



Get started

BlueXP classification

NetApp
August 11, 2025

Table of Contents

Get started	1
Learn about BlueXP classification	1
Features	1
Supported working environments and data sources	2
Cost	2
The BlueXP classification instance	3
How BlueXP classification scanning works	4
What's the difference between Mapping and Classification scans	5
Information that BlueXP classification categorizes	5
Networking overview	6
Access BlueXP classification	6
Deploy BlueXP classification	7
Which BlueXP classification deployment should you use?	7
Deploy BlueXP classification in the cloud using BlueXP	7
Install BlueXP classification on a host that has internet access	17
Install BlueXP classification on a Linux host with no internet access	27
Check that your Linux host is ready to install BlueXP classification	36
Activate scanning on your data sources	41
Scan data sources overview with BlueXP classification	41
Scan Azure NetApp Files volumes with BlueXP classification	45
Scan Amazon FSx for ONTAP volumes with BlueXP classification	48
Scan Cloud Volumes ONTAP and on-premises ONTAP volumes with BlueXP classification	53
Scan database schemas with BlueXP classification	57
Scan file shares with BlueXP classification	59
Scan StorageGRID data with BlueXP classification	63
Integrate your Active Directory with BlueXP classification	65
Supported data sources	66
Connect to your Active Directory server	66
Manage your Active Directory integration	68

Get started

Learn about BlueXP classification

BlueXP classification (Cloud Data Sense) is a data governance service for BlueXP that scans your corporate on-premises and cloud data sources to map and classify data, and to identify private information. This can help reduce your security and compliance risk, decrease storage costs, and assist with your data migration projects.



Beginning with version 1.31, BlueXP classification is available as a core capability with BlueXP. There's no additional charge. No Classification license or subscription is required. If you've been using legacy version 1.30 or earlier, that version is available until your subscription expires. [See a list of deprecated features.](#)

Features

BlueXP classification uses artificial intelligence (AI), natural language processing (NLP), and machine learning (ML) to understand the content that it scans in order to extract entities and categorize the content accordingly. This allows BlueXP classification to provide the following areas of functionality.

[Learn more about the use cases for BlueXP classification.](#)

Maintain compliance

BlueXP classification provides several tools that can help with your compliance efforts. You can use BlueXP classification to:

- Identify Personal Identifiable Information (PII).
- Identify a wide scope of sensitive personal information as required by GDPR, CCPA, PCI, and HIPAA privacy regulations.
- Respond to Data Subject Access Requests (DSAR) based on name or email address.

Strengthen security

BlueXP classification can identify data that is potentially at risk for being accessed for criminal purposes. You can use BlueXP classification to:

- Identify all the files and directories (shares and folders) with open permissions that are exposed to your entire organization or to the public.
- Identify sensitive data that resides outside of the initial, dedicated location.
- Comply with data retention policies.
- Use *Policies* to automatically detect new security issues so security staff can take action immediately.

Optimize storage usage

BlueXP classification provides tools that can help with your storage total cost of ownership (TCO). You can use BlueXP classification to:

- Increase storage efficiency by identifying duplicate or non-business-related data.
- Save storage costs by identifying inactive data that you can tier to less expensive object storage. [Learn more about tiering from Cloud Volumes ONTAP systems.](#) [Learn more about tiering from on-premises](#)

Supported working environments and data sources

BlueXP classification can scan and analyze structured and unstructured data from the following types of working environments and data sources:

Working environments

- Amazon FSx for ONTAP
- Azure NetApp Files
- Cloud Volumes ONTAP (deployed in AWS, Azure, or GCP)
- On-premises ONTAP clusters
- StorageGRID

Data sources

- NetApp file shares
- Databases:
 - Amazon Relational Database Service (Amazon RDS)
 - MongoDB
 - MySQL
 - Oracle
 - PostgreSQL
 - SAP HANA
 - SQL Server (MSSQL)

BlueXP classification supports NFS versions 3.x, 4.0, and 4.1, and CIFS versions 1.x, 2.0, 2.1, and 3.0.

Cost

BlueXP classification is free to use. No Classification license or paid subscription is required.

Infrastructure costs

- Installing BlueXP classification 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 BlueXP classification on an on-premises system.
- BlueXP classification requires that you have deployed a BlueXP Connector. In many cases you already have a Connector because of other storage and services you are using in BlueXP. 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](#). There is no cost if you install the Connector on an on-premises system.

Data transfer costs

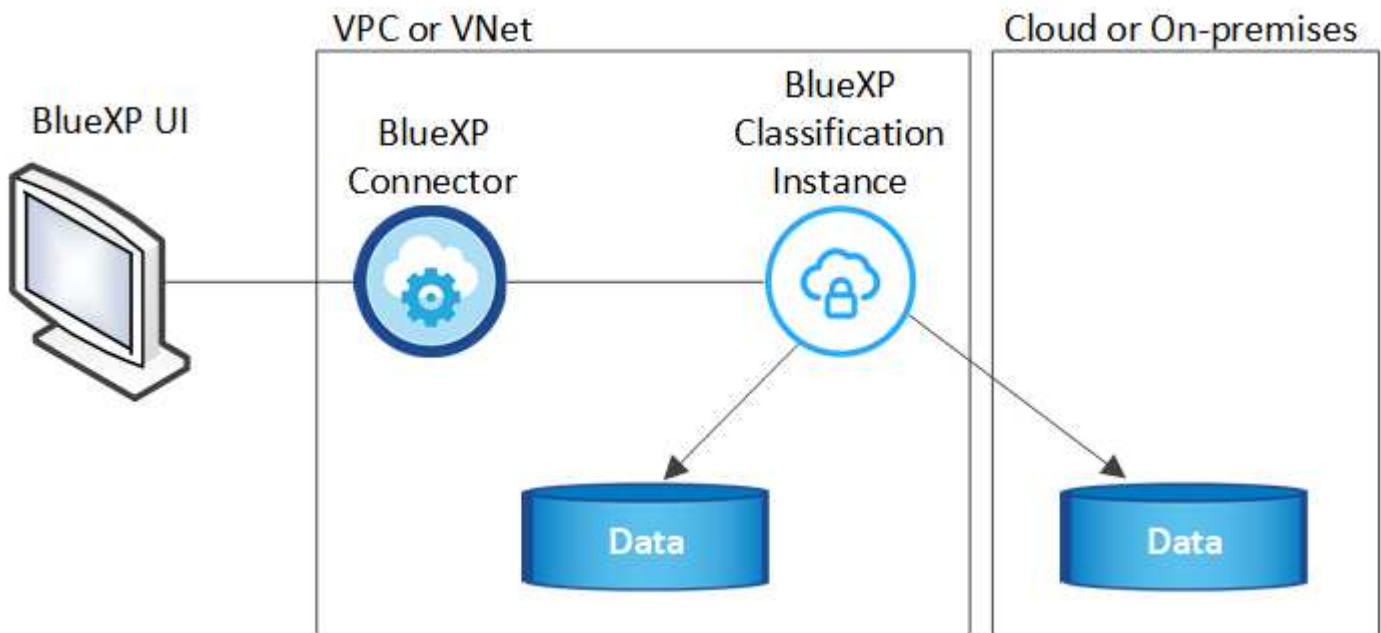
Data transfer costs depend on your setup. If the BlueXP classification instance and data source are in the same Availability Zone and region, then there are no data transfer costs. But if the data source, such as a Cloud Volumes ONTAP system, is in a *different* Availability Zone or region, then you'll be charged by your

cloud provider for data transfer costs. See these links for more details:

- [AWS: Amazon Elastic Compute Cloud \(Amazon EC2\) Pricing](#)
- [Microsoft Azure: Bandwidth Pricing Details](#)
- [Google Cloud: Storage Transfer Service pricing](#)

The BlueXP classification instance

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



Note the following about the default instance:

- In AWS, BlueXP classification runs on an [m6i.4xlarge instance](#) with a 500 GiB GP2 disk. The operating system image is Amazon Linux 2. When deployed in AWS, you can choose a smaller instance size if you are scanning a small amount of data.
- In Azure, BlueXP classification runs on a [Standard_D16s_v3 VM](#) with a 500 GiB disk. The operating system image is Ubuntu 22.04.
- In GCP, BlueXP classification runs on an [n2-standard-16 VM](#) with a 500 GiB Standard persistent disk. The operating system image is Ubuntu 22.04.
- In regions where the default instance isn't available, BlueXP classification runs on an alternate instance. [See the alternate instance types.](#)
- The instance is named *CloudCompliance* with a generated hash (UUID) concatenated to it. For example: *CloudCompliance-16bb6564-38ad-4080-9a92-36f5fd2f71c7*
- Only one BlueXP classification instance is deployed per Connector.

You can also deploy BlueXP classification on a Linux host on your premises or on a host in your preferred cloud provider. The software functions exactly the same way regardless of which installation method you choose. Upgrades of BlueXP classification software are automated as long as the instance has internet access.



The instance should remain running at all times because BlueXP classification continuously scans the data.

Deploy on different instance types

Review the following specifications for instance types:

System size	Specs	Limitations
Extra Large	32 CPUs, 128 GB RAM, 1 TiB SSD	Can scan up to 500 million files.
Large (default)	16 CPUs, 64 GB RAM, 500 GiB SSD	Can scan up to 250 million files.

When deploying BlueXP classification in Azure or GCP, email ng-contact-data-sense@netapp.com for assistance if you want to use a smaller instance type.

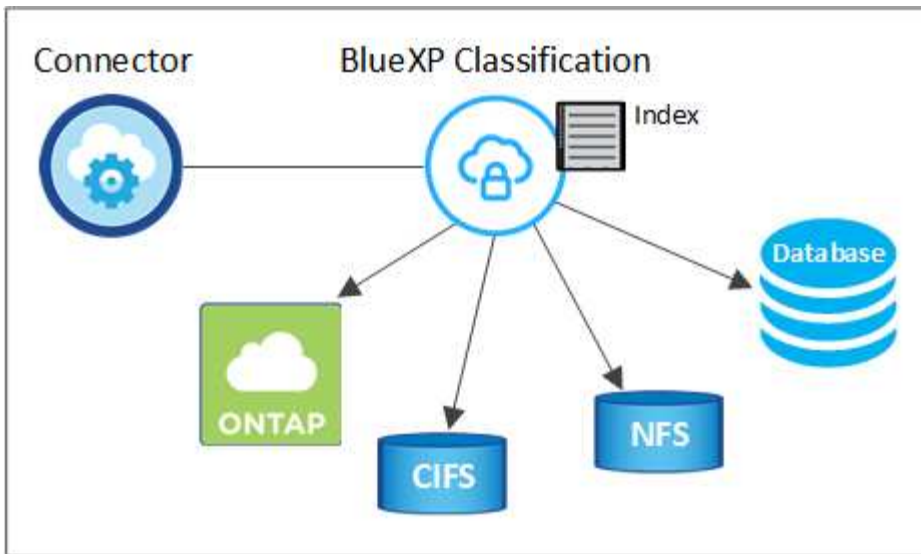
How BlueXP classification scanning works

At a high-level, BlueXP classification scanning works like this:

1. You deploy an instance of BlueXP classification in BlueXP.
2. You enable high-level mapping (called *Mapping only* scans) or deep-level scanning (called *Map & Classify* scans) on one or more data sources.
3. BlueXP classification scans the data using an AI learning process.
4. You use the provided dashboards and reporting tools to help in your compliance and governance efforts.

After you enable BlueXP classification and select the repositories that you want to scan (these are the volumes, database schemas, or other user data), it immediately starts scanning the data to identify personal and sensitive data. You should focus on scanning live production data in most cases instead of backups, mirrors, or DR sites. Then BlueXP classification 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.

BlueXP classification 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.



After the initial scan, BlueXP classification continuously scans your data in a round-robin fashion 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 or the database schema level.



BlueXP classification does not impose a limit on the amount of data it can scan. Each Connector supports scanning and displaying 500 TiB of data. To scan more than 500 TiB of data, [install another Connector](#) then [deploy another classification instance](#).

The BlueXP UI displays data from a single connector. For tips on viewing data from multiple Connectors, see [Work with multiple Connectors](#).

What's the difference between Mapping and Classification scans

You can conduct two types of scans in BlueXP classification:

- **Mapping-only scans** provide only a high-level overview of your data and are performed on selected data sources. Mapping-only scans take less time than map and classify scans because they do not access files to see the data inside. You might want to do this initially to identify areas of research and then perform a Map & Classify scan on those areas.
- **Map & Classify scans** provide deep-level scanning of your data.

For details about the differences between Mapping and Classification scans, see [What's the difference between Mapping and Classification scans?](#).

Information that BlueXP classification categorizes

BlueXP classification collects, indexes, and assigns categories to the following data:

- **Standard metadata** about files: the file type, its size, creation and modification dates, and so on.
- **Personal data:** Personally identifiable information (PII) such as email addresses, identification numbers, or credit card numbers, which BlueXP classification identifies using specific words, strings, and patterns in the files. [Learn more about personal data](#).
- **Sensitive personal data:** Special types of sensitive personal information (SPII), such as health data, ethnic origin, or political opinions, as defined by General Data Protection Regulation (GDPR) and other privacy regulations. [Learn more about sensitive personal data](#).

- **Categories:** BlueXP classification 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:** BlueXP classification takes the data that it scanned and breaks it down by file type. [Learn more about types.](#)
- **Name entity recognition:** BlueXP classification uses AI to extract people's natural names from documents. [Learn about responding to Data Subject Access Requests.](#)

Networking overview

BlueXP classification deploys a single server, or cluster, wherever you choose — in the cloud or on premises. The servers connect via standard protocols to the data sources and index the findings in an Elasticsearch cluster, which is also deployed on the same servers. This enables support for multi-cloud, cross-cloud, private cloud, and on-premises environments.

BlueXP deploys the BlueXP classification instance with a security group that enables inbound HTTP connections from the Connector instance.

When you use BlueXP in SaaS mode, the connection to BlueXP is served over HTTPS, and the private data sent between your browser and the BlueXP classification instance are secured with end-to-end encryption using TLS 1.2, which means NetApp and third parties can't read it.

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

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

Access BlueXP classification

You can access the BlueXP classification service through NetApp BlueXP.

To sign in to BlueXP, you can use your NetApp Support Site credentials or you can sign up for a NetApp cloud login using your email and a password. [Learn more about logging in to BlueXP.](#)

Specific tasks require specific BlueXP user roles. [Learn about BlueXP access roles for all services.](#)

Before you begin

- [You should add a BlueXP Connector.](#)
- [Understand which BlueXP classification deployment style suits your workload.](#)

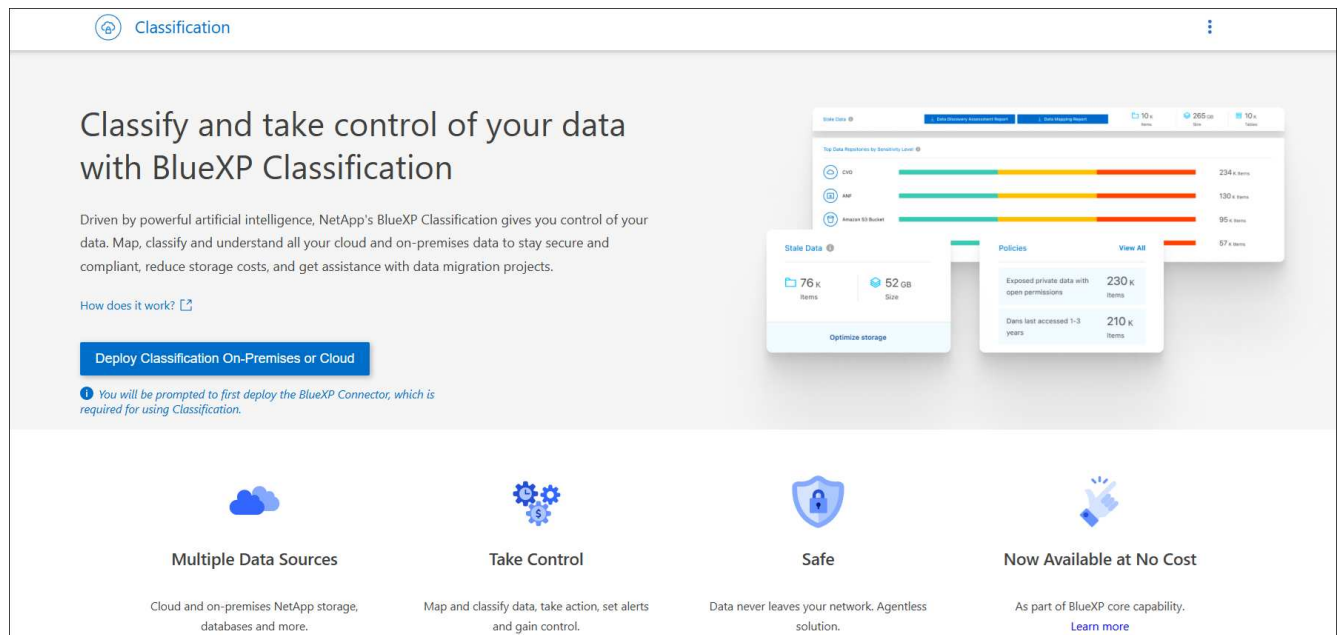
Steps

1. In a web browser, navigate to the [BlueXP console](#).

The NetApp BlueXP login page appears.

2. Sign in to BlueXP.
3. From the BlueXP left navigation menu, select **Governance > Classification**.
4. If this is your first time accessing BlueXP classification, the landing page appears.

Select **Deploy Classification On-Premises or Cloud** to begin deploying your classification instance. For more information, see [Which BlueXP classification deployment should you use?](#)



Otherwise, the BlueXP classification Dashboard appears.

Deploy BlueXP classification

Which BlueXP classification deployment should you use?

You can deploy BlueXP classification in different ways. Learn which method meets your needs.

BlueXP classification can be deployed in the following ways:

- [Deploy in the cloud using BlueXP](#). BlueXP will deploy the BlueXP classification instance in the same cloud provider network as the BlueXP Connector.
- [Install on a Linux host with internet access](#). Install BlueXP classification on a Linux host in your network, or on a Linux host in the cloud, that has internet access. This type of installation may be a good option if you prefer to scan on-premises ONTAP systems using a BlueXP classification instance that's also located on premises—but this is not a requirement.
- [Install on a Linux host in an on-premises site without internet access](#), also known as *private mode*. This type of installation, which uses an installation script, has no connectivity to the BlueXP SaaS layer.

Both the installation on a Linux host with internet access and the on-premises installation on a Linux host without internet access use an installation script. The script starts by checking if the system and environment meet the prerequisites. If the prerequisites are met, the installation starts. If you would like to verify the prerequisites independently of running the BlueXP classification installation, there is a separate software package you can download that only tests for the prerequisites.

Refer to [Check that your Linux host is ready to install BlueXP classification](#).

Deploy BlueXP classification in the cloud using BlueXP

Complete a few steps to deploy BlueXP classification in the cloud. BlueXP will deploy the BlueXP classification instance in the same cloud provider network as the BlueXP

Connector.

Note that you can also [install BlueXP classification on a Linux host that has internet access](#). This type of installation may be a good option if you prefer to scan on-premises ONTAP systems using a BlueXP classification instance that's also located on premises — but this is not a requirement. The 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 now. See [creating a Connector in AWS](#), [creating a Connector in Azure](#), or [creating a Connector in GCP](#).

You can also [install the Connector on-premises](#) on a Linux host in your network or on a Linux host 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 BlueXP classification over port 443, and more. [See the complete list](#).

3

Deploy BlueXP classification

Launch the installation wizard to deploy the BlueXP classification instance in the cloud.

Create a Connector

If you don't already have a Connector, create a Connector in your cloud provider. See [creating a Connector in AWS](#) or [creating a Connector in Azure](#), or [creating a Connector in GCP](#). In most cases you will probably have a Connector set up before you attempt to activate BlueXP classification because most [BlueXP 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 a specific cloud provider:

- When scanning data in Cloud Volumes ONTAP in AWS or Amazon FSx for ONTAP 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.
 - For Azure NetApp Files, it must be deployed in the same region as the volumes you wish to scan.
- When scanning data in Cloud Volumes ONTAP in GCP, you use a Connector in GCP.

On-prem ONTAP systems, NetApp file shares, and databases can be scanned when using any of these cloud Connectors.

Note that you can also [install the Connector on-premises](#) on a Linux host in your network or in the cloud. Some users planning to install BlueXP classification 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](#).



BlueXP classification does not impose a limit on the amount of data it can scan. Each Connector supports scanning and displaying 500 TiB of data. To scan more than 500 TiB of data, [install another Connector](#) then [deploy another classification instance](#). The BlueXP UI displays data from a single connector. For tips on viewing data from multiple Connectors, see [Work with multiple Connectors](#).

Government region support

BlueXP classification is supported when the Connector is deployed in a Government region (AWS GovCloud, Azure Gov, or Azure DoD). When deployed in this manner, BlueXP classification has the following restrictions:

[See more information about deploying the Connector in a Government region.](#)

Review prerequisites

Review the following prerequisites to make sure that you have a supported configuration before you deploy BlueXP classification in the cloud. When you deploy BlueXP classification in the cloud, it's located in the same subnet as the Connector.

Enable outbound internet access from BlueXP classification

BlueXP classification requires outbound internet access. If your virtual or physical network uses a proxy server for internet access, ensure that the BlueXP classification instance has outbound internet access to contact the following endpoints. The proxy must be non-transparent. Transparent proxies are not currently supported.

Review the appropriate table below depending on whether you are deploying BlueXP classification in AWS, Azure, or GCP.

Required endpoints for AWS

Endpoints	Purpose
https://api.bluexp.netapp.com	Communication with the BlueXP service, which includes NetApp accounts.
https://netapp-cloud-account.auth0.com https://auth0.com	Communication with the BlueXP website for centralized user authentication.
https://cloud-compliance-support-netapp.s3.us-west-2.amazonaws.com https://hub.docker.com https://auth.docker.io https://registry-1.docker.io https://index.docker.io/ https://dseasb33srrn.cloudfront.net/ https://production.cloudflare.docker.com/	Provides access to software images, manifests, and templates.
https://kinesis.us-east-1.amazonaws.com	Enables NetApp to stream data from audit records.
https://cognito-idp.us-east-1.amazonaws.com https://cognito-identity.us-east-1.amazonaws.com https://user-feedback-store-prod.s3.us-west-2.amazonaws.com https://customer-data-production.s3.us-west-2.amazonaws.com	Enables BlueXP classification to access and download manifests and templates, and to send logs and metrics.

Required endpoints for Azure

Endpoints	Purpose
https://api.bluexp.netapp.com	Communication with the BlueXP service, which includes NetApp accounts.
https://netapp-cloud-account.auth0.com https://auth0.com	Communication with the BlueXP website for centralized user authentication.
https://support.compliance.api.bluexp.netapp.com/ https://hub.docker.com https://auth.docker.io https://registry-1.docker.io https://index.docker.io/ https://dseasb33srrn.cloudfront.net/ https://production.cloudflare.docker.com/	Provides access to software images, manifests, templates, and to send logs and metrics.
https://support.compliance.api.bluexp.netapp.com/	Enables NetApp to stream data from audit records.

Required endpoints for GCP

Endpoints	Purpose
https://api.bluexp.netapp.com	Communication with the BlueXP service, which includes NetApp accounts.

Endpoints	Purpose
https://netapp-cloud-account.auth0.com https://auth0.com	Communication with the BlueXP website for centralized user authentication.
https://support.compliance.api.blueexp.netapp.com/ https://hub.docker.com https://auth.docker.io https://registry-1.docker.io https://index.docker.io/ https://dseasb33srrn.cloudfront.net/ https://production.cloudflare.docker.com/	Provides access to software images, manifests, templates, and to send logs and metrics.
https://support.compliance.api.blueexp.netapp.com/	Enables NetApp to stream data from audit records.

Ensure that BlueXP has the required permissions

Ensure that BlueXP has permissions to deploy resources and create security groups for the BlueXP classification instance.

- [Google Cloud permissions](#)
- [AWS permissions](#)
- [Azure permissions](#)

Ensure that the BlueXP Connector can access BlueXP classification

Ensure connectivity between the Connector and the BlueXP classification instance. The security group for the Connector must allow inbound and outbound traffic over port 443 to and from the BlueXP classification instance. This connection enables deployment of the BlueXP classification instance and enables you to view information in the Compliance and Governance tabs. BlueXP classification is supported in Government regions in AWS and Azure.

Additional inbound and outbound security group rules are required for AWS and AWS GovCloud deployments. See [Rules for the Connector in AWS](#) for details.

Additional inbound and outbound security group rules are required for Azure and Azure Government deployments. See [Rules for the Connector in Azure](#) for details.

Ensure that you can keep BlueXP classification running

The BlueXP classification instance needs to stay on to continuously scan your data.

Ensure web browser connectivity to BlueXP classification

After BlueXP classification is enabled, ensure that users access the BlueXP interface from a host that has a connection to the BlueXP classification instance.

The BlueXP classification 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 BlueXP must have a connection to that private IP address. That connection can come from a direct connection to your cloud provider (for example, a VPN), or from a host that's inside the same network as the BlueXP classification instance.

Check your vCPU limits

Ensure that your cloud provider's vCPU limit allows for the deployment of an instance with the necessary number of cores. You'll need to verify the vCPU limit for the relevant instance family in the region where BlueXP is running. [See the required instance types](#).

See the following links for more details on vCPU limits:

- [AWS documentation: Amazon EC2 service quotas](#)
- [Azure documentation: Virtual machine vCPU quotas](#)
- [Google Cloud documentation: Resource quotas](#)

Deploy BlueXP classification in the cloud

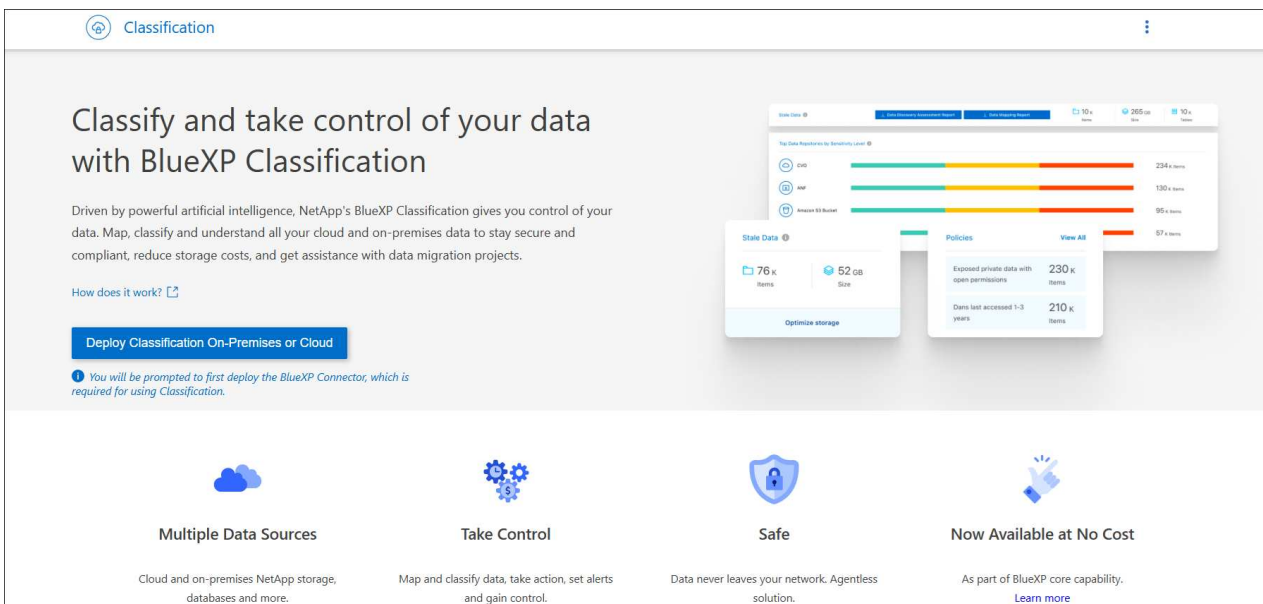
Follow these steps to deploy an instance of BlueXP classification in the cloud. The Connector will deploy the instance in the cloud, and then install BlueXP classification software on that instance.

In regions where the default instance type isn't available, BlueXP classification runs on an [alternate instance type](#).

Deploy in AWS

Steps

1. From the BlueXP left navigation menu, select **Governance > Classification**.
2. Select **Deploy Classification On-Premises or Cloud**.



3. From the *Installation* page, select **Deploy > Deploy** to use the "Large" instance size and start the cloud deployment wizard.
4. The wizard displays progress as it goes through the deployment steps. It will stop and prompt for input if it runs into any issues.



5. When the instance is deployed and BlueXP classification is installed, select **Continue to configuration** to go to the *Configuration* page.

Deploy in Azure

Steps

1. From the BlueXP left navigation menu, select **Governance > Classification**.
2. Select **Deploy Classification On-Premises or Cloud**.

Classification

Classify and take control of your data with BlueXP Classification

Driven by powerful artificial intelligence, NetApp's BlueXP Classification gives you control of your data. Map, classify and understand all your cloud and on-premises data to stay secure and compliant, reduce storage costs, and get assistance with data migration projects.

How does it work? [🔗](#)

[Deploy Classification On-Premises or Cloud](#)

❗ You will be prompted to first deploy the BlueXP Connector, which is required for using Classification.

Multiple Data Sources

Cloud and on-premises NetApp storage, databases and more.

Take Control

Map and classify data, take action, set alerts and gain control.

Safe

Data never leaves your network. Agentless solution.

Now Available at No Cost

As part of BlueXP core capability. [Learn more](#)

3. Select **Deploy** to start the cloud deployment wizard.

Install your Data Sense instance

Select your preferred deployment location:

[Learn more about deploying Data Sense](#) [🔗](#)

Cloud Environment

I want BlueXP to deploy the instance and install Data Sense

[Deploy](#)

> BlueXP will deploy a new machine automatically in the chosen cloud environment.
 > You will be taken to an installation wizard where you can configure your Data Sense installation.

I deployed an instance and I'm ready to install Data Sense

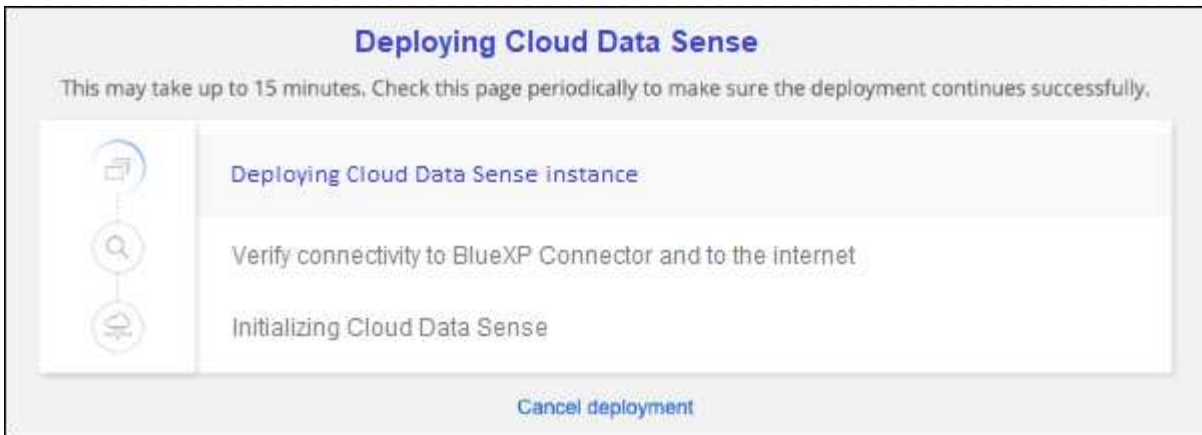
[Deploy](#)

On Premise

I prepared a local machine and I'm ready to install Data Sense

[Deploy](#)

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

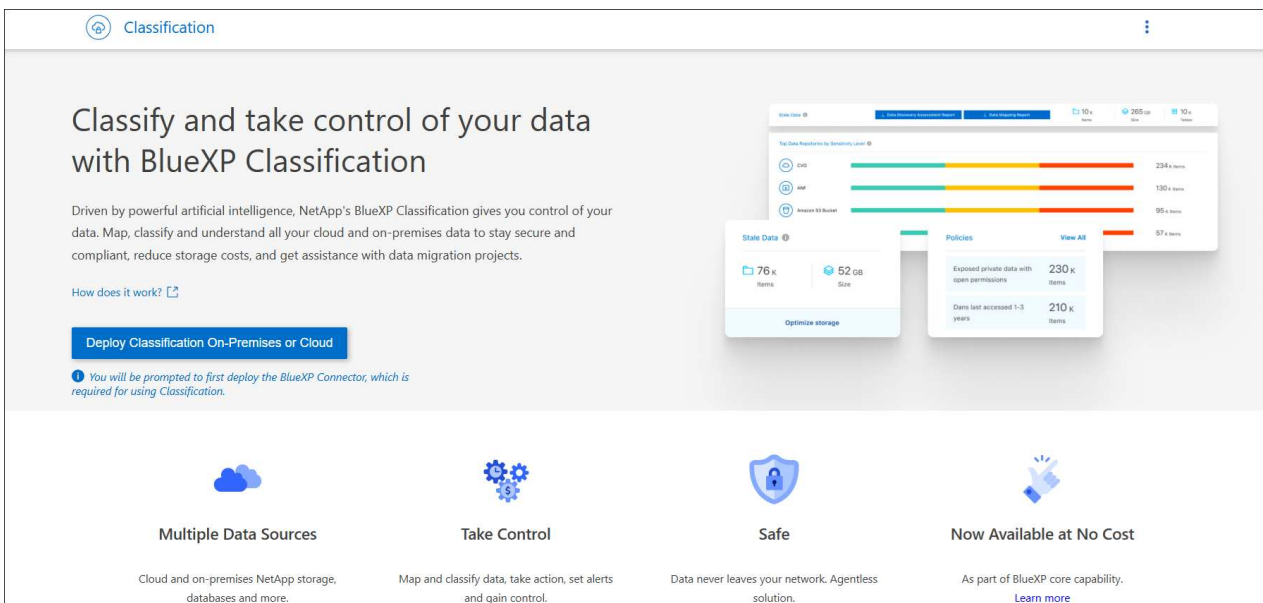


5. When the instance is deployed and BlueXP classification is installed, select **Continue to configuration** to go to the *Configuration* page.

Deploy in Google Cloud

Steps

1. From the BlueXP left navigation menu, select **Governance > Classification**.
2. Select **Deploy Classification On-Premises or Cloud**.





3. Select **Deploy** to start the cloud deployment wizard.



Install your Data Sense instance

Select your preferred deployment location:



[Learn more about deploying Data Sense](#)

Cloud Environment

 **I want BlueXP to deploy the instance and install Data Sense** Deploy 

 **I deployed an instance and I'm ready to install Data Sense** Deploy 

On Premise

 **I prepared a local machine and I'm ready to install Data Sense** Deploy 

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

Deploying Cloud Data Sense

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



Deploying Cloud Data Sense instance

Verify connectivity to BlueXP Connector and to the internet

Initializing Cloud Data Sense

[Cancel deployment](#)

5. When the instance is deployed and BlueXP classification is installed, select **Continue to configuration** to go to the *Configuration* page.

Result

BlueXP deploys the BlueXP classification instance in your cloud provider.

Upgrades to the BlueXP Connector and BlueXP classification software is automated as long as the instances have internet connectivity.

What's Next

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

Install BlueXP classification on a host that has internet access

Complete a few steps to install BlueXP classification on a Linux host in your network, or on a Linux host in the cloud, that has internet access. You'll need to deploy the Linux host manually in your network or in the cloud as part of this installation.

The on-premises installation may be a good option if you prefer to scan on-premises ONTAP systems using a BlueXP classification instance that's also located on premises — but this is not a requirement. The software functions exactly the same way regardless of which installation method you choose.

The BlueXP classification installation script starts by checking if the system and environment meet the required prerequisites. If the prerequisites are all met, then the installation starts. If you would like to verify the prerequisites independently of running the BlueXP classification installation, there is a separate software package you can download that only tests for the prerequisites. [See how to check if your Linux host is ready to install BlueXP classification.](#)

The typical installation on a Linux host *in your premises* has the following components and connections.

The typical installation on a Linux host *in the cloud* has the following components and connections.



For legacy versions 1.30 and earlier, if you need to install BlueXP classification on multiple hosts, refer to [Install BlueXP classification on multiple hosts with no internet access.](#)

You can also [install BlueXP classification in an on-premises site that doesn't have internet access.](#)

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, [deploy the Connector on-premises](#) on a Linux host in your network, or on a Linux host in the cloud.

You can also create a Connector with your cloud provider. See [creating a Connector in AWS](#), [creating a Connector in Azure](#), or [creating a Connector in GCP](#).

2

Review prerequisites

Ensure that your environment can meet the prerequisites. This includes outbound internet access for the instance, connectivity between the Connector and BlueXP classification over port 443, and more. [See the complete list.](#)

You also need a Linux system that meets the [following requirements](#).

3

Download and deploy BlueXP classification

Download the Cloud BlueXP classification software from the NetApp Support Site and copy the installer file to the Linux host you plan to use. Then launch the installation wizard and follow the prompts to deploy the BlueXP classification instance.

Create a Connector

A BlueXP Connector is required before you can install and use BlueXP classification. In most cases you'll probably have a Connector set up before you attempt to activate BlueXP classification because most [BlueXP features require a Connector](#), but there are cases where you'll need to set one up now.

To create one in your cloud provider environment, see [creating a Connector in AWS](#), [creating a Connector in Azure](#), or [creating a Connector in GCP](#).

There are some scenarios where you have to use a Connector that's deployed in a specific cloud provider:

- When scanning data in Cloud Volumes ONTAP in AWS or Amazon FSx for ONTAP, 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.

For Azure NetApp Files, it must be deployed in the same region as the volumes you wish to scan.

- When scanning data in Cloud Volumes ONTAP in GCP, you use a Connector in GCP.

On-prem ONTAP systems, NetApp file shares and database accounts can be scanned using any of these cloud Connectors.

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

You'll need the IP address or host name of the Connector system when installing BlueXP classification. You'll have this information if you installed the Connector in your premises. If the Connector is deployed in the cloud, you can find this information from the BlueXP console: click the Help icon, select **Support**, and click **BlueXP Connector**.

Prepare the Linux host system

BlueXP classification software must run on a host that meets specific operating system requirements, RAM requirements, software requirements, and so on. The Linux host can be in your network, or in the cloud.

Ensure that you can keep BlueXP classification running. The BlueXP classification machine needs to stay on to continuously scan your data.

- BlueXP classification is not supported on a host that is shared with other applications - the host must be a dedicated host.
- When building the host system in your premises, you can choose among these system sizes depending on the size of the dataset that you plan to have BlueXP classification scan.

System size	CPU	RAM (swap memory must be disabled)	Disk
Extra Large	32 CPUs	128 GB RAM	<ul style="list-style-type: none"> • 1 TiB SSD on /, or 100 GiB available on /opt • 895 GiB available on /var/lib/docker • 5 GiB on /tmp • For Podman, 5 GB on /tmp • For Podman, 30 GB on /var/tmp
Large	16 CPUs	64 GB RAM	<ul style="list-style-type: none"> • 500 GiB SSD on /, or 100 GiB available on /opt • 395 GiB available on /var/lib/docker or for Podman /var/lib/containers or for Podman /var/lib/containers • 5 GiB on /tmp • For Podman, 5 GB on /tmp • For Podman, 30 GB on /var/tmp

- When deploying a compute instance in the cloud for your BlueXP classification installation, we recommend a system that meets the "Large" system requirements above:
 - **Amazon Elastic Compute Cloud (Amazon EC2) instance type:** We recommend "m6i.4xlarge". [See additional AWS instance types.](#)
 - **Azure VM size:** We recommend "Standard_D16s_v3". [See additional Azure instance types.](#)
 - **GCP machine type:** We recommend "n2-standard-16". [See additional GCP instance types.](#)
- **UNIX folder permissions:** The following minimum UNIX permissions are required:

Folder	Minimum Permissions
/tmp	rwXrwxrwt
/opt	rwXr-xr-x
/var/lib/docker	rwX-----
/usr/lib/systemd/system	rwXr-xr-x

- **Operating system:**
 - The following operating systems require using the Docker container engine:
 - Red Hat Enterprise Linux version 7.8 and 7.9
 - Ubuntu 22.04 (requires BlueXP classification version 1.23 or greater)
 - Ubuntu 24.04 (requires BlueXP classification version 1.23 or greater)

- The following operating systems require using the Podman container engine, and they require BlueXP classification version 1.30 or greater:
 - Red Hat Enterprise Linux version 8.8, 8.10, 9.0, 9.1, 9.2, 9.3, 9.4, 9.5, and 9.6.
- Advanced Vector Extensions (AVX2) must be enabled on the host system.
- **Red Hat Subscription Management:** The host must be registered with Red Hat Subscription Management. If it's not registered, the system can't access repositories to update required 3rd-party software during installation.
- **Additional software:** You must install the following software on the host before you install BlueXP classification:
 - Depending on the OS you are using, you'll need to install one of the container engines:
 - Docker Engine version 19.3.1 or greater. [View installation instructions.](#)
 - Podman version 4 or greater. To install Podman, enter `(sudo yum install podman netavark -y)`.
- Python version 3.6 or greater. [View installation instructions.](#)
 - **NTP considerations:** NetApp recommends configuring the BlueXP classification system to use a Network Time Protocol (NTP) service. The time must be synchronized between the BlueXP classification system and the BlueXP Connector system.
- **Firewalld considerations:** If you are planning to use `firewalld`, we recommend that you enable it before installing BlueXP classification. Run the following commands to configure `firewalld` so that it is compatible with BlueXP classification:

```
firewall-cmd --permanent --add-service=http
firewall-cmd --permanent --add-service=https
firewall-cmd --permanent --add-port=80/tcp
firewall-cmd --permanent --add-port=8080/tcp
firewall-cmd --permanent --add-port=443/tcp
firewall-cmd --reload
```

If you're planning to use additional BlueXP classification hosts as scanner nodes, add these rules to your primary system at this time:

```
firewall-cmd --permanent --add-port=2377/tcp
firewall-cmd --permanent --add-port=7946/udp
firewall-cmd --permanent --add-port=7946/tcp
firewall-cmd --permanent --add-port=4789/udp
```

Note that you must restart Docker or Podman whenever you enable or update `firewalld` settings.



The IP address of the BlueXP classification host system can't be changed after installation.

Enable outbound internet access from BlueXP classification

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

Endpoints	Purpose
https://api.bluexp.netapp.com	Communication with the BlueXP service, which includes NetApp accounts.
https://netapp-cloud-account.auth0.com https://auth0.com	Communication with the BlueXP website for centralized user authentication.
https://support.compliance.api.bluexp.netapp.com/ https://hub.docker.com https://auth.docker.io https://registry-1.docker.io https://index.docker.io/ https://dseasb33srmrn.cloudfront.net/ https://production.cloudflare.docker.com/	Provides access to software images, manifests, templates, and to send logs and metrics.
https://support.compliance.api.bluexp.netapp.com/	Enables NetApp to stream data from audit records.
https://github.com/docker https://download.docker.com	Provides prerequisite packages for docker installation.
http://packages.ubuntu.com/ http://archive.ubuntu.com	Provides prerequisite packages for Ubuntu installation.

Verify that all required ports are enabled

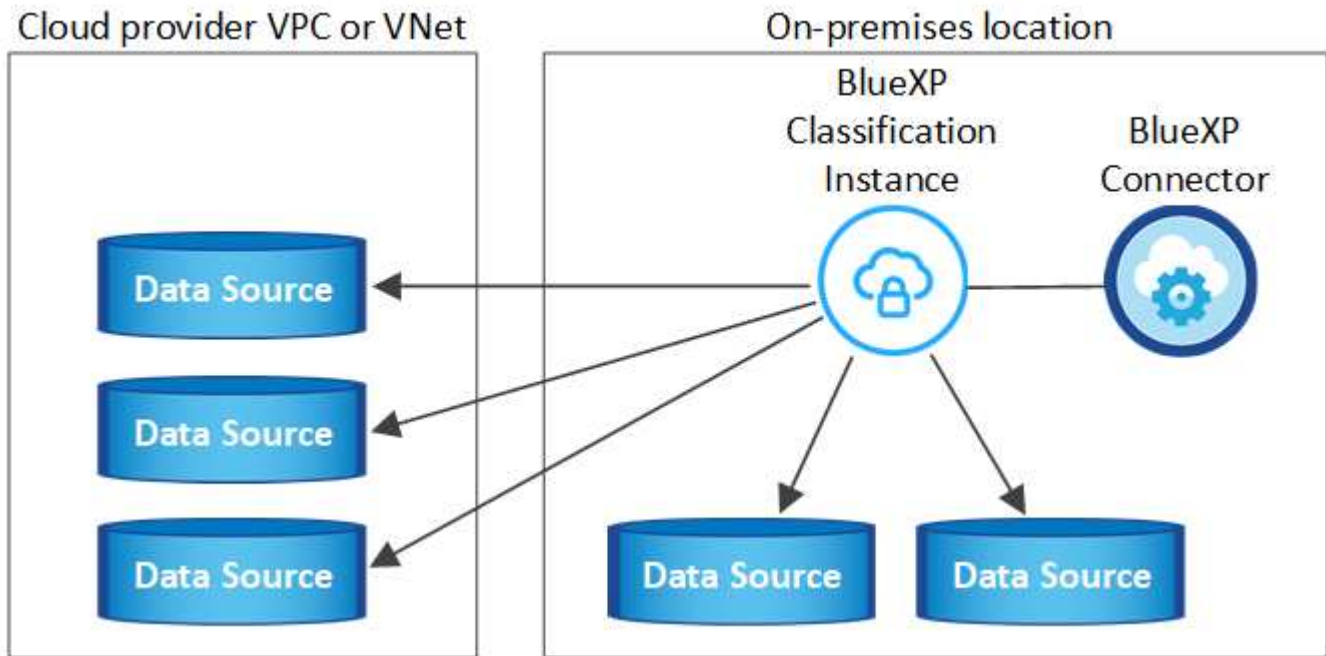
You must ensure that all required ports are open for communication between the Connector, BlueXP classification, Active Directory, and your data sources.

Connection Type	Ports	Description
Connector <> BlueXP classification	8080 (TCP), 443 (TCP), and 80. 9000	<p>The firewall or routing rules for the Connector must allow inbound and outbound traffic over port 443 to and from the BlueXP classification instance.</p> <p>Make sure port 8080 is open so you can see the installation progress in BlueXP.</p> <p>If a firewall is used on the Linux host, port 9000 is required for internal processes within an Ubuntu server.</p>

Connection Type	Ports	Description
Connector <> ONTAP cluster (NAS)	443 (TCP)	<p>BlueXP discovers ONTAP clusters using HTTPS. If you use custom firewall policies, they must meet the following requirements:</p> <ul style="list-style-type: none"> • The Connector host must allow outbound HTTPS access through port 443. If the Connector is in the cloud, all outbound communication is allowed by the predefined firewall or routing rules. • The ONTAP cluster must allow inbound HTTPS access through port 443. The default "mgmt" firewall policy allows inbound HTTPS access from all IP addresses. If you modified this default policy, or if you created your own firewall policy, you must associate the HTTPS protocol with that policy and enable access from the Connector host.
BlueXP classification <> ONTAP cluster	<ul style="list-style-type: none"> • For NFS - 111 (TCP\UDP) and 2049 (TCP\UDP) • For CIFS - 139 (TCP\UDP) and 445 (TCP\UDP) 	<p>BlueXP classification needs a network connection to each Cloud Volumes ONTAP subnet or on-prem ONTAP system. Firewalls or routing rules for Cloud Volumes ONTAP must allow inbound connections from the BlueXP classification instance.</p> <p>Make sure these ports are open to the BlueXP classification instance:</p> <ul style="list-style-type: none"> • For NFS - 111 and 2049 • For CIFS - 139 and 445 <p>NFS volume export policies must allow access from the BlueXP classification instance.</p>
BlueXP classification <> Active Directory	389 (TCP & UDP), 636 (TCP), 3268 (TCP), and 3269 (TCP)	<p>You must have an Active Directory already set up for the users in your company. Additionally, BlueXP classification needs Active Directory credentials to scan CIFS volumes.</p> <p>You must have the information for the Active Directory:</p> <ul style="list-style-type: none"> • DNS Server IP Address, or multiple IP Addresses • User Name and Password for the server • Domain Name (Active Directory Name) • Whether you are using secure LDAP (LDAPS) or not • LDAP Server Port (typically 389 for LDAP, and 636 for secure LDAP)

Install BlueXP classification on the Linux host

For typical configurations you'll install the software on a single host system. [See those steps here.](#)



See [Preparing the Linux host system](#) and [Reviewing prerequisites](#) for the full list of requirements before you deploy BlueXP classification.

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



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

Single-host installation for typical configurations

Review the requirements and follow these steps when installing BlueXP classification software on a single on-premises host.

[Watch this video](#) to see how to install BlueXP classification.

Note that all installation activities are logged when installing BlueXP classification. If you run into any issues during installation, you can view the contents of the installation audit log. It is written to `/opt/netapp/install_logs/`. [See more details here.](#)

Before you begin

- Verify that your Linux system meets the [host requirements](#).
- Verify that the system has the two prerequisite software packages installed (Docker Engine or Podman, and Python 3).
- Make sure you have root privileges on the Linux system.
- If you're using a proxy for access to the internet:

- You'll need the proxy server information (IP address or host name, connection port, connection scheme: https or http, user name and password).
- If the proxy is performing TLS interception, you'll need to know the path on the BlueXP classification Linux system where the TLS CA certificates are stored.
- The proxy must be non-transparent. BlueXP classification does not currently support transparent proxies.
- The user must be a local user. Domain users are not supported.

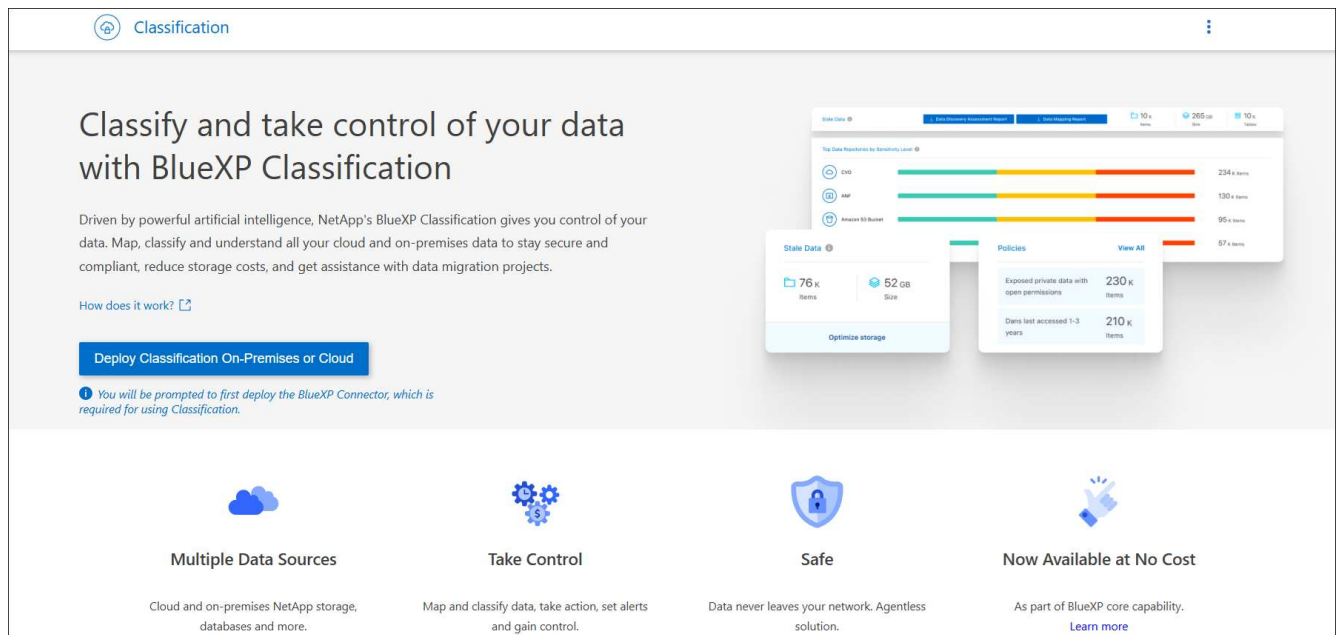
- Verify that your offline environment meets the required [permissions and connectivity](#).

Steps

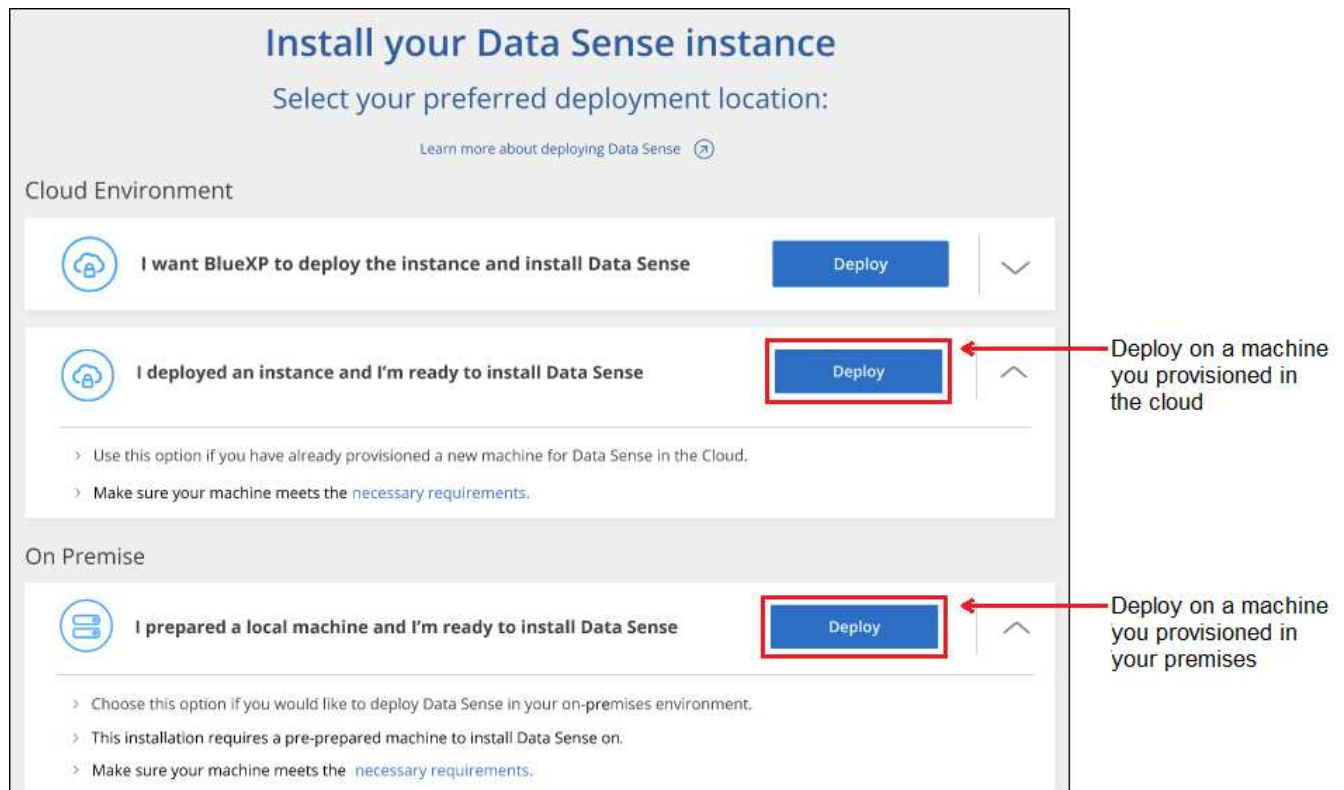
1. Download the BlueXP classification software from the [NetApp Support Site](#). The file you should select is named **DATASENSE-INSTALLER-<version>.tar.gz**.
2. Copy the installer file to the Linux host you plan to use (using `scp` or some other method).
3. Unzip the installer file on the host machine, for example:

```
tar -xzf DATASENSE-INSTALLER-V1.25.0.tar.gz
```

4. In BlueXP, select **Governance > Classification**.
5. Select **Deploy Classification On-Premises or Cloud**.



6. Depending on whether you are installing BlueXP classification on an instance you prepared in the cloud or on an instance you prepared in your premises, click the appropriate **Deploy** button to start the BlueXP classification installation.



7. The *Deploy Data Sense On Premises* dialog is displayed. Copy the provided command (for example: `sudo ./install.sh -a 12345 -c 27AG75 -t 2198qq`) and paste it in a text file so you can use it later. Then click **Close** to dismiss the dialog.
8. On the host machine, enter the command you copied and then follow a series of prompts, or you can provide the full command including all required parameters as command line arguments.

Note that the installer performs a pre-check to make sure your system and networking requirements are in place for a successful installation. [Watch this video](#) to understand the pre-check messages and implications.

Enter parameters as prompted:	Enter the full command:
<ol style="list-style-type: none"> 1. Paste the command you copied from step 7: <pre>sudo ./install.sh -a <account_id> -c <client_id> -t <user_token></pre> <p>If you are installing on a cloud instance (not on your premises), add <code>--manual-cloud-install <cloud_provider></code>.</p> 2. Enter the IP address or host name of the BlueXP classification host machine so it can be accessed by the Connector system. 3. Enter the IP address or host name of the BlueXP Connector host machine so it can be accessed by the BlueXP classification system. 4. Enter proxy details as prompted. If your BlueXP Connector already uses a proxy, there is no need to enter this information again here since BlueXP classification will automatically use the proxy used by the Connector. 	<p>Alternatively, you can create the whole command in advance, providing the necessary host and proxy parameters:</p> <pre>sudo ./install.sh -a <account_id> -c <client_id> -t <user_token> --host <ds_host> --manager-host <cm_host> --manual-cloud-install <cloud_provider> --proxy-host <proxy_host> --proxy-port <proxy_port> --proxy-scheme <proxy_scheme> --proxy -user <proxy_user> --proxy-password <proxy_password> --cacert-folder-path <ca_cert_dir></pre>

Variable values:

- *account_id* = NetApp Account ID
- *client_id* = Connector Client ID (add the suffix "clients" to the client ID if it not already there)
- *user_token* = JWT user access token
- *ds_host* = IP address or host name of the BlueXP classification Linux system.
- *cm_host* = IP address or host name of the BlueXP Connector system.
- *cloud_provider* = When installing on a cloud instance, enter "AWS", "Azure", or "Gcp" depending on cloud provider.
- *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. The user must be a local user - domain users are not supported.
- *proxy_password* = Password for the user name that you specified.
- *ca_cert_dir* = Path on the BlueXP classification Linux system containing additional TLS CA certificate bundles. Only required if the proxy is performing TLS interception.

Result

The BlueXP classification installer installs packages, registers the installation, and installs BlueXP classification. Installation can take 10 to 20 minutes.

If there is connectivity over port 8080 between the host machine and the Connector instance, you'll see the installation progress in the BlueXP classification tab in BlueXP.

What's Next

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

Install BlueXP classification on a Linux host with no internet access

Complete a few steps to install BlueXP classification on a Linux host in an on-premises site that doesn't have internet access - also known as *private mode*. This type of installation, which uses an installation script, has no connectivity to the BlueXP SaaS layer.

[Learn about the different deployment modes for the BlueXP Connector and BlueXP classification.](#)

You can also [deploy BlueXP classification in an on-premises site that has internet access.](#)

The BlueXP classification installation script starts by checking if the system and environment meet the required prerequisites. If the prerequisites are all met, then the installation starts. If you would like to verify the prerequisites independently of running the BlueXP classification installation, there is a separate software package you can download that only tests for the prerequisites. [See how to check if your Linux host is ready to install BlueXP classification.](#)



For legacy versions 1.30 and earlier, if you need to install BlueXP classification on multiple hosts, refer to [Install BlueXP classification on multiple hosts with no internet access.](#)

Supported data sources

When installed private mode (sometimes called an "offline" or "dark" site), BlueXP classification can only scan data from data sources that are also local to the on-premises site. At this time, BlueXP classification can scan the following **local** data sources:

- On-premises ONTAP systems
- Database schemas

There is no support currently for scanning Cloud Volumes ONTAP, Azure NetApp Files, or FSx for ONTAP accounts when BlueXP classification is deployed in private mode.

Limitations

Most BlueXP classification features work when it is deployed in a site with no internet access. However, certain features that require internet access are not supported, for example:

- Setting BlueXP roles for different users (for example, Account Admin or Compliance Viewer)
- Copying and synchronizing source files using BlueXP copy and sync
- Automated software upgrades from BlueXP

Both the BlueXP Connector and BlueXP classification will require periodic manual upgrades to enable new features. You can see the BlueXP classification version at the bottom of the BlueXP classification UI pages. Check the [BlueXP classification Release Notes](#) to see the new features in each release and whether you want those features. Then you can follow the steps to [upgrade the BlueXP Connector](#) and [upgrade your BlueXP classification software.](#)

Quick start

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

1

Install the BlueXP Connector

If you don't already have a Connector installed in private mode, [deploy the Connector](#) on a Linux host now.

2

Review BlueXP classification prerequisites

Ensure that your Linux system meets the [host requirements](#), that it has all required software installed, and that your offline environment meets the required [permissions and connectivity](#).

3

Download and deploy BlueXP classification

Download the BlueXP classification software from the NetApp Support Site and copy the installer file to the Linux host you plan to use. Then launch the installation wizard and follow the prompts to deploy the BlueXP classification instance.

Install the BlueXP Connector

If you don't already have a BlueXP Connector installed in private mode, [deploy the Connector](#) on a Linux host in your offline site.

Prepare the Linux host system

BlueXP classification software must run on a host that meets specific operating system requirements, RAM requirements, software requirements, and so on.

- BlueXP classification is not supported on a host that is shared with other applications - the host must be a dedicated host.
- When building the host system in your premises, you can choose among these system sizes depending on the size of the dataset that you plan to have BlueXP classification scan.

System size	CPU	RAM (swap memory must be disabled)	Disk
Extra Large	32 CPUs	128 GB RAM	<ul style="list-style-type: none">• 1 TiB SSD on /, or 100 GiB available on /opt• 895 GiB available on /var/lib/docker• 5 GiB on /tmp• For Podman, 5 GB on /tmp• For Podman, 30 GB on /var/tmp

System size	CPU	RAM (swap memory must be disabled)	Disk
Large	16 CPUs	64 GB RAM	<ul style="list-style-type: none"> • 500 GiB SSD on /, or 100 GiB available on /opt • 395 GiB available on /var/lib/docker or for Podman /var/lib/containers or for Podman /var/lib/containers • 5 GiB on /tmp • For Podman, 5 GB on /tmp • For Podman, 30 GB on /var/tmp

- When deploying a compute instance in the cloud for your BlueXP classification installation, we recommend a system that meets the "Large" system requirements above:
 - **Amazon Elastic Compute Cloud (Amazon EC2) instance type:** We recommend "m6i.4xlarge". [See additional AWS instance types.](#)
 - **Azure VM size:** We recommend "Standard_D16s_v3". [See additional Azure instance types.](#)
 - **GCP machine type:** We recommend "n2-standard-16". [See additional GCP instance types.](#)
- **UNIX folder permissions:** The following minimum UNIX permissions are required:

Folder	Minimum Permissions
/tmp	rwXrwxrwt
/opt	rwXr-xr-x
/var/lib/docker	rwX-----
/usr/lib/systemd/system	rwXr-xr-x

- **Operating system:**
 - The following operating systems require using the Docker container engine:
 - Red Hat Enterprise Linux version 7.8 and 7.9
 - Ubuntu 22.04 (requires BlueXP classification version 1.23 or greater)
 - Ubuntu 24.04 (requires BlueXP classification version 1.23 or greater)
 - The following operating systems require using the Podman container engine, and they require BlueXP classification version 1.30 or greater:
 - Red Hat Enterprise Linux version 8.8, 8.10, 9.0, 9.1, 9.2, 9.3, 9.4, 9.5, and 9.6.
 - Advanced Vector Extensions (AVX2) must be enabled on the host system.
- **Red Hat Subscription Management:** The host must be registered with Red Hat Subscription Management. If it's not registered, the system can't access repositories to update required 3rd-party software during installation.
- **Additional software:** You must install the following software on the host before you install BlueXP classification:

- Depending on the OS you are using, you'll need to install one of the container engines:
 - Docker Engine version 19.3.1 or greater. [View installation instructions.](#)
 - Podman version 4 or greater. To install Podman, enter `(sudo yum install podman netavark -y)`.
- Python version 3.6 or greater. [View installation instructions.](#)
 - **NTP considerations:** NetApp recommends configuring the BlueXP classification system to use a Network Time Protocol (NTP) service. The time must be synchronized between the BlueXP classification system and the BlueXP Connector system.
- **Firewalld considerations:** If you are planning to use `firewalld`, we recommend that you enable it before installing BlueXP classification. Run the following commands to configure `firewalld` so that it is compatible with BlueXP classification:

```
firewall-cmd --permanent --add-service=http
firewall-cmd --permanent --add-service=https
firewall-cmd --permanent --add-port=80/tcp
firewall-cmd --permanent --add-port=8080/tcp
firewall-cmd --permanent --add-port=443/tcp
firewall-cmd --reload
```

Note that you must restart Docker or Podman whenever you enable or update `firewalld` settings.



The IP address of the BlueXP classification host system can't be changed after installation.

Verify BlueXP and BlueXP classification prerequisites

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

- Ensure that the Connector has permissions to deploy resources and create security groups for the BlueXP classification instance. You can find the latest BlueXP permissions in [the policies provided by NetApp](#).
- Ensure that you can keep BlueXP classification running. The BlueXP classification instance needs to stay on to continuously scan your data.
- Ensure web browser connectivity to BlueXP classification. After BlueXP classification is enabled, ensure that users access the BlueXP interface from a host that has a connection to the BlueXP classification instance.

The BlueXP classification instance uses a private IP address to ensure that the indexed data isn't accessible to others. As a result, the web browser that you use to access BlueXP must have a connection to that private IP address. That connection can come from a host that's inside the same network as the BlueXP classification instance.

Verify that all required ports are enabled

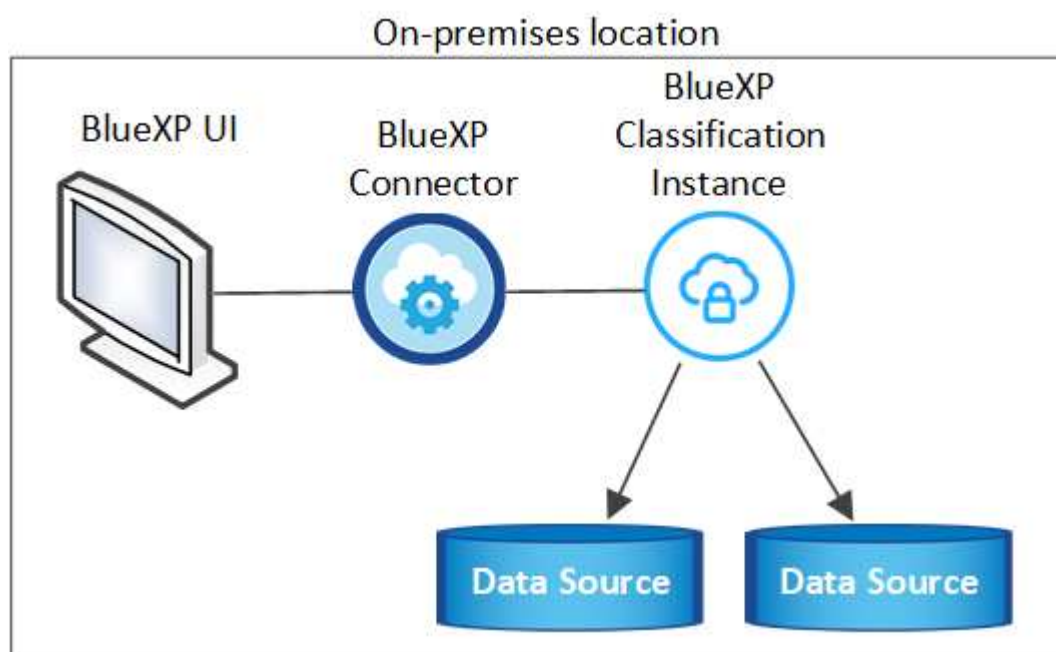
You must ensure that all required ports are open for communication between the Connector, BlueXP classification, Active Directory, and your data sources.

Connection Type	Ports	Description
Connector <> BlueXP classification	8080 (TCP), 6000 (TCP), 443 (TCP), and 80. 9000	<p>The security group for the Connector must allow inbound and outbound traffic over ports 6000 and 443 to and from the BlueXP classification instance.</p> <ul style="list-style-type: none"> • Port 6000 is required so that the BlueXP classification BYOL license works in a dark site. • Port 8080 should be open so you can see the installation progress in BlueXP. • If a firewall is used on the Linux host, port 9000 is required for internal processes within an Ubuntu server.
Connector <> ONTAP cluster (NAS)	443 (TCP)	<p>BlueXP discovers ONTAP clusters using HTTPS. If you use custom firewall policies, they must meet the following requirements:</p> <ul style="list-style-type: none"> • The Connector host must allow outbound HTTPS access through port 443. If the Connector is in the cloud, all outbound communication is allowed by the predefined security group. • The ONTAP cluster must allow inbound HTTPS access through port 443. The default "mgmt" firewall policy allows inbound HTTPS access from all IP addresses. If you modified this default policy, or if you created your own firewall policy, you must associate the HTTPS protocol with that policy and enable access from the Connector host.
BlueXP classification <> ONTAP cluster	<ul style="list-style-type: none"> • For NFS - 111 (TCP\UDP) and 2049 (TCP\UDP) • For CIFS - 139 (TCP\UDP) and 445 (TCP\UDP) 	<p>BlueXP classification needs a network connection to each Cloud Volumes ONTAP subnet or on-prem ONTAP system. Security groups for Cloud Volumes ONTAP must allow inbound connections from the BlueXP classification instance.</p> <p>Make sure these ports are open to the BlueXP classification instance:</p> <ul style="list-style-type: none"> • For NFS - 111 and 2049 • For CIFS - 139 and 445 <p>NFS volume export policies must allow access from the BlueXP classification instance.</p>

Connection Type	Ports	Description
BlueXP classification <> Active Directory	389 (TCP & UDP), 636 (TCP), 3268 (TCP), and 3269 (TCP)	<p>You must have an Active Directory already set up for the users in your company. Additionally, BlueXP classification needs Active Directory credentials to scan CIFS volumes.</p> <p>You must have the information for the Active Directory:</p> <ul style="list-style-type: none"> • DNS Server IP Address, or multiple IP Addresses • User Name and Password for the server • Domain Name (Active Directory Name) • Whether you are using secure LDAP (LDAPS) or not • LDAP Server Port (typically 389 for LDAP, and 636 for secure LDAP)
If a firewall used on Linux host	9000	Needed for internal processes within an Ubuntu server.

Install BlueXP classification on the on-premises Linux host

For typical configurations you'll install the software on a single host system.



Single-host installation for typical configurations

Follow these steps when installing BlueXP classification software on a single on-premises host in an offline environment.

Note that all installation activities are logged when installing BlueXP classification. If you run into any issues during installation, you can view the contents of the installation audit log. It is written to

/opt/netapp/install_logs/. [See more details here.](#)

Before you begin

- Verify that your Linux system meets the [host requirements](#).
- Verify that you have installed the two prerequisite software packages (Docker Engine or Podman, and Python 3).
- Make sure you have root privileges on the Linux system.
- Verify that your offline environment meets the required [permissions and connectivity](#).

Steps

1. On an internet-configured system, download the BlueXP classification software from the [NetApp Support Site](#). The file you should select is named **DataSense-offline-bundle-<version>.tar.gz**.
2. Copy the installer bundle to the Linux host you plan to use in private mode.
3. Unzip the installer bundle on the host machine, for example:

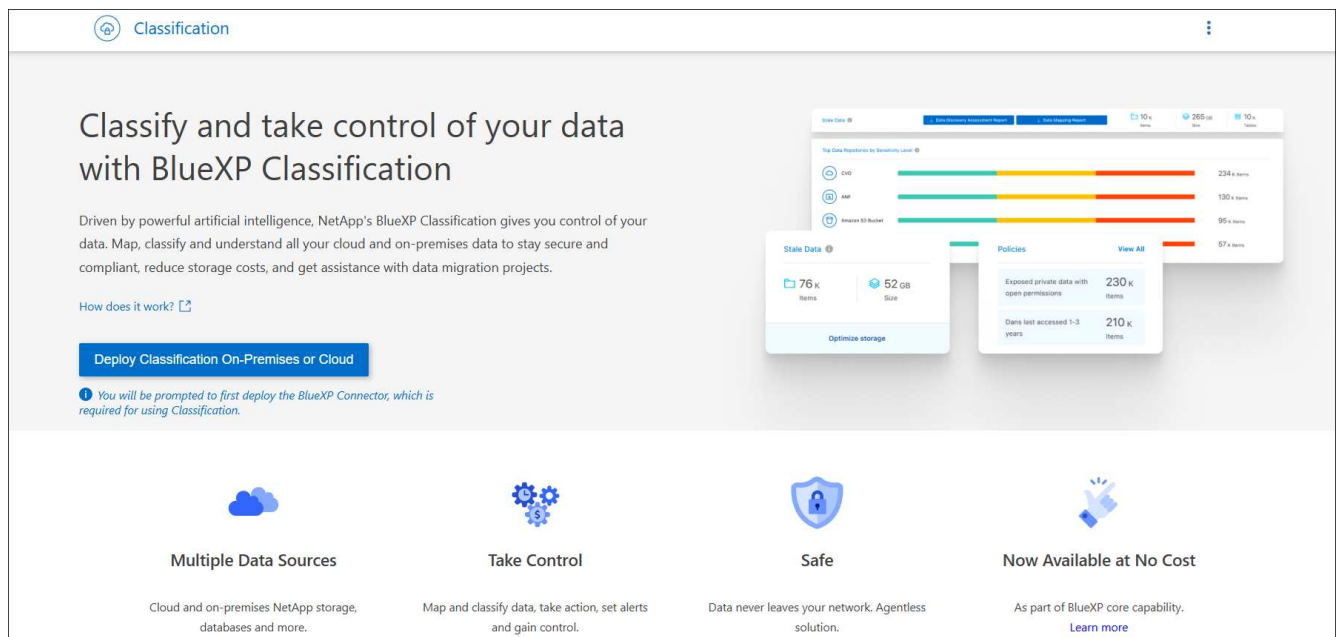
```
tar -xzf DataSense-offline-bundle-v1.25.0.tar.gz
```

This extracts required software and the actual installation file **cc_onprem_installer.tar.gz**.

4. Unzip the installation file on the host machine, for example:

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

5. Launch BlueXP and select **Governance > Classification**.
6. Select **Deploy Classification On-Premises or Cloud**.




7. Click **Deploy** to start the on-prem installation.


Install your Data Sense instance

Select your preferred deployment location:


[Learn more about deploying Data Sense](#)

Cloud Environment


I want BlueXP to deploy the instance and install Data Sense
Deploy
▼


I deployed an instance and I'm ready to install Data Sense
Deploy
▼

On Premise


I prepared a local machine and I'm ready to install Data Sense
Deploy
▲

➤ Choose this option if you would like to deploy Data Sense in your on-premises environment.

➤ This installation requires a pre-prepared machine to install Data Sense on.

➤ Make sure your machine meets the [necessary requirements](#).

8. The *Deploy Data Sense On Premises* dialog is displayed. Copy the provided command (for example: `sudo ./install.sh -a 12345 -c 27AG75 -t 2198qq --darksite`) and paste it in a text file so you can use it later. Then click **Close** to dismiss the dialog.
9. On the host machine, enter the command you copied and then follow a series of prompts, or you can provide the full command including all required parameters as command line arguments.

Note that the installer performs a pre-check to make sure your system and networking requirements are in place for a successful installation.

Enter parameters as prompted:	Enter the full command:
<ol style="list-style-type: none"> 1. Paste the information you copied from step 8: <code>sudo ./install.sh -a <account_id> -c <client_id> -t <user_token> --darksite</code> 2. Enter the IP address or host name of the BlueXP classification host machine so it can be accessed by the Connector system. 3. Enter the IP address or host name of the BlueXP Connector host machine so it can be accessed by the BlueXP classification system. 	<p>Alternatively, you can create the whole command in advance, providing the necessary host parameters:</p> <pre>sudo ./install.sh -a <account_id> -c <client_id> -t <user_token> --host <ds_host> --manager-host <cm_host> --no-proxy --darksite</pre>

Variable values:

- *account_id* = NetApp Account ID

- *client_id* = Connector Client ID (add the suffix "clients" to the client ID if it not already there)
- *user_token* = JWT user access token
- *ds_host* = IP address or host name of the BlueXP classification system.
- *cm_host* = IP address or host name of the BlueXP Connector system.

Result

The BlueXP classification installer installs packages, registers the installation, and installs BlueXP classification. Installation can take 10 to 20 minutes.

If there is connectivity over port 8080 between the host machine and the Connector instance, you'll see the installation progress in the BlueXP classification tab in BlueXP.

What's Next

From the Configuration page you can select the local [on-prem ONTAP clusters](#) and [databases](#) that you want to scan.

Upgrade BlueXP classification software

Since BlueXP classification software is updated with new features on a regular basis, you should get into a routine to check for new versions periodically to make sure you're using the newest software and features. You'll need to upgrade BlueXP classification software manually because there's no internet connectivity to perform the upgrade automatically.

Before you begin

- We recommend that your BlueXP Connector software is upgraded to the newest available version. [See the Connector upgrade steps](#).
- Starting with BlueXP classification version 1.24 you can perform upgrades to any future version of software.

If your BlueXP classification software is running a version prior to 1.24, you can upgrade only one major version at a time. For example, if you have version 1.21.x installed, you can upgrade only to 1.22.x. If you are a few major versions behind, you'll need to upgrade the software multiple times.

Steps

1. On an internet-configured system, download the BlueXP classification software from the [NetApp Support Site](#). The file you should select is named **DataSense-offline-bundle-<version>.tar.gz**.
2. Copy the software bundle to the Linux host where BlueXP classification is installed in the dark site.
3. Unzip the software bundle on the host machine, for example:

```
tar -xvf DataSense-offline-bundle-v1.25.0.tar.gz
```

This extracts the installation file **cc_onprem_installer.tar.gz**.

4. Unzip the installation file on the host machine, for example:

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

This extracts the upgrade script **start_darksite_upgrade.sh** and any required third-party software.

5. Run the upgrade script on the host machine, for example:

```
start_darksite_upgrade.sh
```

Result

The BlueXP classification software is upgraded on your host. The update can take 5 to 10 minutes.

You can verify that the software has been updated by checking the version at the bottom of the BlueXP classification UI pages.

Check that your Linux host is ready to install BlueXP classification

Before installing BlueXP classification manually on a Linux host, optionally run a script on the host to verify that all the prerequisites are in place for installing BlueXP classification. You can run this script on a Linux host in your network, or on a Linux host in the cloud. The host can be connected to the internet, or the host can reside in a site that doesn't have internet access (a *dark site*).

There is also a prerequisite test script that is part of the BlueXP classification installation script. The script described here is specifically designed for users who want to verify the Linux host independently of running the BlueXP classification installation script.

Getting Started

You'll perform the following tasks.

1. Optionally, install a BlueXP Connector if you don't already have one installed. You can run the test script without having a Connector installed, but the script checks for connectivity between the Connector and the BlueXP classification host machine - so it is recommended that you have a Connector.
2. Prepare the host machine and verify that it meets all the requirements.
3. Enable outbound internet access from the BlueXP classification host machine.
4. Verify that all required ports are enabled on all systems.
5. Download and run the Prerequisite test script.

Create a Connector

A BlueXP Connector is required before you can install and use BlueXP classification. You can, however, run the Prerequisites script without a Connector.

You can [install the Connector on-premises](#) on a Linux host in your network or on a Linux host in the cloud. Some users planning to install BlueXP classification on-prem may also choose to install the Connector on-prem.

To create a Connector in your cloud provider environment, see [creating a Connector in AWS](#), [creating a Connector in Azure](#), or [creating a Connector in GCP](#).

You'll need the IP address or host name of the Connector system when running the Prerequisites script. You'll

have this information if you installed the Connector in your premises. If the Connector is deployed in the cloud, you can find this information from the BlueXP console: click the Help icon, select **Support**, and click **BlueXP Connector**.

Verify host requirements

BlueXP classification software must run on a host that meets specific operating system requirements, RAM requirements, software requirements, and so on.

- BlueXP classification is not supported on a host that is shared with other applications - the host must be a dedicated host.
- When building the host system in your premises, you can choose among these system sizes depending on the size of the dataset that you plan to have BlueXP classification scan.

System size	CPU	RAM (swap memory must be disabled)	Disk
Extra Large	32 CPUs	128 GB RAM	<ul style="list-style-type: none">• 1 TiB SSD on /, or 100 GiB available on /opt• 895 GiB available on /var/lib/docker• 5 GiB on /tmp• For Podman, 5 GB on /tmp• For Podman, 30 GB on /var/tmp
Large	16 CPUs	64 GB RAM	<ul style="list-style-type: none">• 500 GiB SSD on /, or 100 GiB available on /opt• 395 GiB available on /var/lib/docker or for Podman /var/lib/containers or for Podman /var/lib/containers• 5 GiB on /tmp• For Podman, 5 GB on /tmp• For Podman, 30 GB on /var/tmp

- When deploying a compute instance in the cloud for your BlueXP classification installation, we recommend a system that meets the "Large" system requirements above:
 - **Amazon Elastic Compute Cloud (Amazon EC2) instance type:** We recommend "m6i.4xlarge". [See additional AWS instance types.](#)
 - **Azure VM size:** We recommend "Standard_D16s_v3". [See additional Azure instance types.](#)
 - **GCP machine type:** We recommend "n2-standard-16". [See additional GCP instance types.](#)
- **UNIX folder permissions:** The following minimum UNIX permissions are required:

Folder	Minimum Permissions
/tmp	rwXrwxrwt
/opt	rwXr-Xr-X
/var/lib/docker	rwX-----
/usr/lib/systemd/system	rwXr-Xr-X

- **Operating system:**

- The following operating systems require using the Docker container engine:
 - Red Hat Enterprise Linux version 7.8 and 7.9
 - Ubuntu 22.04 (requires BlueXP classification version 1.23 or greater)
 - Ubuntu 24.04 (requires BlueXP classification version 1.23 or greater)
- The following operating systems require using the Podman container engine, and they require BlueXP classification version 1.30 or greater:
 - Red Hat Enterprise Linux version 8.8, 8.10, 9.0, 9.1, 9.2, 9.3, 9.4, 9.5, and 9.6.
- Advanced Vector Extensions (AVX2) must be enabled on the host system.

- **Red Hat Subscription Management:** The host must be registered with Red Hat Subscription Management. If it's not registered, the system can't access repositories to update required 3rd-party software during installation.

- **Additional software:** You must install the following software on the host before you install BlueXP classification:

- Depending on the OS you are using, you'll need to install one of the container engines:
 - Docker Engine version 19.3.1 or greater. [View installation instructions.](#)
 - Podman version 4 or greater. To install Podman, enter `(sudo yum install podman netavark -y)`.

- Python version 3.6 or greater. [View installation instructions.](#)

- **NTP considerations:** NetApp recommends configuring the BlueXP classification system to use a Network Time Protocol (NTP) service. The time must be synchronized between the BlueXP classification system and the BlueXP Connector system.

- **Firewalld considerations:** If you are planning to use `firewalld`, we recommend that you enable it before installing BlueXP classification. Run the following commands to configure `firewalld` so that it is compatible with BlueXP classification:

```
firewall-cmd --permanent --add-service=http
firewall-cmd --permanent --add-service=https
firewall-cmd --permanent --add-port=80/tcp
firewall-cmd --permanent --add-port=8080/tcp
firewall-cmd --permanent --add-port=443/tcp
firewall-cmd --reload
```

If you're planning to use additional BlueXP classification hosts as scanner nodes (in a distributed model),

add these rules to your primary system at this time:

```
firewall-cmd --permanent --add-port=2377/tcp
firewall-cmd --permanent --add-port=7946/udp
firewall-cmd --permanent --add-port=7946/tcp
firewall-cmd --permanent --add-port=4789/udp
```

Note that you must restart Docker or Podman whenever you enable or update `firewalld` settings.

Enable outbound internet access from BlueXP classification

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



This section is not required for host systems installed in sites without internet connectivity.

Endpoints	Purpose
https://api.bluexp.netapp.com	Communication with the BlueXP service, which includes NetApp accounts.
https://netapp-cloud-account.auth0.com https://auth0.com	Communication with the BlueXP website for centralized user authentication.
https://support.compliance.api.bluexp.netapp.com/ https://hub.docker.com https://auth.docker.io https://registry-1.docker.io https://index.docker.io/ https://dseasb33srrn.cloudfront.net/ https://production.cloudflare.docker.com/	Provides access to software images, manifests, templates, and to send logs and metrics.
https://support.compliance.api.bluexp.netapp.com/	Enables NetApp to stream data from audit records.
https://github.com/docker https://download.docker.com	Provides prerequisite packages for docker installation.
http://packages.ubuntu.com/ http://archive.ubuntu.com	Provides prerequisite packages for Ubuntu installation.

Verify that all required ports are enabled

You must ensure that all required ports are open for communication between the Connector, BlueXP classification, Active Directory, and your data sources.

Connection Type	Ports	Description
Connector <> BlueXP classification	8080 (TCP), 443 (TCP), and 80. 9000	<p>The firewall or routing rules for the Connector must allow inbound and outbound traffic over port 443 to and from the BlueXP classification instance.</p> <p>Make sure port 8080 is open so you can see the installation progress in BlueXP.</p> <p>If a firewall is used on the Linux host, port 9000 is required for internal processes within an Ubuntu server.</p>
Connector <> ONTAP cluster (NAS)	443 (TCP)	BlueXP discovers ONTAP clusters using HTTPS. If you use custom firewall policies, the Connector host must allow outbound HTTPS access through port 443. If the Connector is in the cloud, all outbound communication is allowed by the predefined firewall or routing rules.

Run the BlueXP classification Prerequisites script

Follow these steps to run the BlueXP classification Prerequisites script.

[Watch this video](#) to see how to run the Prerequisites script and interpret the results.

Before you begin

- Verify that your Linux system meets the [host requirements](#).
- Verify that the system has the two prerequisite software packages installed (Docker Engine or Podman, and Python 3).
- Make sure you have root privileges on the Linux system.

Steps

1. Download the BlueXP classification Prerequisites script from the [NetApp Support Site](#). The file you should select is named **standalone-pre-requisite-tester-<version>**.
2. Copy the file to the Linux host you plan to use (using `scp` or some other method).
3. Assign permissions to run the script.

```
chmod +x standalone-pre-requisite-tester-v1.25.0
```

4. Run the script using the following command.

```
./standalone-pre-requisite-tester-v1.25.0 <--darksite>
```

Add the option "--darksite" only if you are running the script on a host that doesn't have internet access. Certain prerequisite tests are skipped when the host is not connected to the internet.

5. The script prompts you for the IP address of the BlueXP classification host machine.

- Enter the IP address or host name.
- 6. The script prompts whether you have an installed BlueXP Connector.
 - Enter **N** if you do not have an installed Connector.
 - Enter **Y** if you do have an installed Connector. And then enter the IP address or host name of the BlueXP Connector so the test script can test this connectivity.
- 7. The script runs a variety of tests on the system and it displays results as it progresses. When it finishes it writes a log of the session to a file named `prerequisites-test-<timestamp>.log` in the directory `/opt/netapp/install_logs`.

Result

If all the prerequisites tests ran successfully, you can install BlueXP classification on the host when you are ready.

If any issues were discovered, they are categorized as "Recommended" or "Required" to be fixed. Recommended issues are typically items that would make the BlueXP classification scanning and categorizing tasks run slower. These items do not need to be corrected - but you may want to address them.

If you have any "Required" issues, you should fix the issues and run the Prerequisites test script again.

Activate scanning on your data sources

Scan data sources overview with BlueXP classification

BlueXP classification scans the data in the repositories (the volumes, database schemas, or other user data) that you select to identify personal and sensitive data. BlueXP classification then maps your organizational data, categorizes each file, and identifies predefined patterns in the data. The result of the scan is an index of personal information, sensitive personal information, data categories, and file types.

After the initial scan, BlueXP classification continuously scans your data in a round-robin fashion 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 or at the database schema level.

What's the difference between Mapping and Classification scans

You can conduct two types of scans in BlueXP classification:

- **Mapping-only scans** provide only a high-level overview of your data and are performed on selected data sources. Mapping-only scans take less time than map and classify scans because they do not access files to see the data inside. You might want to do this initially to identify areas of research and then perform a Map & Classify scan on those areas.
- **Map & Classify scans** provide deep-level scanning of your data.

The table below shows some of the differences:

Feature	Map & classify scans	Mapping-only scans
Scan speed	Slow	Fast

Feature	Map & classify scans	Mapping-only scans
Pricing	Free	Free
Capacity	Limited to 500 TiB*	Limited to 500 TiB*
List of file types and used capacity	Yes	Yes
Number of files and used capacity	Yes	Yes
Age and size of files	Yes	Yes
Ability to run a Data Mapping Report	Yes	Yes
Data Investigation page to view file details	Yes	No
Search for names within files	Yes	No
Create saved searches that provide custom search results	Yes	No
Ability to run other reports	Yes	No
Ability to see metadata from files*	No	Yes

* include::_include/connector-limit.adoc[]

*The following metadata is extracted from files during mapping scans:

- Working environment
- Working environment type
- Storage repository
- File type
- Used capacity
- Number of files
- File size
- File creation
- File last access
- File last modified
- File discovered time
- Permissions extraction

Governance dashboard differences:

Feature	Map & Classify	Map
Stale data	Yes	Yes
Non-business data	Yes	Yes
Duplicated files	Yes	Yes
Predefined saved searches	Yes	No
Default saved searches	Yes	Yes
DDA report	Yes	Yes
Mapping report	Yes	Yes
Sensitivity level detection	Yes	No
Sensitive data with wide permissions	Yes	No
Open permissions	Yes	Yes
Age of data	Yes	Yes
Size of data	Yes	Yes
Categories	Yes	No
File types	Yes	Yes

Compliance dashboard differences:

Feature	Map & Classify	Map
Personal information	Yes	No
Sensitive personal information	Yes	No
Privacy risk assessment report	Yes	No
HIPAA report	Yes	No
PCI DSS report	Yes	No

Investigation filters differences:

Feature	Map & Classify	Map
Saved searches	Yes	Yes
Working environment type	Yes	Yes
Working environment	Yes	Yes
Storage repository	Yes	Yes
File type	Yes	Yes
File size	Yes	Yes
Created time	Yes	Yes
Discovered time	Yes	Yes
Last modified	Yes	Yes
Last access	Yes	Yes
Open permissions	Yes	Yes
File directory path	Yes	Yes
Category	Yes	No
Sensitivity level	Yes	No
Number of identifiers	Yes	No
Personal data	Yes	No
Sensitive personal data	Yes	No
Data subject	Yes	No
Duplicates	Yes	Yes
Classification status	Yes	Status is always "Limited insights"
Scan analysis event	Yes	Yes
File hash	Yes	Yes
Number of users with access	Yes	Yes
User/group permissions	Yes	Yes
File owner	Yes	Yes
Directory type	Yes	Yes

How quickly does BlueXP classification scan data

The scan speed is affected by network latency, disk latency, network bandwidth, environment size, and file distribution sizes.

- When performing Mapping-only scans, BlueXP classification can scan between 100-150 TiBs of data per

day.

- When performing Map & classify scans, BlueXP classification can scan between 15-40 TiBs of data per day.

Scan Azure NetApp Files volumes with BlueXP classification

Complete a few steps to get started with BlueXP classification for Azure NetApp Files.

Discover the Azure NetApp Files system that you want to scan

If the Azure NetApp Files system you want to scan is not already in BlueXP as a working environment, you can add it to the canvas at this time.

[See how to discover the Azure NetApp Files system in BlueXP.](#)

Deploy the BlueXP classification instance

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

BlueXP classification 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.

Note: Deploying BlueXP classification in an on-premises location is not currently supported when scanning Azure NetApp Files volumes.

Enable BlueXP classification in your working environments

You can enable BlueXP classification on your Azure NetApp Files volumes.

1. From the BlueXP left navigation menu, click **Governance > Classification**.
2. From the BlueXP classification menu, select **Configuration**.



3. Select how you want to scan the volumes in each working environment. [Learn about mapping and classification scans:](#)
 - To map all volumes, select **Map all Volumes**.
 - To map and classify all volumes, select **Map & Classify all Volumes**.
 - To customize scanning for each volume, select **Or select scanning type for each volume**, and then choose the volumes you want to map and/or classify.

See [Enable and disable compliance scans on volumes](#) for details.

4. In the confirmation dialog box, select **Approve** to have BlueXP classification start scanning your volumes.

Result

BlueXP classification starts scanning the volumes you selected in the working environment. Results will be available in the Compliance dashboard as soon as BlueXP classification finishes the initial scans. The time that it takes depends on the amount of data—it could be a few minutes or hours. You can track the progress of the initial scan by navigating to the **Configuration** menu then selecting the **Working Environment configuration**. The progress of each scan is shown as a progress bar. You can also hover over the progress bar to see the number of files scanned relative to the total files in the volume.



- By default, if BlueXP classification doesn't have write attributes permissions in CIFS, or write permissions in NFS, the system won't scan the files in your volumes because BlueXP classification can't revert the "last access time" to the original timestamp. If you don't care if the last access time is reset, click **Or select scanning type for each volume**. The resulting page has a setting you can enable so that BlueXP classification will scan the volumes regardless of permissions.
- BlueXP classification scans only one file share under a volume. If you have multiple shares in your volumes, you'll need to scan those other shares separately as a shares group. [See more details about this BlueXP classification limitation.](#)

Verify that BlueXP classification has access to volumes

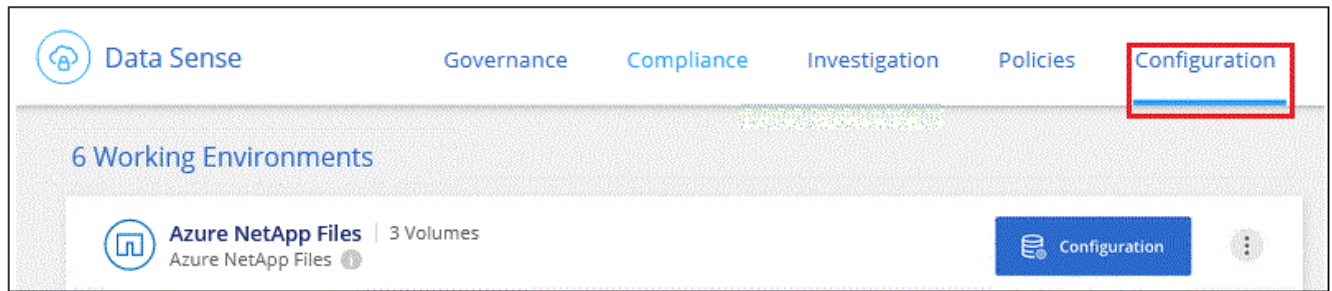
Make sure that BlueXP classification can access volumes by checking your networking, security groups, and export policies. You'll need to provide BlueXP classification with CIFS credentials so it can access CIFS volumes.



For Azure NetApp Files, BlueXP classification can only scan volumes that are in the same region as BlueXP.

Steps

1. Make sure that there's a network connection between the BlueXP classification instance and each network that includes volumes for Azure NetApp Files.
2. Ensure the following ports are open to the BlueXP classification instance:
 - For NFS – ports 111 and 2049.
 - For CIFS – ports 139 and 445.
3. Ensure that NFS volume export policies include the IP address of the BlueXP classification instance so it can access the data on each volume.
4. If you use CIFS, provide BlueXP classification with Active Directory credentials so it can scan CIFS volumes.
 - a. From the BlueXP left navigation menu, select **Governance > Classification**.
5. From the BlueXP classification menu, select **Configuration**.

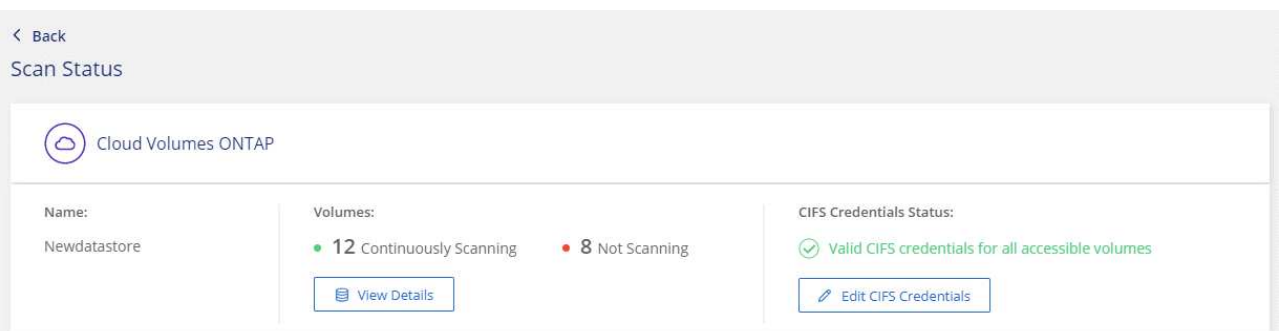


- a. For each working environment, select **Edit CIFS Credentials** and enter the user name and password that BlueXP classification needs to access CIFS volumes on the system.

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

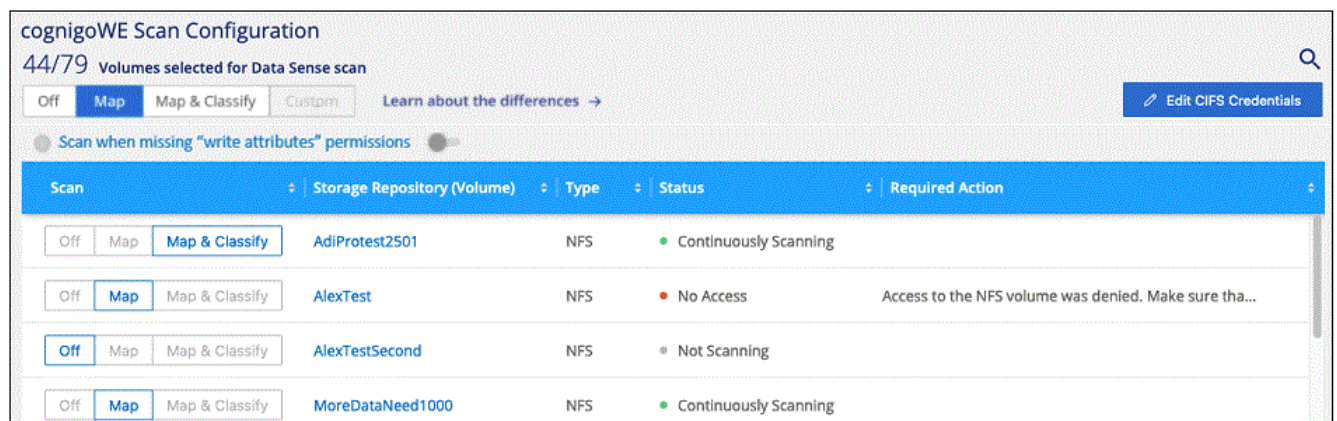
If you want to make sure your files "last accessed times" are unchanged by BlueXP classification scans, it's recommended the user has Write Attributes permissions in CIFS or write permissions in NFS. If possible, configure the Active Directory user as part of a parent group in the organization which has permissions to all files.

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



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

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



Enable and disable compliance scans on volumes

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



New volumes added to the working environment are automatically scanned only when you have set the **Map** or **Map & Classify** setting in the heading area. When set to **Custom** or **Off** in the heading area, you'll need to activate mapping and/or full scanning on each new volume you add in the working environment.

The switch at the top of the page for **Scan when missing "write attributes" permissions** is disabled by default. This means that if BlueXP classification doesn't have write attributes permissions in CIFS, or write permissions in NFS, that the system won't scan the files because BlueXP classification can't revert the "last access time" to the original timestamp. If you don't care if the last access time is reset, turn the switch ON and all files are scanned regardless of the permissions. [Learn more](#).

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

Steps

1. From the BlueXP classification menu, select **Configuration**.
2. Do one of the following:
 - To enable mapping-only scans on a volume, in the volume area, select **Map**. To enable on all volumes, in the heading area, select **Map**.
 - To enable full scanning on a volume, in the volume area, select **Map & Classify**. To enable on all volumes, in the heading area, select **Map & Classify**.
 - To disable scanning on a volume, in the volume area, select **Off**. To disable scanning on all volumes, in the heading area, select **Off**.

Scan Amazon FSx for ONTAP volumes with BlueXP classification

Complete a few steps to get started scanning Amazon FSx for ONTAP volume with BlueXP classification.

Before you begin

- You need an active Connector in AWS to deploy and manage BlueXP classification.
- The security group you selected when creating the working environment must allow traffic from the BlueXP

classification instance. You can find the associated security group using the ENI connected to the FSx for ONTAP file system and edit it using the AWS Management Console.

[AWS security groups for Linux instances](#)

[AWS security groups for Windows instances](#)

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

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

Deploy the BlueXP classification instance

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

You should deploy BlueXP classification in the same AWS network as the Connector for AWS and the FSx volumes you wish to scan.

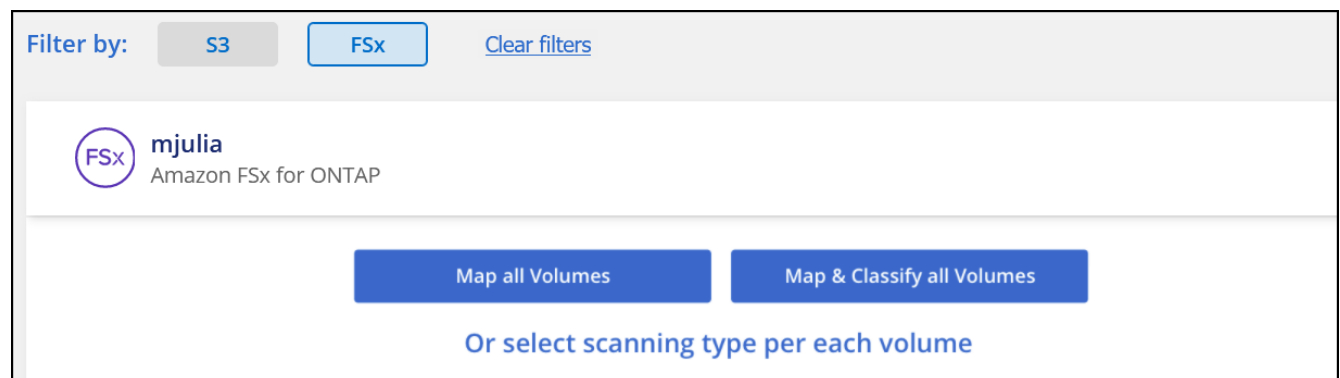
Note: Deploying BlueXP classification in an on-premises location is not currently supported when scanning FSx volumes.

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

Enable BlueXP classification in your working environments

You can enable BlueXP classification for FSx for ONTAP volumes.

1. From the BlueXP left navigation menu, select **Governance > Classification**.
2. From the BlueXP classification menu, select **Configuration**.



3. Select how you want to scan the volumes in each working environment. [Learn about mapping and classification scans](#):
 - To map all volumes, click **Map all Volumes**.
 - To map and classify all volumes, click **Map & Classify all Volumes**.
 - To customize scanning for each volume, click **Or select scanning type for each volume**, and then choose the volumes you want to map and/or classify.
4. In the confirmation dialog box, click **Approve** to have BlueXP classification start scanning your volumes.

Result

BlueXP classification starts scanning the volumes you selected in the working environment. Results will be available in the Compliance dashboard as soon as BlueXP classification finishes the initial scans. The time that it takes depends on the amount of data—it could be a few minutes or hours. You can track the progress of the initial scan by navigating to the **Configuration** menu then selecting the **Working Environment configuration**. The progress of each scan is shown as a progress bar. You can also hover over the progress bar to see the number of files scanned relative to the total files in the volume.



- By default, if BlueXP classification doesn't have write attributes permissions in CIFS, or write permissions in NFS, the system won't scan the files in your volumes because BlueXP classification can't revert the "last access time" to the original timestamp. If you don't care if the last access time is reset, click **Or select scanning type for each volume**. The resulting page has a setting you can enable so that BlueXP classification will scan the volumes regardless of permissions.
- BlueXP classification scans only one file share under a volume. If you have multiple shares in your volumes, you'll need to scan those other shares separately as a shares group. [See more details about this BlueXP classification limitation.](#)

Verify that BlueXP classification has access to volumes

Make sure BlueXP classification can access volumes by checking your networking, security groups, and export policies.

You'll need to provide BlueXP classification with CIFS credentials so it can access CIFS volumes.

Steps

1. From the BlueXP classification menu, select **Configuration**.
2. On the Configuration page, select **View Details** to review the status and correct any errors.

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

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

3. Make sure there's a network connection between the BlueXP classification instance and each network that includes volumes for FSx for ONTAP.



For FSx for ONTAP, BlueXP classification can scan volumes only in the same region as BlueXP.

4. Ensure NFS volume export policies include the IP address of the BlueXP classification instance so it can access the data on each volume.
5. If you use CIFS, provide BlueXP classification with Active Directory credentials so it can scan CIFS volumes.
 - a. From the BlueXP classification menu, select **Configuration**.
 - b. For each working environment, select **Edit CIFS Credentials** and enter the user name and password that BlueXP classification needs to access CIFS volumes on the system.

The credentials can be read-only, but providing admin credentials ensures that BlueXP classification

can read any data that requires elevated permissions. The credentials are stored on the BlueXP classification instance.

If you want to make sure your files "last accessed times" are unchanged by BlueXP classification scans, it's recommended the user has Write Attributes permissions in CIFS or write permissions in NFS. If possible, configure the Active Directory user as part of a parent group in the organization which has permissions to all files.

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

Enable and disable compliance scans on volumes

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

The switch at the top of the page for **Scan when missing "write attributes" permissions** is disabled by default. This means that if BlueXP classification doesn't have write attributes permissions in CIFS, or write permissions in NFS, that the system won't scan the files because BlueXP classification can't revert the "last access time" to the original timestamp. If you don't care if the last access time is reset, turn the switch ON and all files are scanned regardless of the permissions. [Learn more.](#)

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

1. From the BlueXP classification menu, select **Configuration**.
2. In the Configuration page, locate the working environment with the volumes you want to scan.
3. Do one of the following:
 - To enable mapping-only scans on a volume, in the volume area, select **Map**. Or, to enable on all volumes, in the heading area, select **Map**.
 - To enable full scanning on a volume, in the volume area, select **Map & Classify**. Or, to enable on all volumes, in the heading area, select **Map & Classify**.
 - To disable scanning on a volume, in the volume area, select **Off**. To disable scanning on all volumes, in the heading area, select **Off**.



New volumes added to the working environment are automatically scanned only when you have set the **Map** or **Map & Classify** setting in the heading area. When set to **Custom** or **Off** in the heading area, you'll need to activate mapping and/or full scanning on each new volume you add in the working environment.

Scan data protection volumes

By default, data protection (DP) volumes are not scanned because they are not exposed externally and BlueXP classification cannot access them. These are the destination volumes for SnapMirror operations from an FSx for ONTAP file system.

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

The screenshot shows the 'Working Environment Name' Configuration page. At the top, it says '22/28 Volumes selected for compliance scan'. There are buttons for 'Enable Access to DP Volumes' (highlighted with a red box) and 'Edit CIFS Credentials'. Below this is a table with columns: Scan, Storage Repository (Volume), Type, Status, and Required Action. The table lists three volumes: VolumeName1 (Type DP, Status Not Scanning, Required Action Enable access to DP Volumes), VolumeName2 (Type NFS, Status Continuously Scanning), and VolumeName3 (Type CIFS, Status Not Scanning). Each volume has buttons for 'Off', 'Map', and 'Map & Classify'.

Steps

If you want to scan these data protection volumes:

1. From the BlueXP classification menu, select **Configuration**.
2. Select **Enable Access to DP volumes** at the top of the page.
3. Review the confirmation message and select **Enable Access to DP volumes** again.
 - Volumes that were initially created as NFS volumes in the source FSx for ONTAP file system are enabled.
 - Volumes that were initially created as CIFS volumes in the source FSx for ONTAP file system require that you enter CIFS credentials to scan those DP volumes. If you already entered Active Directory credentials so that BlueXP classification can scan CIFS volumes you can use those credentials, or you can specify a different set of Admin credentials.

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

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

4. Activate each DP volume that you want to scan.

Result

Once enabled, BlueXP classification creates an NFS share from each DP volume that was activated for

scanning. The share export policies only allow access from the BlueXP classification instance.

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



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

Scan Cloud Volumes ONTAP and on-premises ONTAP volumes with BlueXP classification

Complete a few steps to start scanning your Cloud Volumes ONTAP and on-premises ONTAP volumes using BlueXP classification.

Prerequisites

Before you enable BlueXP classification, make sure you have a supported configuration.

- If you are scanning Cloud Volumes ONTAP and on-premises ONTAP systems that are accessible over the internet, you can [deploy BlueXP classification in the cloud](#) or [in an on-premises location that has internet access](#).
- If you are scanning on-premises ONTAP systems that have 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.

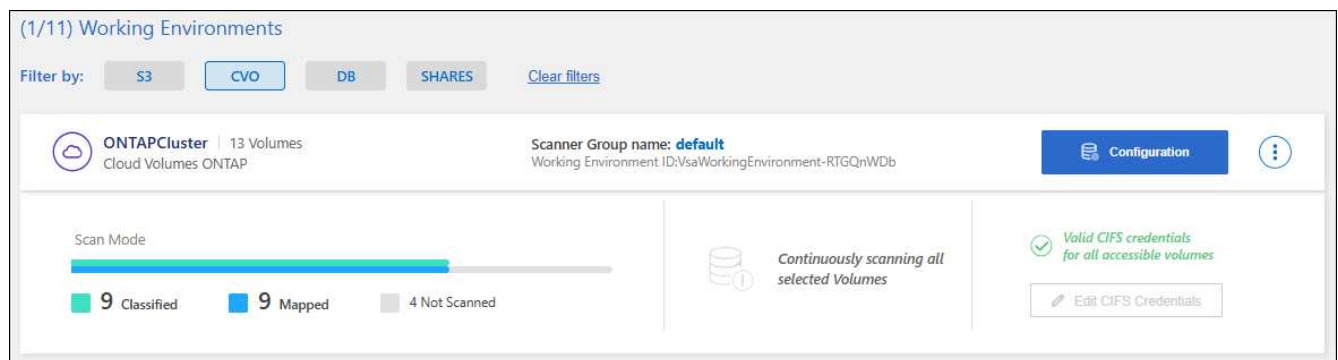
Enable BlueXP classification scanning in your working environments

You can enable BlueXP classification scanning on Cloud Volumes ONTAP systems in any supported cloud provider, and on on-premises ONTAP clusters.

Steps

1. From the BlueXP left navigation menu, select **Governance > Classification**.
2. From the BlueXP classification menu, select **Configuration**.

The Configuration page shows multiple working environments.



3. Choose a working environment and select **Configuration**.

ONTAPCluster Scan Configuration

Volumes selected for Classification scan (9/13)

Off Map Map & Classify Custom Mapping vs. Classification → Retry All Edit CIFS Credentials

Scan when missing "write" permissions ☐

Scan	Storage Repository (Volume)	Type	Mapping status	Scan progress	Required Action
Off Map Map & Classify	bank_statements	NFS	<div> <div>Error 2025-01-09 18:53</div> <div>Last full cycle: 2025-01-09 18:48</div> </div>	Mapped 210 Classified 210	<div> <div>✖ Retry</div> <div>...</div> </div>
Off Map Map & Classify	cifs_labs	CIFS			...
Off Map Map & Classify	cifs_labs_second	CIFS			...
Off Map Map & Classify	datasence	NFS	<div> <div>Error 2025-01-12 06:11</div> <div>Last full cycle: 2025-01-12 06:06</div> </div>	Mapped 127K Classified 127K	<div> <div>✖ Retry</div> <div>...</div> </div>
Off Map Map & Classify	german_data	NFS	<div> <div>Error 2024-10-10 01:35</div> <div>Last full cycle: 2024-10-10 01:29</div> </div>	Mapped 13 Classified 13	<div> <div>✖ Retry</div> <div>...</div> </div>
Off Map Map & Classify	german_data_share	CIFS			...

1-13 of 13

- If you don't care if the last access time is reset, turn the **Scan when missing "write attributes" permissions** switch ON and all files are scanned regardless of the permissions.

The switch at the top of the page for **Scan when missing "write attributes" permissions** is disabled by default. This means that if BlueXP classification doesn't have write attributes permissions in CIFS, or write permissions in NFS, that the system won't classify the files because BlueXP classification can't revert the "last access time" to the original timestamp. [Learn more](#).

- Select how you want to scan the volumes in each working environment. [Learn about mapping and classification scans](#):
 - To map all volumes, select **Map**.
 - To map and classify all volumes, select **Map & Classify**.
 - To customize scanning for each volume, select **Custom**, and then choose the volumes you want to map and/or classify.
- In the confirmation dialog box, select **Approve** to have BlueXP classification start scanning your volumes.

Result

BlueXP classification starts scanning the volumes you selected in the working environment. Results start to appear in the Compliance dashboard as soon as BlueXP classification starts the scan. The time that it takes to complete depends on the amount of data—it could be a few minutes or hours.



BlueXP classification scans only one file share under a volume. If you have multiple shares in your volumes, you'll need to scan those other shares separately as a shares group. [See more details about this BlueXP classification limitation](#).

Verify that BlueXP classification has access to volumes

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

Steps

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

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

3. Ensure that NFS volume export policies include the IP address of the BlueXP classification instance so it can access the data on each volume.
4. If you use CIFS, provide BlueXP classification with Active Directory credentials so it can scan CIFS volumes.
 - a. From the BlueXP left navigation menu, select **Governance > Classification**.
 - b. From the BlueXP classification menu, select **Configuration**.

ONTAPCluster Scan Configuration

Volumes selected for Classification scan (9/13)

Off Map Map & Classify Custom Mapping vs. Classification → Retry All Edit CIFS Credentials

Scan when missing "write" permissions

Scan	Storage Repository (Volume)	Type	Mapping status	Scan progress	Required Action
Off Map Map & Classify	bank_statements	NFS	Error 2025-01-09 18:53 Last full cycle: 2025-01-09 18:48	Mapped 210 Classified 210	Retry
Off Map Map & Classify	cifs_labs	CIFS			
Off Map Map & Classify	cifs_labs_second	CIFS			
Off Map Map & Classify	datasence	NFS	Error 2025-01-12 06:11 Last full cycle: 2025-01-12 06:06	Mapped 127K Classified 127K	Retry
Off Map Map & Classify	german_data	NFS	Error 2024-10-10 01:35 Last full cycle: 2024-10-10 01:29	Mapped 13 Classified 13	Retry
Off Map Map & Classify	german_data_share	CIFS			

1-13 of 13

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

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

If you want to make sure your files "last accessed times" are unchanged by BlueXP classification scans, it's recommended the user has Write Attributes permissions in CIFS or write permissions in NFS. If possible, configure the Active Directory user as part of a parent group in the organization which has permissions to all files.

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

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

Disable compliance scans on volumes

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



New volumes added to the working environment are automatically scanned only when you have set the **Map** or **Map & Classify** setting in the heading area. When the option is set to **Custom** or **Off** in the heading area, you'll need to activate mapping and/or full scanning on each new volume you add in the working environment.

Steps

- From the BlueXP classification menu, select **Configuration**.
- Select the **Configuration** button for the working environment that you want to change.

ONTAPCluster Scan Configuration

Volumes selected for Classification scan (9/13)

Off Map Map & Classify Custom Mapping vs. Classification → Retry All Edit CIFS Credentials

Scan when missing "write" permissions

Scan	Storage Repository (Volume)	Type	Mapping status	Scan progress	Required Action
Off Map Map & Classify	bank_statements	NFS	Error 2025-01-09 18:53 Last full cycle: 2025-01-09 18:48	Mapped 210 Classified 210	Retry
Off Map Map & Classify	cifs_labs	CIFS			
Off Map Map & Classify	cifs_labs_second	CIFS			
Off Map Map & Classify	datasence	NFS	Error 2025-01-12 06:11 Last full cycle: 2025-01-12 06:06	Mapped 127K Classified 127K	Retry
Off Map Map & Classify	german_data	NFS	Error 2024-10-10 01:35 Last full cycle: 2024-10-10 01:29	Mapped 13 Classified 13	Retry
Off Map Map & Classify	german_data_share	CIFS			

1-13 of 13

- Do one of the following:

- To disable scanning on a volume, in the volume area, select **Off**.
- To disable scanning on all volumes, in the heading area, select **Off**.

Scan database schemas with BlueXP classification

Complete a few steps to start scanning your database schemas with BlueXP classification.

Review prerequisites

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

Supported databases

BlueXP classification can scan schemas from the following databases:

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



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

Database requirements

Any database with connectivity to the BlueXP classification 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 BlueXP classification system with all the required permissions.

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

Deploy the BlueXP classification instance

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

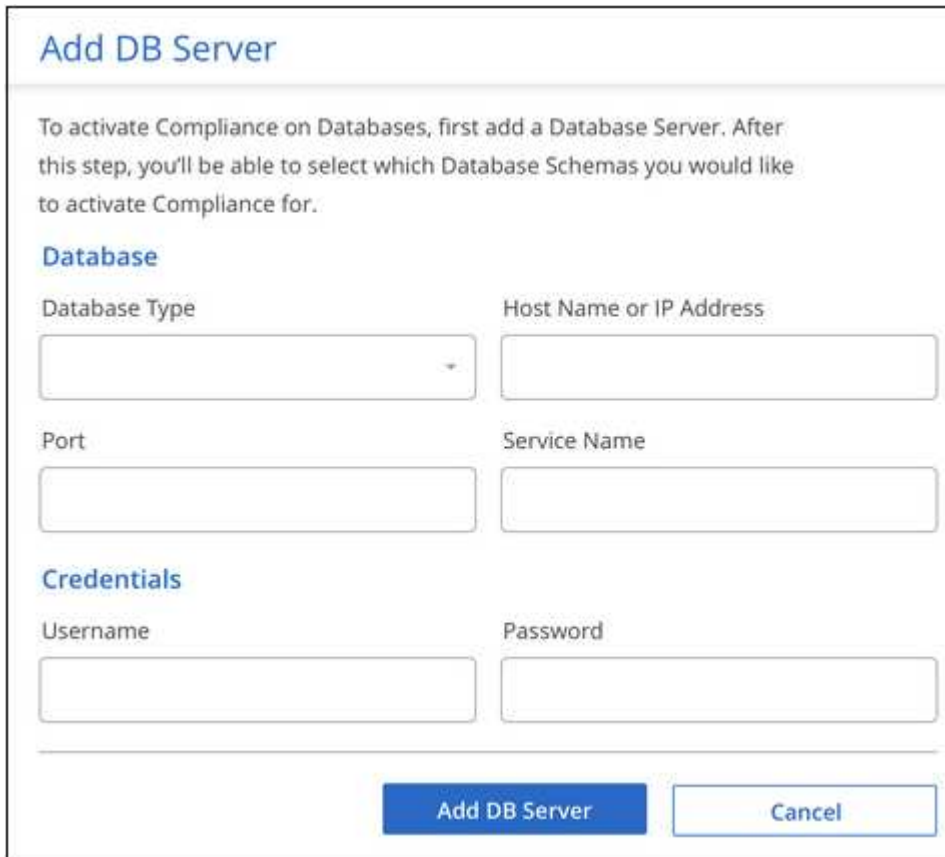
If you are scanning database schemas that are accessible over the internet, you can [deploy BlueXP classification in the cloud](#) or [deploy BlueXP classification in an on-premises location that has internet access](#).

If you are scanning database schemas that have 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.

Add the database server

Add the database server where the schemas reside.

1. From the BlueXP classification menu, select **Configuration**.
2. From the Configuration page, select **Add Working Environment > Add Database Server**.
3. 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 BlueXP classification can access the server.
 - e. Click **Add DB Server**.



The screenshot shows a web form titled "Add DB Server". At the top, there is a blue header bar with the title. Below the header, a paragraph explains that to activate compliance on databases, a database server must first be added. The form is divided into three sections: "Database", "Credentials", and a bottom action bar. The "Database" section contains four input fields: "Database Type" (a dropdown menu), "Host Name or IP Address", "Port", and "Service Name". The "Credentials" section contains two input fields: "Username" and "Password". At the bottom, there are two buttons: "Add DB Server" (a solid blue button) and "Cancel" (a white button with a blue border).

The database is added to the list of working environments.

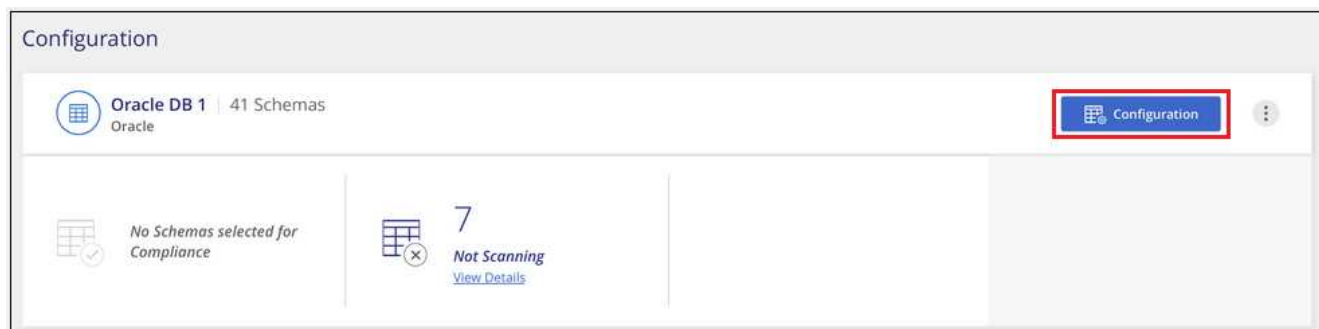
Enable and disable compliance scans on database schemas

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

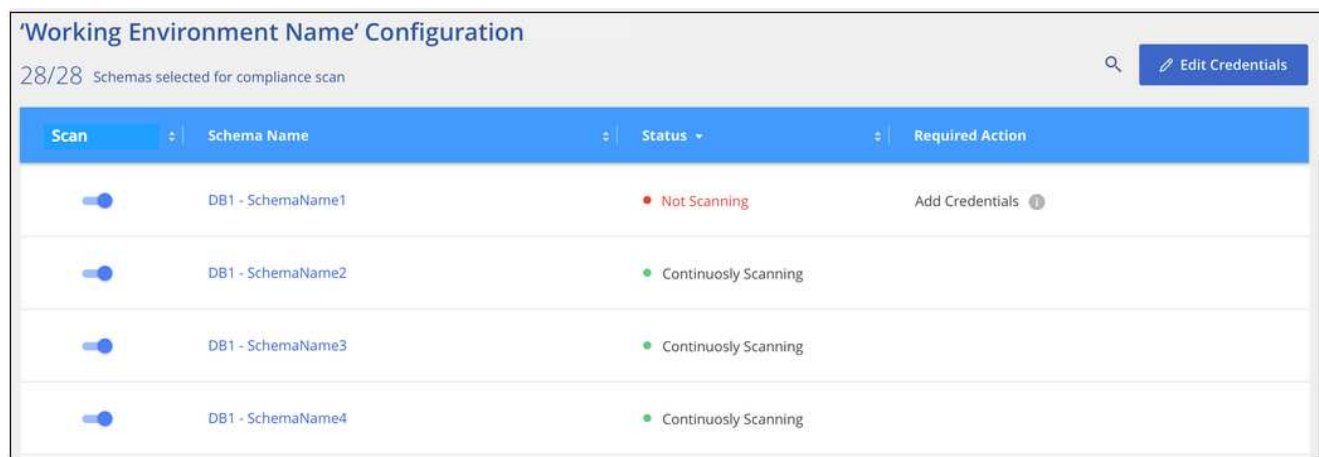


There is no option to select mapping-only scans for database schemas.

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



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



Result

BlueXP classification starts scanning the database schemas that you enabled. You can track the progress of the initial scan by navigating to the **Configuration** menu then selecting the **Working Environment configuration**. The progress of each scan is shown as a progress bar. You can also hover over the progress bar to see the number of files scanned relative to the total files in the volume. If there are any errors, they'll appear in the Status column, alongside the required action to fix the error.

BlueXP classification scans your databases once per day; databases are not continuously scanned like other data sources.

Scan file shares with BlueXP classification

To scan file shares, you must first create a file shares group in BlueXP classification. File shares groups are for NFS or CIFS (SMB) shares hosted on-premises or in the cloud.



Scanning data from non-NetApp file shares is not supported in the BlueXP classification core version.

Prerequisites

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

- The shares can be hosted anywhere, including in the cloud or on-premises. CIFS shares from older NetApp 7-Mode storage systems can be scanned as file shares.
 - BlueXP classification can't extract permissions or the "last access time" from 7-Mode systems.
 - Because of a known issue between some Linux versions and CIFS shares on 7-Mode systems, you must configure the share to use only SMBv1 with NTLM authentication enabled.
- There needs to be network connectivity between the BlueXP classification instance and the shares.
- You can add a DFS (Distributed File System) share as a regular CIFS share. Because BlueXP classification is unaware that the share is built upon multiple servers/volumes combined as a single CIFS share, you might receive permission or connectivity errors about the share when the message really only applies to one of the folders/shares that is located on a different server/volume.
- For CIFS (SMB) shares, ensure that you have Active Directory credentials that provide read access to the shares. Admin credentials are preferred in case BlueXP classification needs to scan any data that requires elevated permissions.

If you want to make sure your files "last accessed times" are unchanged by BlueXP classification scans, it's recommended the user has Write Attributes permissions in CIFS or write permissions in NFS. If possible, configure the Active Directory user as part of a parent group in the organization which has permissions to all files.

- All CIFS file shares in a group must use the same Active Directory credentials.
- You can mix NFS and CIFS (using either Kerberos or NTLM) shares. You must add the shares to the group separately. That is, you must complete the process twice—once per protocol.
 - You cannot create a file shares group that mixes CIFS authentication types (Kerberos and NTLM).
- If you're using CIFS with Kerberos authentication, ensure the IP address provided is accessible to the BlueXP classification service. The file shares can't be added if the IP address is unreachable.

Create a file shares group

When you add file shares to the group, you must use the format `<host_name>:/<share_path>`.

+

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

Steps

1. From the BlueXP classification menu, select **Configuration**.
2. From the Configuration page, select **Add Working Environment > Add File Shares Group**.
3. In the Add File Shares Group dialog, enter the name for the group of shares then select **Continue**.
4. Select the protocol for the file shares you are adding.

.If you're adding CIFS shares with NTLM authentication, enter the Active Directory credentials to access the CIFS volumes. Although read-only credentials are supported, it's recommended you provide full access with administrator credentials. Select Save.

1. Add the file shares that you want to scan (one file share per line). Then select **Continue**.
2. 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. If the issue pertains to a naming convention, you can re-add the share with a corrected name.

3. Configure scanning on the volume:

- To enable mapping-only scans on file shares, select **Map**.
- To enable full scans on file shares, select **Map & Classify**.
- To disable scanning on file shares, select **Off**.



The switch at the top of the page for **Scan when missing "write attributes" permissions** is disabled by default. This means that if BlueXP classification doesn't have write attributes permissions in CIFS or write permissions in NFS, the system won't scan the files because BlueXP classification can't revert the "last access time" to the original timestamp.

If you switch **Scan when missing "write attributes" permissions** to **On**, the scan resets the last accessed time and scans all files regardless of permissions.

To learn more about the last accessed time stamp, see [xref:./Metadata collected from data sources in BlueXP classification](#).

Result

BlueXP classification starts scanning the files in the file shares you added. You can [Track the scanning progress](#) and view the results of the scan in the **Dashboard**.



If the scan doesn't complete successfully for a CIFS configuration with Kerberos authentication, check the **Configuration** tab for errors.

Edit a file shares group

After you create a file shares group, you can edit the CIFS protocol or add and remove file shares.

Edit the CIFS protocol configuration

1. From the BlueXP classification menu, select **Configuration**.
2. From the Configuration page, select the file shares group you want to modify.
3. Select **Edit CIFS Credentials**.

Edit CIFS Authentication

Classification requires Active Directory credentials to access CIFS Volumes in Micky.

The credentials can be read-only, but providing admin credentials ensures that Classification can read any data that requires elevated permissions.

Select Authentication Method

☒ NTLM

☐ Kerberos

Username ⓘ

Password

domain\user or user@domain

Password

Save

Cancel

4. Choose the authentication method: **NTLM** or **Kerberos**.
5. Enter the Active Directory **Username** and **Password**.
6. Select **Save** to complete the process.

Add file shares to compliance scans

1. From the BlueXP classification menu, select **Configuration**.
2. From the Configuration page, select the file shares group you want to modify.
3. Select **+ Add shares**.
4. Select the protocol for the file shares you are adding.

If you're adding file shares to a protocol you've already configured, no changes are required.

If you're adding file shares with a second protocol, ensure you've properly configured the authentication properly as detailed in the [prerequisites](#).

5. Add the file shares you want to scan (one file share per line) using the format `<host_name>:/<share_path>`.
6. Select **Continue** to complete adding the file shares.

Remove a file share from compliance scans

1. From the BlueXP classification menu, select **Configuration**.
2. Select the working environment from which you want to remove file shares.
3. Select **Configuration**.
4. From the Configuration page, select the Actions **...** for the file share you want to remove.
5. From the Actions menu, select **Remove Share**.

Track the scanning progress

You can track the progress of the initial scan.

1. Select the **Configuration** menu.
2. Select the **Working Environment Configuration**.

The progress of each scan is shown as a progress bar.

3. Hover over the progress bar to see the number of files scanned relative to the total files in the volume.

Scan StorageGRID data with BlueXP classification

Complete a few steps to start scanning data within StorageGRID directly with BlueXP classification.

Review StorageGRID 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 StorageGRID so that BlueXP classification can access the buckets.

Deploy the BlueXP classification instance

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

If you are scanning data from StorageGRID 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 StorageGRID 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.

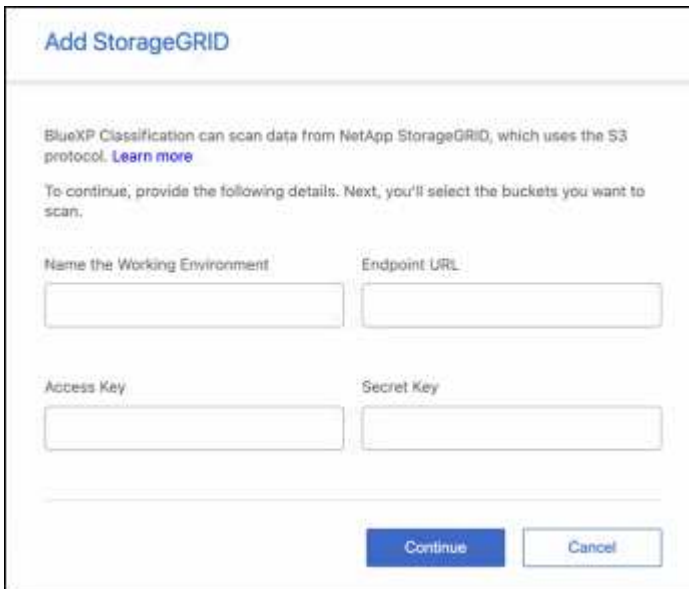
Add the StorageGRID service to BlueXP classification

Add the StorageGRID service.

Steps

1. From the BlueXP classification menu, select the **Configuration** option.
2. From the Configuration page, select **Add Working Environment > Add StorageGRID**.
3. In the Add StorageGRID Service dialog, enter the details for the StorageGRID service and click **Continue**.

- a. Enter the name you want to use for the Working Environment. This name should reflect the name of the StorageGRID 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 StorageGRID.



The screenshot shows a web form titled "Add StorageGRID". Below the title, there is explanatory text: "BlueXP Classification can scan data from NetApp StorageGRID, which uses the S3 protocol. [Learn more](#)". This is followed by instructions: "To continue, provide the following details. Next, you'll select the buckets you want to scan." The form contains four input fields arranged in two rows. The first row has "Name the Working Environment" and "Endpoint URL". The second row has "Access Key" and "Secret Key". At the bottom right of the form are two buttons: "Continue" (in blue) and "Cancel" (in light blue).

Result

StorageGRID is added to the list of working environments.

Enable and disable compliance scans on StorageGRID buckets

After you enable BlueXP classification on StorageGRID, 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, locate the StorageGRID working environment.
2. On the StorageGRID working environment tile, select **Configuration**.

Buckets selected for Classification scan (5/8)

Scan	Storage Repository (Bucket)	Mapping status	Classification status	Required Action
Off Map Map & Classify	bucketadipro	Finished 2024-09-05 10:33 Last full cycle: 2024-09-05 10:33	Mapped: 84 Classified: 5	...
Off Map Map & Classify	datasense-0-files	Finished 2024-09-05 08:00 Last full cycle: 2024-09-05 08:00		...
Off Map Map & Classify	datasense-10tb	Running 2024-09-04 07:25	Mapped: 3.7M Classified: 2.1M	...
Off Map Map & Classify	datasense-1tb	Running 2024-09-05 09:05 Last full cycle: 2024-09-05 03:04	Mapped: 1.3M	...
Off Map Map & Classify	datasense-1tb-2	Running 2024-09-05 09:06 Last full cycle: 2024-09-05 03:05	Mapped: 1.3M	...
Off Map Map & Classify	datasense-1tb-3	Not scanning		...

3. Complete one of the following steps to enable or disable scanning:

- To enable mapping-only scans on a bucket, select **Map**.
- To enable full scans on a bucket, select **Map & Classify**.
- To disable scanning on a bucket, select **Off**.

Result

BlueXP classification starts scanning the buckets that you enabled. You can track the progress of the initial scan by navigating to the **Configuration** menu then selecting the **Working Environment configuration**. The progress of each scan is show as a progress bar. You can also hover over the progress bar to see the number of files scanned relative to the total files in the volume. If there are any errors, they'll appear in the Status column, alongside the required action to fix the error.

Integrate your Active Directory with BlueXP classification

You can integrate a global Active Directory with BlueXP classification to enhance the results that BlueXP classification reports about file owners and which users and groups have access to your files.

When you set up certain data sources (listed below), you need to enter Active Directory credentials in order for BlueXP classification to scan CIFS volumes. This integration provides BlueXP classification with file owner and permissions details for the data that resides in those data sources. The Active Directory entered for those data sources might differ from the global Active Directory credentials you enter here. BlueXP classification will look in all integrated Active Directories for user and permission details.

This integration provides additional information in the following locations in BlueXP classification:

- You can use the "File Owner" [filter](#) and see results in the file's metadata in the Investigation pane. Instead of the file owner containing the SID (Security IDentifier), it is populated with the actual user name.

You can also view more details about the file owner: account name, email address, and SAM account name, or view items owned by that user.

- You can see [full file permissions](#) for each file and directory when you click the "View all Permissions"

button.

- In the [Governance dashboard](#), the Open Permissions panel will show a greater level of detail about your data.



Local user SIDs, and SIDs from unknown domains, are not translated to the actual user name.

Supported data sources

An Active Directory integration with BlueXP classification can identify data from within the following data sources:

- On-premises ONTAP systems
- Cloud Volumes ONTAP
- Azure NetApp Files
- FSx for ONTAP
- OneDrive accounts and SharePoint accounts (for legacy versions 1.30 and earlier)

There is no support for identifying user and permission information from Database schemas, Google Drive accounts, Amazon S3 accounts, or Object Storage that uses the Simple Storage Service (S3) protocol.

Connect to your Active Directory server

After you've deployed BlueXP classification and have activated scanning on your data sources, you can integrate BlueXP classification with your Active Directory. Active Directory can be accessed using a DNS Server IP address or an LDAP Server IP address.

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

For CIFS volumes/file shares, if you want to make sure your files "last accessed times" are unchanged by BlueXP classification scans, we recommend that the user has Write Attributes permission. If possible, we recommend making the Active Directory configured user part of a parent group in the organization which has permissions to all files.

Requirements

- You must have an Active Directory already set up for the users in your company.
- You must have the information for the Active Directory:
 - DNS Server IP address, or multiple IP addressesor
 - LDAP Server IP address, or multiple IP addresses
 - User Name and Password to access the server
 - Domain Name (Active Directory Name)
 - Whether you are using secure LDAP (LDAPS) or not
 - LDAP Server Port (typically 389 for LDAP, and 636 for secure LDAP)

- The following ports must be open for outbound communication by the BlueXP classification instance:

Protocol	Port	Destination	Purpose
TCP & UDP	389	Active Directory	LDAP
TCP	636	Active Directory	LDAP over SSL
TCP	3268	Active Directory	Global Catalog
TCP	3269	Active Directory	Global Catalog over SSL

Steps

1. From the BlueXP classification Configuration page, click **Add Active Directory**.

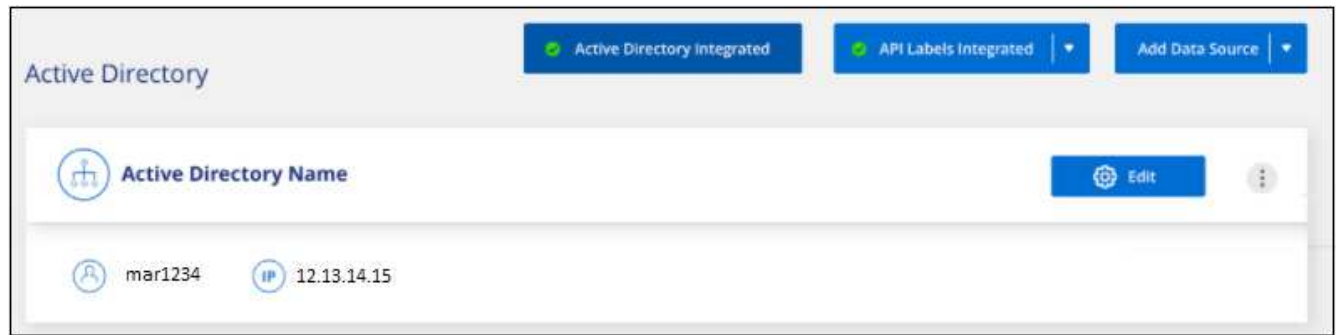


2. In the Connect to Active Directory dialog, enter the Active Directory details and click **Connect**.

You can add multiple IP addresses, if required, by clicking **Add IP**.


 A screenshot of the 'Connect to Active Directory' dialog box. It has a title bar at the top. Below the title, there are two input fields: 'Username' with the value 'mar1234' and 'Password' with masked characters. Below these, there are two sections. The first section is for 'DNS Server IP address' and 'Domain Name'. The 'DNS Server IP address' field contains '12.20.70.00' and has a '+ Add IP' button next to it. The 'Domain Name' field contains 'mar@netapp.com'. The second section is for 'LDAP Server IP Address' and 'LDAP Server Port'. The 'LDAP Server IP Address' field is empty and has a '+ Add IP' button next to it. The 'LDAP Server Port' field contains '389'. There is also a checkbox for 'LDAP Secure Connection' which is currently unchecked. At the bottom of the dialog, there are two buttons: 'Connect' and 'Cancel'. The 'Connect' button is highlighted with a red rectangular box.

BlueXP classification integrates to the Active Directory, and a new section is added to the Configuration page.



Manage your Active Directory integration

If you need to modify any values in your Active Directory integration, click the **Edit** button and make the changes.

You can also delete the integration selecting the  button then **Remove Active Directory**.

Copyright information

Copyright © 2025 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.

LIMITED RIGHTS LEGEND: Use, duplication, or disclosure by the government is subject to restrictions as set forth in subparagraph (b)(3) of the Rights in Technical Data -Noncommercial Items at DFARS 252.227-7013 (FEB 2014) and FAR 52.227-19 (DEC 2007).

Data contained herein pertains to a commercial product and/or commercial service (as defined in FAR 2.101) and is proprietary to NetApp, Inc. All NetApp technical data and computer software provided under this Agreement is commercial in nature and developed solely at private expense. The U.S. Government has a non-exclusive, non-transferrable, nonsublicensable, worldwide, limited irrevocable license to use the Data only in connection with and in support of the U.S. Government contract under which the Data was delivered. Except as provided herein, the Data may not be used, disclosed, reproduced, modified, performed, or displayed without the prior written approval of NetApp, Inc. United States Government license rights for the Department of Defense are limited to those rights identified in DFARS clause 252.227-7015(b) (FEB 2014).

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.