locations.

Remove a Google Drive account from compliance scans

Since only a single user's Google Drive files are part of a single Google Drive account, if you want to stop scanning files from a user's Google Drive account, then you should delete the Google Drive account from BlueXP classification.

Scan object storage that uses S3 protocol

Complete a few steps to start scanning data within object storage directly with BlueXP classification. BlueXP classification 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, Amazon S3, and more.

NOTE This information is relevant only for BlueXP classification legacy versions 1.30 and earlier.

Quick start

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

1 **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 BlueXP classification can access the buckets.

Deploy the BlueXP classification instance

Deploy BlueXP classification if there isn't already an instance deployed.

Add the Object Storage Service

Add the object storage service to BlueXP classification.

Select the buckets to scan

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

Reviewing object storage requirements

Review the following prerequisites to make sure that you have a supported configuration before you enable BlueXP classification.

- 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 BlueXP classification can access the buckets.

Deploying the BlueXP classification instance

Deploy BlueXP classification if there isn't already an instance deployed.

If you are scanning data from S3 object storage that is accessible over the internet, you can deploy BlueXP classification in the cloud or deploy BlueXP classification in an on-premises location that has internet access.

If you are scanning data from S3 object storage that has been installed in a dark site that has no internet access, you need to deploy BlueXP classification in the same on-premises location that has no internet access. This also requires that the BlueXP Connector is deployed in that same on-premises location.

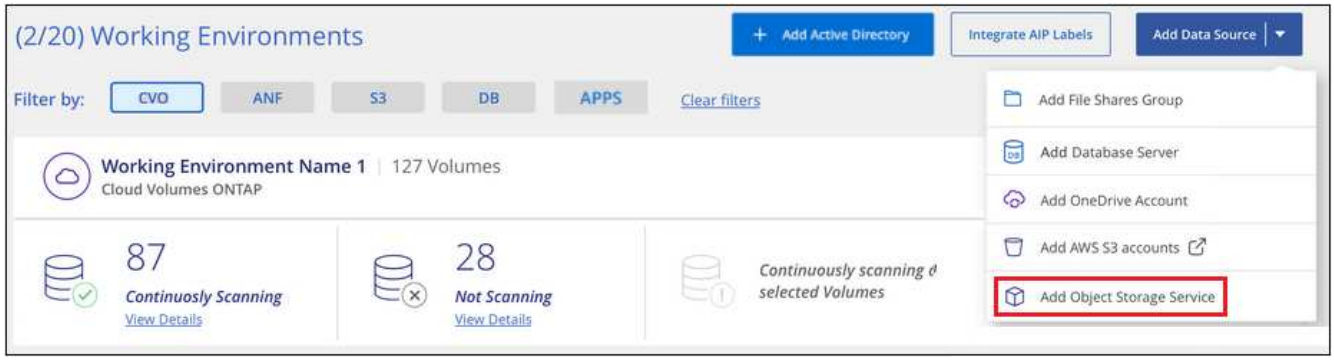
Upgrades to BlueXP classification software is automated as long as the instance has internet connectivity.

Adding the object storage service to BlueXP classification

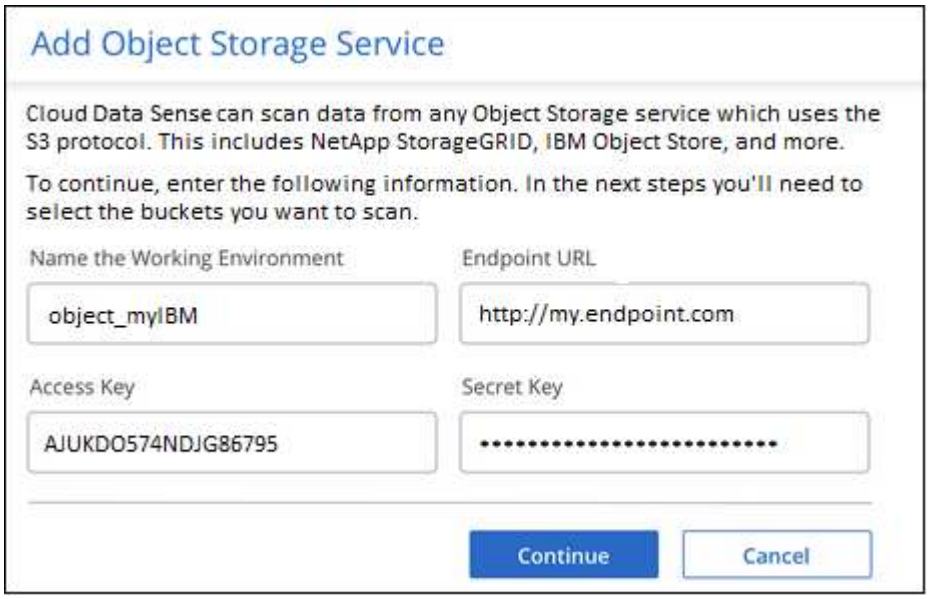
Add the object storage service.

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 BlueXP classification can access the buckets in the object storage.



Result

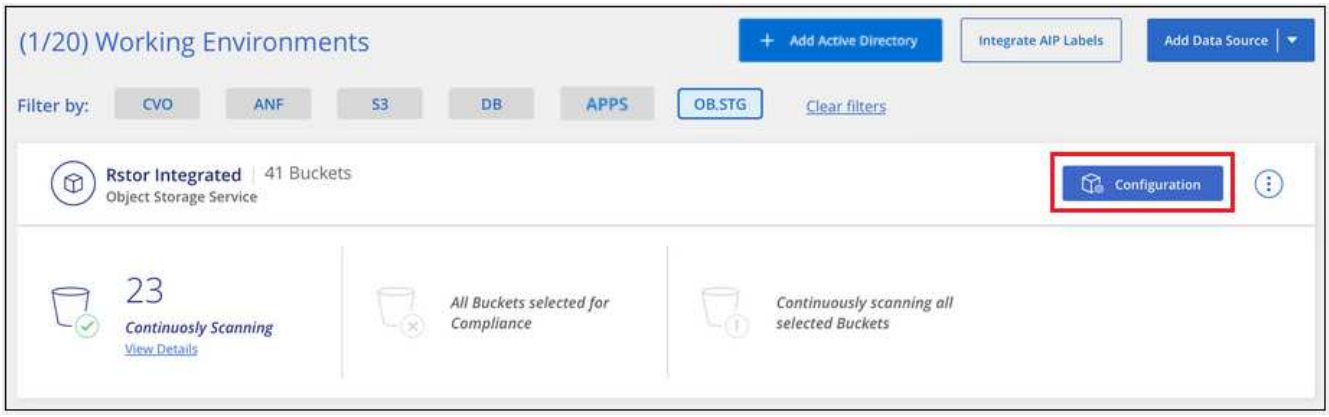
The new Object Storage Service is added to the list of working environments.

Enabling and disabling compliance scans on object storage buckets

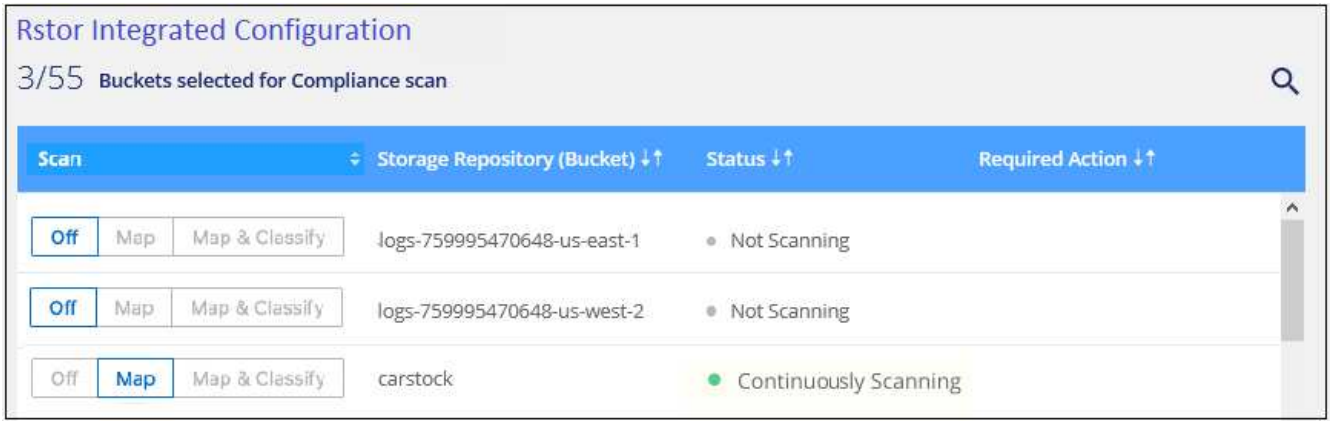
After you enable BlueXP classification on your Object Storage Service, the next step is to configure the buckets that you want to scan. BlueXP classification discovers those buckets and displays them in the working environment you created.

Steps

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



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



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

Result

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

Copyright information

Trademark information