



部署数据分类 NetApp Data Classification

NetApp
January 26, 2026

目录

部署数据分类	1
您应该使用哪种NetApp Data Classification部署?	1
使用NetApp Console在云中部署NetApp Data Classification	1
快速启动	1
创建控制台代理	2
前提条件	2
在云中部署数据分类	6
在可以访问互联网的主机上安装NetApp Data Classification	8
快速启动	9
创建控制台代理	10
准备 Linux 主机系统	10
启用数据分类的出站互联网访问	12
验证所有必需的端口均已启用	13
在 Linux 主机上安装数据分类	14
在没有互联网访问的 Linux 主机上安装NetApp Data Classification	17
检查您的 Linux 主机是否已准备好安装NetApp Data Classification	17
入门指南	18
创建控制台代理	18
验证主机要求	18
启用数据分类的出站互联网访问	20
验证所有必需的端口均已启用	21
运行数据分类先决条件脚本	21

部署数据分类

您应该使用哪种NetApp Data Classification部署？

您可以通过不同的方式部署NetApp Data Classification。了解哪种方法可以满足您的需求。

数据分类可以通过以下方式部署：

- ["使用控制台在云中部署"](#)。控制台将数据分类实例部署在与控制台代理相同的云提供商网络中。
- ["在可以访问互联网的 Linux 主机上安装"](#)。在您的网络中的 Linux 主机上或云中的 Linux 主机上安装数据分类，前提是该主机可以访问互联网。如果您希望使用同样位于本地的数据分类实例来扫描本地ONTAP系统，则这种类型的安装可能是一个不错的选择，尽管这不是必需的。
- ["在没有互联网访问的本地站点的 Linux 主机上安装"](#)，也称为_私人模式_。这种安装类型使用安装脚本，与控制台 SaaS 层没有连接。



BlueXP私有模式（传统BlueXP接口）通常用于没有互联网连接的本地环境和安全云区域，其中包括 AWS Secret Cloud、AWS Top Secret Cloud 和 Azure IL6。NetApp继续通过传统的BlueXP界面支持这些环境。有关旧版BlueXP界面中的私有模式文档，请参阅["BlueXP私人模式的 PDF 文档"](#)。

无论是在有互联网访问的 Linux 主机上安装，还是在没有互联网访问的 Linux 主机上进行本地安装，都使用安装脚本。脚本首先检查系统和环境是否满足先决条件。如果满足先决条件，则安装开始。如果您想独立于运行数据分类安装来验证先决条件，您可以下载一个单独的软件包，该软件包仅测试先决条件。

请参阅["检查您的 Linux 主机是否已准备好安装数据分类"](#)。

使用NetApp Console在云中部署NetApp Data Classification

您可以使用NetApp Console在云中部署NetApp Data Classification。控制台将数据分类实例部署在与控制台代理相同的云提供商网络中。

请注意，您还可以["在可以访问互联网的 Linux 主机上安装数据分类"](#)。如果您希望使用同样位于本地的数据分类实例来扫描本地ONTAP系统，则这种类型的安装可能是一个不错的选择 - 但这不是必需的。无论您选择哪种安装方法，软件的功能都完全相同。

快速启动

按照以下步骤快速开始，或者向下滚动到其余部分以获取完整详细信息。



创建控制台代理

如果您还没有控制台代理，请创建一个。看["在 AWS 中创建控制台代理"](#)，["在 Azure 中创建控制台代理"](#)，或者["在 GCP 中创建控制台代理"](#)。

您也可以["在本地安装控制台代理"](#)在您网络中的 Linux 主机上，或者在云端的 Linux 主机上。

2

先决条件

确保您的环境能够满足先决条件。这包括实例的出站互联网访问、Console 代理与 Data Classification 在端口 443 上的连接等。[查看完整列表](#)。

3

部署数据分类

启动安装向导以在云中部署数据分类实例。

创建控制台代理

如果您还没有控制台代理，请在您的云提供商中创建一个控制台代理。看 ["在 AWS 中创建控制台代理"](#) 或者 ["在 Azure 中创建控制台代理"](#)，或者 ["在 GCP 中创建控制台代理"](#)。大多数情况下，您可能需要在尝试激活数据分类之前设置好控制台代理，因为大多数 ["控制台功能需要控制台代理"](#) 但有些情况下，你需要现在就设置一个。

在某些情况下，您必须使用部署在特定云提供商中的控制台代理：

- 在 AWS 中的 Cloud Volumes ONTAP 或 Amazon FSx for ONTAP 存储桶中扫描数据时，您可以使用 AWS 中的控制台代理。
- 在 Azure 中的 Cloud Volumes ONTAP 或 Azure NetApp Files 中扫描数据时，您可以使用 Azure 中的控制台代理。
 - 对于 Azure NetApp Files，它必须部署在与您要扫描的卷相同的区域中。
- 在 GCP 中的 Cloud Volumes ONTAP 中扫描数据时，您可以使用 GCP 中的控制台代理。

使用任何这些云控制台代理时，都可以扫描本地 ONTAP 系统、NetApp 文件共享和数据库。

请注意，您也可以 ["在本地安装控制台代理"](#) 在网络或云端的 Linux 主机上。一些计划在本地安装数据分类的用户可能还会选择在本地安装控制台代理。

可能有些情况下你需要使用 ["多个控制台代理"](#)。



数据分类不会对其可以扫描的数据量施加限制。每个控制台代理支持扫描和显示 500 TiB 的数据。要扫描超过 500 TiB 的数据，["安装另一个控制台代理"](#) 然后 ["部署另一个数据分类实例"](#)。+ 控制台 UI 显示来自单个连接器的数据。有关查看来自多个控制台代理的数据的提示，请参阅 ["使用多个控制台代理"](#)。

政府区域支持

当控制台代理部署在政府区域（AWS GovCloud、Azure Gov 或 Azure DoD）时，支持数据分类。以这种方式部署时，数据分类具有以下限制：

["了解如何在政府区域部署控制台代理"](#)。

前提条件

在云中部署数据分类之前，请查看以下先决条件，以确保您具有受支持的配置。当您在云中部署数据分类时，它与控制台代理位于同一子网中。

启用数据分类的出站互联网访问

数据分类需要出站互联网访问。如果您的虚拟或物理网络使用代理服务器进行互联网访问，请确保数据分类实例具有出站互联网访问权限以联系以下端点。代理必须是非透明的。目前不支持透明代理。

根据您是在 AWS、Azure 还是 GCP 中部署数据分类，查看下面的相应表格。

AWS 所需的终端节点

端点	目的
\ https://api.console.netapp.com	与控制台服务（包括NetApp帐户）的通信。
\ https://netapp-cloud-account.auth0.com \ https://auth0.com	与控制台网站通信，实现集中用户身份验证。
\ 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://dseasb33srnrn.cloudfront.net/ \ https://production.cloudflare.docker.com/	提供对软件映像、清单和模板的访问。
\ https://kinesis.us-east-1.amazonaws.com	使NetApp能够从审计记录中流式传输数据。
\ 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	启用数据分类来访问和下载清单和模板，以及发送日志和指标。

Azure 所需的终结点

端点	目的
\ https://api.console.netapp.com	与控制台服务（包括NetApp帐户）的通信。
\ https://netapp-cloud-account.auth0.com \ https://auth0.com	与控制台网站通信，实现集中用户身份验证。
\ https://support.compliance.api.console.netapp.com/ \ https://hub.docker.com \ https://auth.docker.io \ https://registry-1.docker.io \ https://index.docker.io/ \ https://dseasb33srnrn.cloudfront.net/ \ https://production.cloudflare.docker.com/	提供对软件映像、清单、模板的访问以及发送日志和指标。
\ https://support.compliance.api.console.netapp.com/	使NetApp能够从审计记录中流式传输数据。

GCP 所需的端点

端点	目的
\ https://api.console.netapp.com	与控制台服务（包括NetApp帐户）的通信。
\ https://netapp-cloud-account.auth0.com \ https://auth0.com	与控制台网站通信，实现集中用户身份验证。

端点	目的
\ https://support.compliance.api.console.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/	提供对软件映像、清单、模板的访问以及发送日志和指标。
\ https://support.compliance.api.console.netapp.com/	使NetApp能够从审计记录中流式传输数据。

确保数据分类具有所需的权限

确保数据分类具有部署资源和为数据分类实例创建安全组的权限。

- ["Google Cloud 权限"](#)
- ["AWS 权限"](#)
- ["Azure 权限"](#)

确保控制台代理可以访问数据分类

确保控制台代理和数据分类实例之间的连接。控制台代理的安全组必须允许通过端口 443 进出数据分类实例的入站和出站流量。此连接支持部署数据分类实例，并允许您查看“合规性和治理”选项卡中的信息。AWS 和 Azure 的政府区域支持数据分类。

AWS 和 AWS GovCloud 部署需要额外的入站和出站安全组规则。看 ["AWS 中的控制台代理规则"](#)了解详情。

Azure 和 Azure 政府部署需要额外的入站和出站安全组规则。看 ["Azure 中的控制台代理规则"](#)了解详情。

确保数据分类能够持续运行

数据分类实例需要保持开启状态以持续扫描您的数据。

确保 Web 浏览器连接到数据分类

启用数据分类后，确保用户从与数据分类实例有连接的主机访问控制台界面。

数据分类实例使用私有 IP 地址来确保索引数据无法被互联网访问。因此，您用来访问控制台的 Web 浏览器必须连接到该私有 IP 地址。该连接可以来自与云提供商的直接连接（例如 VPN），也可以来自与数据分类实例位于同一网络内的主机。

检查您的 vCPU 限制

确保您的云提供商的 vCPU 限制允许部署具有必要数量的核心的实例。您需要验证控制台运行区域中相关实例系列的 vCPU 限制。["查看所需的实例类型"](#)。

有关 vCPU 限制的更多详细信息，请参阅以下链接：

- ["AWS 文档：Amazon EC2 服务配额"](#)
- ["Azure 文档：虚拟机 vCPU 配额"](#)
- ["Google Cloud 文档：资源配额"](#)

在云中部署数据分类

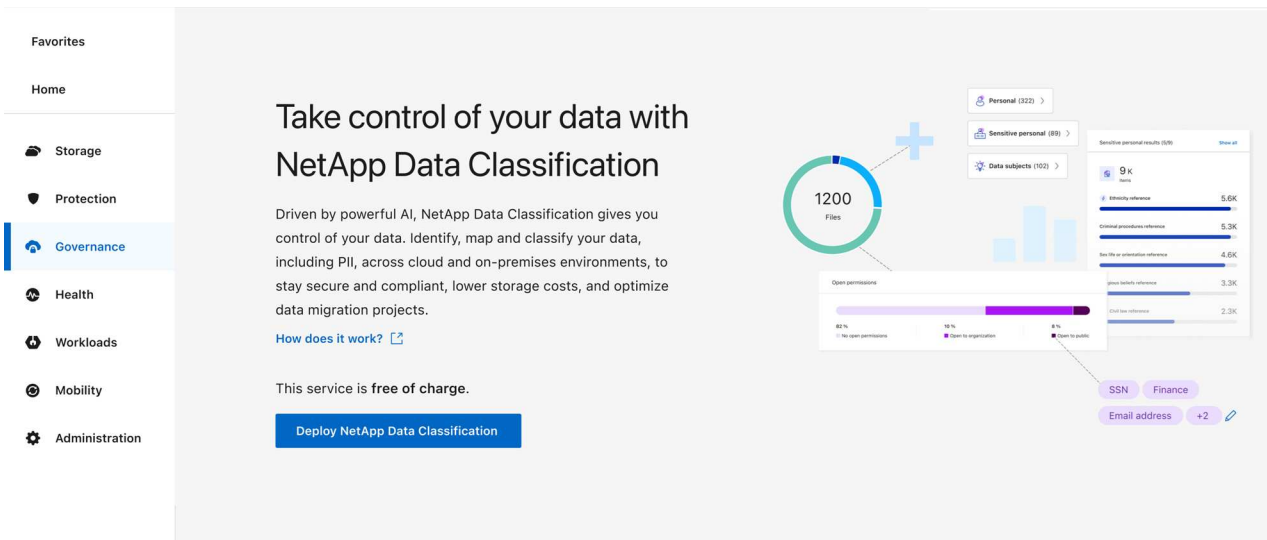
按照以下步骤在云中部署数据分类实例。控制台代理将在云中部署实例，然后在该实例上安装数据分类软件。

在默认实例类型不可用的区域中，数据分类在["备用实例类型"](#)。

在 AWS 中部署

步骤

1. 从数据分类主页中，选择*在本地或云中部署分类*。

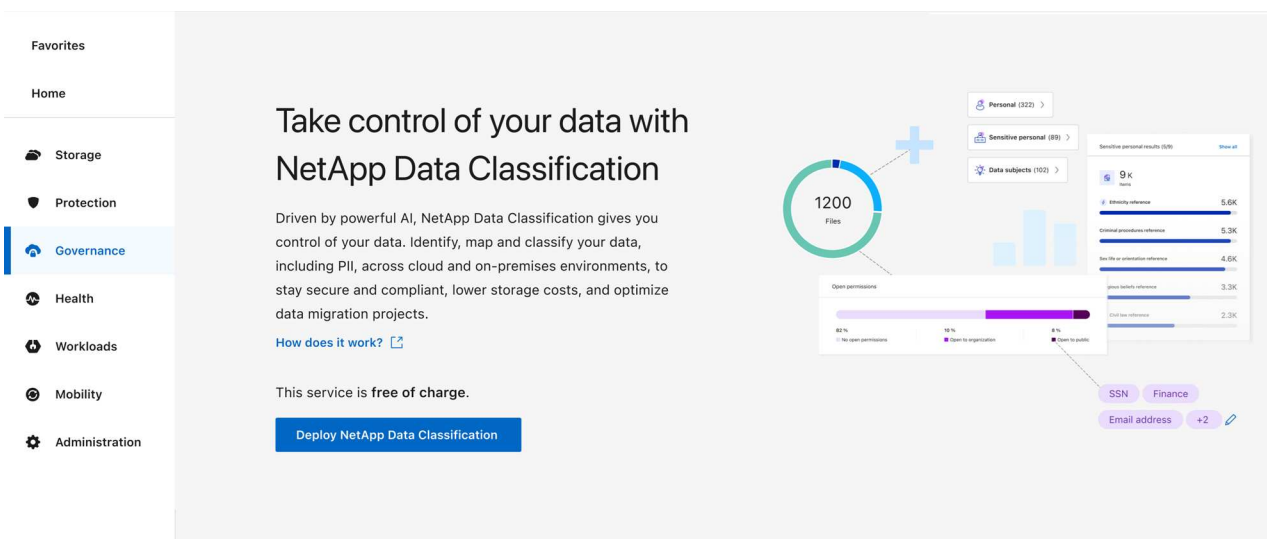


2. 从“安装”页面中，选择“部署”>“部署”以使用“大型”实例大小并启动云部署向导。
3. 向导在执行部署步骤时会显示进度。当需要输入或遇到问题时，系统会提示您。
4. 当实例部署完毕并安装数据分类后，选择“继续配置”进入“配置”页面。

在 Azure 中部署

步骤

1. 从数据分类主页中，选择*在本地或云中部署分类*。

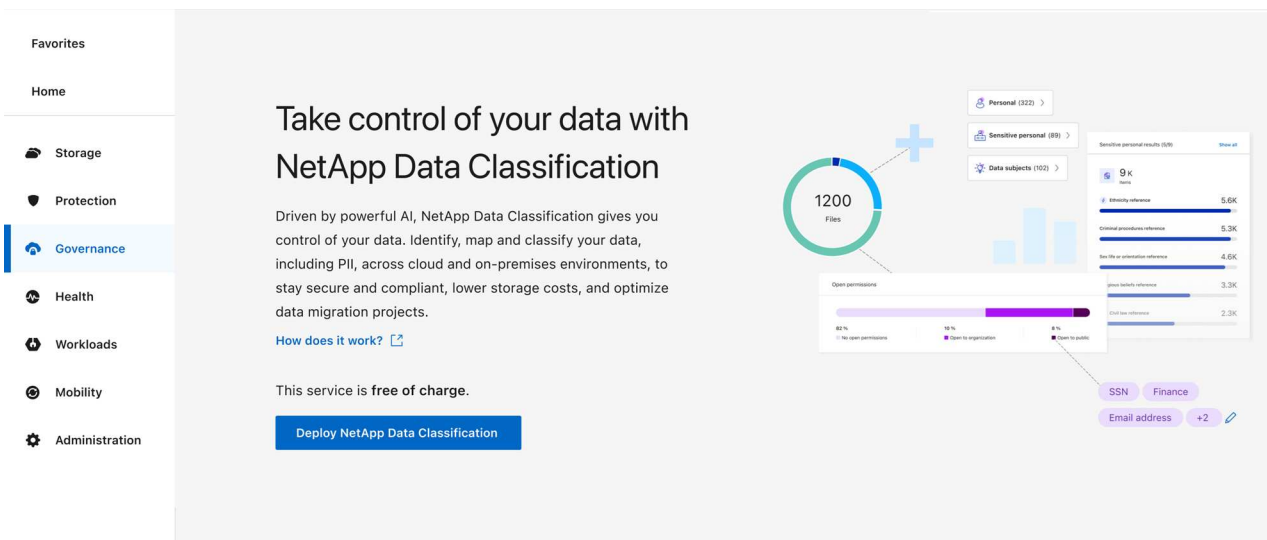


2. 选择*部署*以启动云部署向导。
3. 向导在执行部署步骤时会显示进度。如果遇到任何问题，它将停止并提示输入。
4. 当实例部署完毕并安装数据分类后，选择“继续配置”进入“配置”页面。

在 Google Cloud 中部署

步骤

1. 从数据分类主页中，选择*治理>分类*。
2. 选择*在本地或云中部署分类*。



3. 选择*部署*以启动云部署向导。
4. 向导在执行部署步骤时会显示进度。如果遇到任何问题，它将停止并提示输入。
5. 当实例部署完毕并安装数据分类后，选择“继续配置”进入“配置”页面。

结果

控制台在您的云提供商中部署数据分类实例。

只要实例具有互联网连接，控制台代理和数据分类软件的升级就会自动进行。

下一步

您可以从配置页面选择要扫描的数据源。

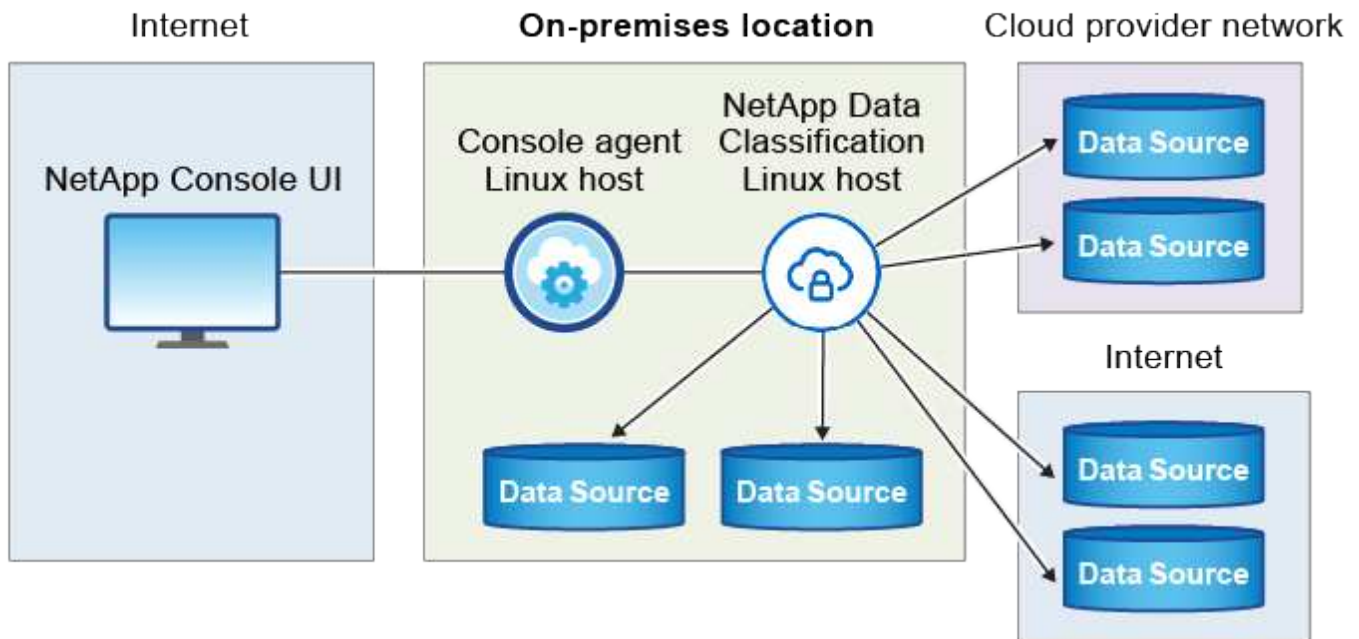
在可以访问互联网的主机上安装NetApp Data Classification

要在您的网络中的 Linux 主机或具有 Internet 访问权限的云中的 Linux 主机上部署NetApp Data Classification，您需要在您的网络或云中手动部署 Linux 主机。

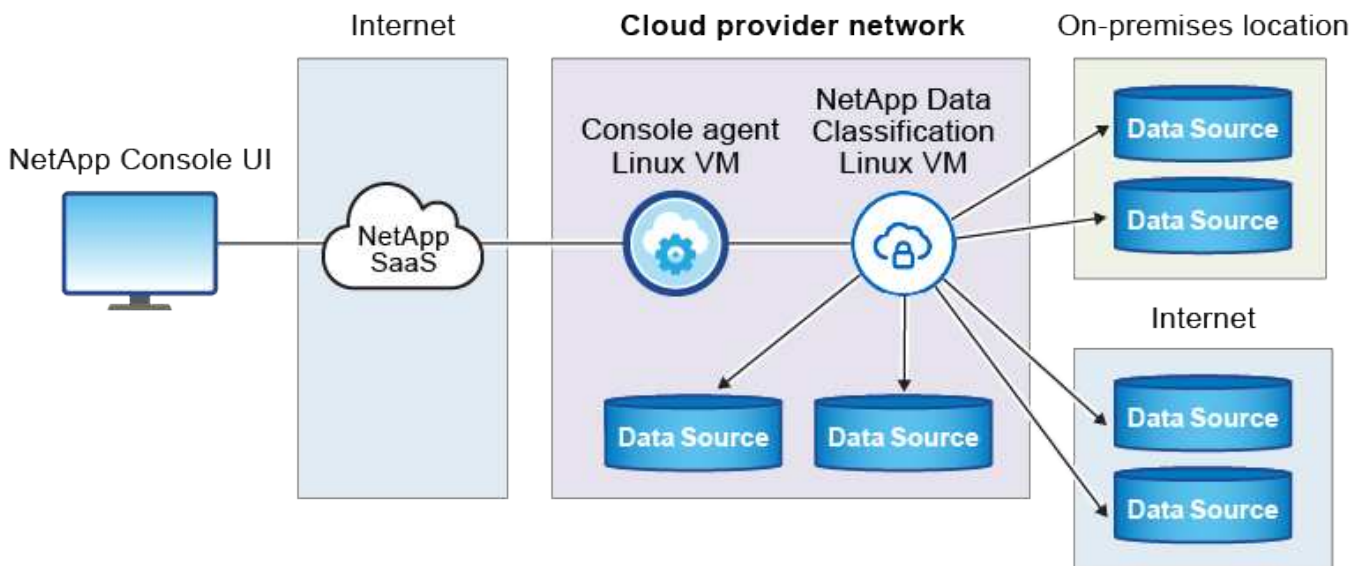
如果您希望使用同样位于本地的数据分类实例来扫描本地ONTAP系统，则本地安装是一个不错的选择。这不是必需的。无论选择哪种安装方法，软件的功能都是相同的。

数据分类安装脚本首先检查系统和环境是否满足所需的先决条件。如果所有先决条件都满足，则安装开始。如果您想独立于运行数据分类安装来验证先决条件，您可以下载一个单独的软件包，该软件包仅测试先决条件。[了解如何检查您的 Linux 主机是否已准备好安装数据分类](#)。

您所在场所的 Linux 主机上的典型安装具有以下组件和连接。



云端 Linux 主机上的典型安装具有以下组件和连接。



快速启动

按照以下步骤快速开始，或者向下滚动到其余部分以获取完整详细信息。

1

创建控制台代理

如果您还没有控制台代理，["在本地部署控制台代理"](#)在您的网络中的 Linux 主机上，或在云中的 Linux 主机上。

您还可以与您的云提供商一起创建控制台代理。看 ["在 AWS 中创建控制台代理"](#)，["在 Azure 中创建控制台代理"](#)，或者 ["在 GCP 中创建控制台代理"](#)。

2

审查先决条件

确保您的环境能够满足先决条件。这包括实例的出站互联网访问、控制台代理和数据分类之间通过端口 443 的连接等等。 [查看完整列表](#)。

您还需要一个满足以下条件的 Linux 系统[遵循要求](#)。

3

下载并部署数据分类

从NetApp支持站点下载云数据分类软件，并将安装程序文件复制到您计划使用的 Linux 主机。然后启动安装向导并按照提示部署数据分类实例。

创建控制台代理

您需要先安装控制台代理，然后才能安装和使用数据分类。在大多数情况下，您可能会在尝试激活数据分类之前设置控制台代理，因为大多数 ["控制台功能需要控制台代理"](#)，但有些情况下您需要立即设置一个。

要在您的云提供商环境中创建一个，请参阅 ["在 AWS 中创建控制台代理"](#)，["在 Azure 中创建控制台代理"](#)，或者 ["在 GCP 中创建控制台代理"](#)。

在某些情况下，您必须使用部署在特定云提供商中的控制台代理：

- 在 AWS 或Amazon FSx for ONTAP中的Cloud Volumes ONTAP中扫描数据时，您可以使用 AWS 中的控制台代理。
- 在 Azure 中的Cloud Volumes ONTAP或Azure NetApp Files中扫描数据时，您可以使用 Azure 中的控制台代理。

对于Azure NetApp Files，它必须部署在与您要扫描的卷相同的区域中。

- 在 GCP 中的Cloud Volumes ONTAP中扫描数据时，您可以使用 GCP 中的控制台代理。

可以使用任何这些云控制台代理来扫描本地ONTAP系统、NetApp文件共享和数据库帐户。

请注意，您还可以 ["在本地部署控制台代理"](#)在您的网络中的 Linux 主机上或云中的 Linux 主机上。一些计划在本地安装数据分类的用户可能还会选择在本地安装控制台代理。

安装数据分类时，您将需要控制台代理系统的 IP 地址或主机名。如果您在您的场所安装了控制台代理，您将获得此信息。如果控制台代理部署在云中，您可以从控制台中找到此信息：选择帮助图标，然后选择*支持*，然后选择控制台代理。

准备 Linux 主机系统

数据分类软件必须在满足特定操作系统要求、RAM 要求、软件要求等的主机上运行。Linux 主机可以在您的网络中，也可以在云中。

确保您可以保持数据分类运行。数据分类机器需要保持开启状态以持续扫描您的数据。

- 数据分类必须运行在专用主机上。主机不能与其他应用程序或第三方软件（例如防病毒软件）共享。
- 选择与您计划使用数据分类扫描的数据集相符的大小。

系统大小	CPU	RAM（必须禁用交换内存）	磁盘
超大	32 个 CPU	128 GB 内存	• / 上 1 TiB SSD，或 /opt 上 100 GiB 可用 • /var/lib/docker 上可用 895 GiB • /tmp 上 5 GiB • 对于 Podman，/var/tmp 上有 30 GB
大的	16 个 CPU	64 GB 内存	• / 上 500 GiB SSD，或 /opt 上 100 GiB 可用 • /var/lib/docker 或 Podman /var/lib/containers 上可用 400 GiB • /tmp 上 5 GiB • 对于 Podman，/var/tmp 上有 30 GB

- 在云中为数据分类安装部署计算实例时，建议您使用满足上述“大型”系统要求的系统：
 - **Amazon Elastic Compute Cloud (Amazon EC2)** 实例类型：“m6i.4xlarge”。["查看其他 AWS 实例类型"](#)。
 - **Azure VM** 大小：“Standard_D16s_v3”。["查看其他 Azure 实例类型"](#)。
 - **GCP** 机器类型：“n2-standard-16”。["查看其他 GCP 实例类型"](#)。
- **UNIX** 文件夹权限：需要以下最低 UNIX 权限：

文件夹	最低权限
/tmp	rwxrwxrwt
/选择	rwxr-xr-x
/var/lib/docker	rwx-----
/usr/lib/systemd/系统	rwxr-xr-x

- 操作系统：
 - 以下操作系统需要使用 Docker 容器引擎：
 - Red Hat Enterprise Linux 版本 7.8 和 7.9
 - Ubuntu 22.04（需要数据分类版本 1.23 或更高版本）
 - Ubuntu 24.04（需要数据分类版本 1.23 或更高版本）
 - 以下操作系统需要使用 Podman 容器引擎，并且需要数据分类版本 1.30 或更高版本：
 - Red Hat Enterprise Linux 版本 8.8、8.10、9.0、9.1、9.2、9.3、9.4、9.5 和 9.6。
 - 必须在主机系统上启用高级矢量扩展 (AVX2)。

- **Red Hat 订阅管理**：主机必须在 Red Hat 订阅管理中注册。如果未注册，系统将无法访问存储库来在安装期间更新所需的第三方软件。
- **附加软件**：安装数据分类之前，必须在主机上安装以下软件：
 - 根据您使用的操作系统，您需要安装其中一个容器引擎：
 - Docker Engine 版本 19.3.1 或更高版本。 ["查看安装说明"](#)。
 - Podman 版本 4 或更高版本。要安装 Podman，请输入(sudo yum install podman netavark -y)。
- **Python 版本 3.6 或更高版本**。 ["查看安装说明"](#)。
 - **NTP 注意事项**：NetApp建议配置数据分类系统以使用网络时间协议 (NTP) 服务。数据分类系统和控制台代理系统之间的时间必须同步。
- **Firewalld 注意事项**：如果您计划使用 firewalld，我们建议您在安装数据分类之前启用它。运行以下命令进行配置 `firewalld` 以便与数据分类兼容：

```
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
```

如果您计划使用其他数据分类主机作为扫描器节点，请在此时将规则添加到您的主系统：

```
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
```

请注意，每次启用或更新时都必须重新启动 Docker 或 Podman `firewalld` 设置。



安装后，数据分类主机系统的 IP 地址无法更改。

启用数据分类的出站互联网访问

数据分类需要出站互联网访问。如果您的虚拟或物理网络使用代理服务器进行互联网访问，请确保数据分类实例具有出站互联网访问权限以联系以下端点。

端点	目的
\ https://api.console.netapp.com	与控制台的通信，其中包括NetApp帐户。
\ https://netapp-cloud-account.auth0.com \ https://auth0.com	与控制台网站通信，实现集中用户身份验证。

端点	目的
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/	提供对软件映像、清单、模板的访问以及发送日志和指标。
https://support.compliance.api.blueexp.netapp.com/	使NetApp能够从审计记录中流式传输数据。
https://github.com/docker https://download.docker.com	提供docker安装的必备包。
http://packages.ubuntu.com/ \ http://archive.ubuntu.com	提供 Ubuntu 安装的必备软件包。

验证所有必需的端口均已启用

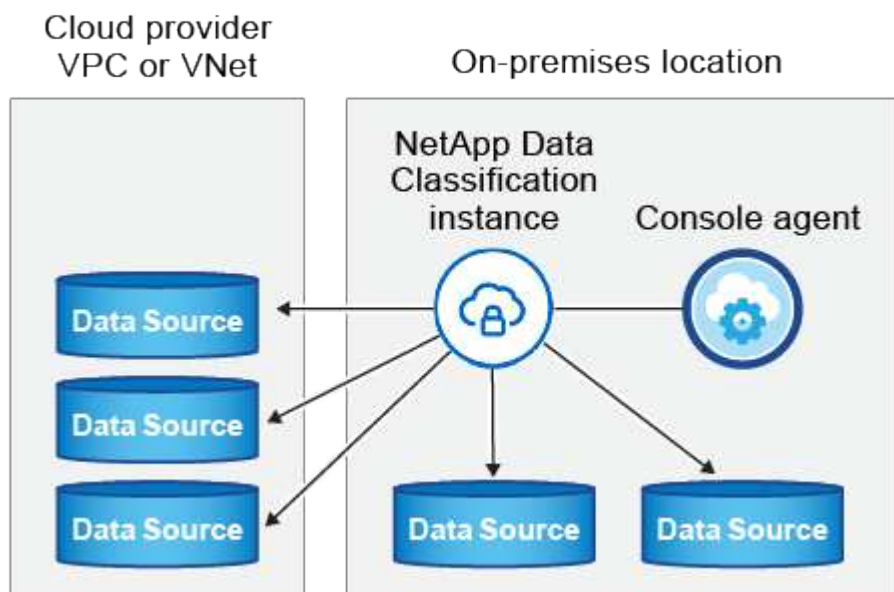
您必须确保所有必需的端口都已打开，以便控制台代理、数据分类、Active Directory 和数据源之间进行通信。

连接类型	端口	描述
控制台代理<>数据分类	8080 (TCP)、443 (TCP) 和 80。9000	控制台代理的防火墙或路由规则必须允许通过端口 443 进出数据分类实例的入站和出站流量。确保端口 8080 已打开，以便您可以在控制台中看到安装进度。如果 Linux 主机上使用防火墙，则 Ubuntu 服务器内的内部进程需要端口 9000。
控制台代理<> ONTAP 集群 (NAS)	443 (TCP)	控制台使用 HTTPS 发现ONTAP集群。如果您使用自定义防火墙策略，则它们必须满足以下要求： • 控制台代理主机必须允许通过端口 443 进行出站 HTTPS 访问。如果控制台代理位于云中，则预定义的防火墙或路由规则允许所有出站通信。 • ONTAP集群必须允许通过端口 443 进行入站 HTTPS 访问。默认的“mgmt”防火墙策略允许来自所有 IP 地址的入站 HTTPS 访问。如果您修改了此默认策略，或者创建了自己的防火墙策略，则必须将 HTTPS 协议与该策略关联并启用从控制台代理主机的访问。

连接类型	端口	描述
数据分类 <> ONTAP集群	- 对于 NFS - 111 (TCP\UDP) 和 2049 (TCP\UDP) - 对于 CIFS - 139 (TCP\UDP) 和 445 (TCP\UDP)	数据分类需要与每个Cloud Volumes ONTAP子网或本地ONTAP系统建立网络连接。Cloud Volumes ONTAP的防火墙或路由规则必须允许来自数据分类实例的入站连接。 确保这些端口对数据分类实例开放： - 对于 NFS - 111 和 2049 - 对于 CIFS - 139 和 445 NFS 卷导出策略必须允许从数据分类实例进行访问。
数据分类<> Active Directory	389 (TCP 和 UDP)、636 (TCP)、3268 (TCP) 和 3269 (TCP)	您必须已经为公司用户设置了 Active Directory。此外，数据分类需要 Active Directory 凭据来扫描 CIFS 卷。 您必须具有 Active Directory 的信息： - DNS 服务器 IP 地址，或多个 IP 地址 - 服务器的用户名和密码 - 域名 (Active Directory 名称) - 您是否使用安全 LDAP (LDAPS) - LDAP 服务器端口 (LDAP 通常为 389，安全 LDAP 通常为 636)

在 Linux 主机上安装数据分类

对于典型配置，您将在单个主机系统上安装该软件。[请参阅此处的步骤](#)。



看[准备 Linux 主机系统](#)和[审查先决条件](#)了解部署数据分类之前的完整要求列表。

只要实例具有互联网连接，数据分类软件的升级就会自动进行。



当软件安装在本地时，数据分类当前无法扫描 S3 存储桶、Azure NetApp Files 或 FSx for ONTAP。在这些情况下，您需要在云中部署单独的控制台代理和数据分类实例，并且 ["在连接器之间切换"](#)适用于不同的数据源。

典型配置的单主机安装

在单个本地主机上安装数据分类软件时，请查看要求并遵循以下步骤。

["观看此视频"](#)了解如何安装数据分类。

请注意，安装数据分类时会记录所有安装活动。如果您在安装过程中遇到任何问题，您可以查看安装审计日志的内容。它被写给 `/opt/netapp/install_logs/`。

开始之前

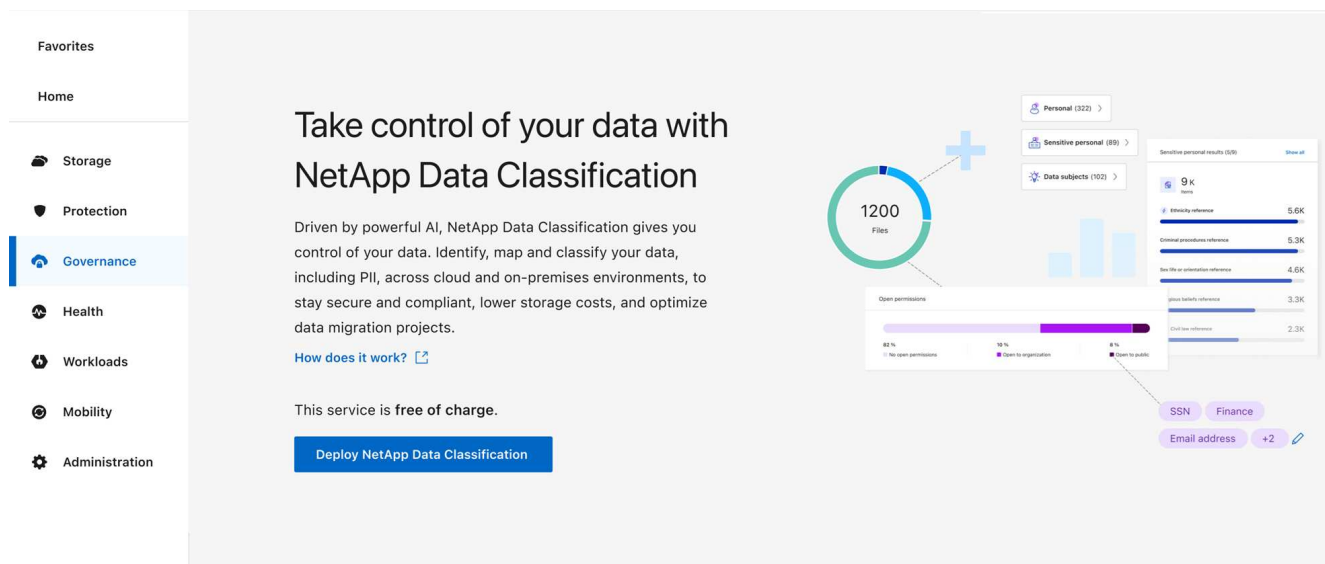
- 验证您的 Linux 系统是否满足[主机要求](#)。
- 验证系统是否安装了两个必备软件包（Docker Engine 或 Podman 和 Python 3）。
- 确保您在 Linux 系统上拥有 root 权限。
- 如果您使用代理访问互联网：
 - 您将需要代理服务器信息（IP 地址或主机名、连接端口、连接方案：https 或 http、用户名和密码）。
 - 如果代理正在执行 TLS 拦截，您需要知道数据分类 Linux 系统上存储 TLS CA 证书的路径。
 - 代理必须是非透明的。数据分类目前不支持透明代理。
 - 该用户必须是本地用户。不支持域用户。
- 验证您的离线环境是否满足要求[权限和连接性](#)。

步骤

1. 从下载数据分类软件 ["NetApp 支持站点"](#)。您应该选择的文件名为 **DATASENSE-INSTALLER-`<version>`.tar.gz**。
2. 将安装程序文件复制到您计划使用的 Linux 主机（使用 ``scp`` 或其他方法）。
3. 在主机上解压安装程序文件，例如：

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

4. 在控制台中，选择*治理>分类*。
5. 选择*在本地或云中部署分类*。



6. 根据您是在云中准备的实例上还是在本地准备的实例上安装数据分类，选择适当的*部署*选项来启动数据分类安装。
7. 将显示“在本地部署数据分类”对话框。复制提供的命令（例如： `sudo ./install.sh -a 12345 -c 27AG75 -t 2198qq` ）并将其粘贴到文本文件中，以便稍后使用。然后选择*关闭*以关闭对话框。
8. 在主机上，输入您复制的命令，然后按照一系列提示进行操作，或者您可以提供包含所有必需参数的完整命令作为命令行参数。

请注意，安装程序会执行预检查以确保您的系统和网络要求满足，以便成功安装。 ["观看此视频"](#)了解预检信息和含义。

根据提示输入参数：	输入完整命令：
a. 粘贴从步骤 7 复制的命令： <pre>sudo ./install.sh -a <account_id> -c <client_id> -t <user_token></pre> 如果您在云实例上安装（而不是在您的本地），请添加 <code>--manual-cloud-install <cloud_provider></code>。 b. 输入数据分类主机的 IP 地址或主机名，以便控制台代理系统可以访问它。 c. 输入控制台代理主机的 IP 地址或主机名，以便数据分类系统可以访问它。 d. 根据提示输入代理详细信息。如果您的控制台代理已经使用代理，则无需在此处再次输入此信息，因为数据分类将自动使用控制台代理所使用的代理。	或者，您可以提前创建整个命令，提供必要的主机和代理参数： <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>

变量值：

- `account_id` = NetApp帐户 ID

- `client_id` = 控制台代理客户端 ID（如果客户端 ID 中没有后缀“clients”，则添加后缀）
- `user_token` = JWT 用户访问令牌
- `ds_host` = 数据分类 Linux 系统的 IP 地址或主机名。
- `cm_host` = 控制台代理系统的 IP 地址或主机名。
- `cloud_provider` = 在云实例上安装时，根据云提供商输入“AWS”、“Azure”或“Gcp”。
- `proxy_host` = 如果主机位于代理服务器后面，则为代理服务器的 IP 或主机名。
- `proxy_port` = 连接到代理服务器的端口（默认为 80）。
- `proxy_scheme` = 连接方案：https 或 http（默认 http）。
- `proxy_user` = 如果需要基本身份验证，则经过身份验证的用户连接到代理服务器。用户必须是本地用户 - 不支持域用户。
- `proxy_password` = 您指定的用户名的密码。
- `ca_cert_dir` = 数据分类 Linux 系统上包含附加 TLS CA 证书包的路径。仅当代理执行 TLS 拦截时才需要。

结果

数据分类安装程序安装包、注册安装并安装数据分类。安装可能需要 10 到 20 分钟。

如果主机和控制台代理实例之间通过端口 8080 建立连接，您将在控制台的“数据分类”选项卡中看到安装进度。

下一步

您可以从配置页面选择要扫描的数据源。

在没有互联网访问的 Linux 主机上安装NetApp Data Classification

在没有互联网访问权限的本地站点的 Linux 主机上安装NetApp Data Classification称为_私有模式_。这种安装类型使用安装脚本，与NetApp Console SaaS 层没有连接。



BlueXP私有模式（传统BlueXP接口）通常用于没有互联网连接的本地环境和安全云区域，其中包括 AWS Secret Cloud、AWS Top Secret Cloud 和 Azure IL6。NetApp继续通过传统的BlueXP界面支持这些环境。有关旧版BlueXP界面中的私有模式文档，请参阅["BlueXP私人模式的 PDF 文档"](#)。

检查您的 Linux 主机是否已准备好安装NetApp Data Classification

在 Linux 主机上手动安装NetApp Data Classification之前，可以选择在主机上运行脚本以验证安装数据分类的所有先决条件是否都已具备。您可以在网络中的 Linux 主机上或云中的 Linux 主机上运行此脚本。主机可以连接到互联网，或者主机可以驻留在没有互联网访问的站点（暗站）。

数据分类安装脚本包含一个测试脚本，以确保您的环境满足要求。您可以单独运行此脚本来验证 Linux 主机是否

已准备就绪，然后再运行安装脚本。

入门指南

您将执行以下任务。

- 如果您尚未安装控制台代理，则可以选择安装一个。您可以在未安装控制台代理的情况下运行测试脚本，但脚本会检查控制台代理和数据分类主机之间的连接 - 因此建议您安装控制台代理。
- 准备主机并验证其是否满足所有要求。
- 启用数据分类主机的出站互联网访问。
- 验证所有系统上的所有必需端口是否都已启用。
- 下载并运行先决条件测试脚本。

创建控制台代理

您需要先安装控制台代理，然后才能安装和使用数据分类。但是，您可以在没有控制台代理的情况下运行先决条件脚本。

您可以 ["在本地安装控制台代理"](#) 在您网络中的 Linux 主机上，或者在云端的 Linux 主机上。如果控制台代理安装在本地，您也可以在本地上安装数据分类。

要在您的云提供商环境中创建控制台代理，请参阅：

- ["在 AWS 中创建控制台代理"](#)
- ["在 Azure 中创建控制台代理"](#)
- ["在 GCP 中创建控制台代理"](#)

运行先决条件脚本时，需要控制台代理系统的 IP 地址或主机名。如果您在本地安装了控制台代理，则可以获取此信息。如果控制台代理部署在云端，您可以从控制台中找到此信息：选择帮助图标，然后选择“支持”；在“代理和审核”部分，选择“转到代理”。

验证主机要求

数据分类软件必须运行在满足特定操作系统要求、内存要求和软件要求的主机上。

- 数据分类必须运行在专用主机上。主机不能与其他应用程序或第三方软件（例如防病毒软件）共享。
- 选择与您计划使用数据分类扫描的数据集相符的大小。

系统大小	CPU	RAM（必须禁用交换内存）	磁盘
超大	32 个 CPU	128 GB 内存	• / 上 1 TiB SSD，或 /opt 上 100 GiB 可用 • /var/lib/docker 上可用 895 GiB • /tmp 上 5 GiB • 对于 Podman，/var/tmp 上有 30 GB

系统大小	CPU	RAM（必须禁用交换内存）	磁盘
大的	16 个 CPU	64 GB 内存	• / 上 500 GiB SSD，或 /opt 上 100 GiB 可用 • /var/lib/docker 或 Podman /var/lib/containers 上可用 400 GiB • /tmp 上 5 GiB • 对于 Podman，/var/tmp 上有 30 GB

- 在云中为数据分类安装部署计算实例时，建议您使用满足上述“大型”系统要求的系统：
 - **Amazon Elastic Compute Cloud (Amazon EC2)** 实例类型：“m6i.4xlarge”。["查看其他 AWS 实例类型"](#)。
 - **Azure VM** 大小：“Standard_D16s_v3”。["查看其他 Azure 实例类型"](#)。
 - **GCP** 机器类型：“n2-standard-16”。["查看其他 GCP 实例类型"](#)。

- **UNIX** 文件夹权限：需要以下最低 UNIX 权限：

文件夹	最低权限
/tmp	rwxrwxrwt
/选择	rwxr-xr-x
/var/lib/docker	rwx-----
/usr/lib/systemd/系统	rwxr-xr-x

- 操作系统：
 - 以下操作系统需要使用 Docker 容器引擎：
 - Red Hat Enterprise Linux 版本 7.8 和 7.9
 - Ubuntu 22.04（需要数据分类版本 1.23 或更高版本）
 - Ubuntu 24.04（需要数据分类版本 1.23 或更高版本）
 - 以下操作系统需要使用 Podman 容器引擎，并且需要数据分类版本 1.30 或更高版本：
 - Red Hat Enterprise Linux 版本 8.8、8.10、9.0、9.1、9.2、9.3、9.4、9.5 和 9.6。
 - 必须在主机系统上启用高级矢量扩展 (AVX2)。
- **Red Hat** 订阅管理：主机必须在 Red Hat 订阅管理中注册。如果未注册，系统将无法访问存储库来在安装期间更新所需的第三方软件。
- 附加软件：安装数据分类之前，必须在主机上安装以下软件：
 - 根据您使用的操作系统，您需要安装其中一个容器引擎：
 - Docker Engine 版本 19.3.1 或更高版本。["查看安装说明"](#)。
 - Podman 版本 4 或更高版本。要安装 Podman，请输入(sudo yum install podman netavark -y)。

- Python 版本 3.6 或更高版本。"查看安装说明"。
 - **NTP** 注意事项：NetApp建议配置数据分类系统以使用网络时间协议 (NTP) 服务。数据分类系统和控制台代理系统之间的时间必须同步。
- **Firewalld** 注意事项：如果您计划使用 firewalld，我们建议您在安装数据分类之前启用它。运行以下命令进行配置 `firewalld` 以便与数据分类兼容：

```
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
```

如果您计划使用其他数据分类主机作为扫描器节点（在分布式模型中），请在此时将规则添加到您的主系统：

```
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
```

请注意，每次启用或更新时都必须重新启动 Docker 或 Podman `firewalld` 设置。

启用数据分类的出站互联网访问

数据分类需要出站互联网访问。如果您的虚拟或物理网络使用代理服务器进行互联网访问，请确保数据分类实例具有出站互联网访问权限以联系以下端点。



对于安装在没有互联网连接的站点的主机系统，不需要本节。

端点	目的
\ https://api.console.netapp.com	与控制台服务（包括NetApp帐户）的通信。
\ https://netapp-cloud-account.auth0.com \ https://auth0.com	与控制台网站通信，实现集中用户身份验证。
\ https://support.compliance.api.console.netapp.com/ \ https://hub.docker.com/ \ https://auth.docker.io/ \ https://registry-1.docker.io/ \ https://index.docker.io/ \ https://dseasb33srrrn.cloudfront.net/ \ https://production.cloudflare.docker.com/	提供对软件映像、清单、模板的访问以及发送日志和指标。

端点	目的
\ https://support.compliance.api.console.netapp.com/	使NetApp能够从审计记录中流式传输数据。
https://github.com/docker https://download.docker.com	提供docker安装的必备包。
\ http://packages.ubuntu.com/ \ http://archive.ubuntu.com	提供 Ubuntu 安装的必备软件包。

验证所有必需的端口均已启用

您必须确保所有必需的端口都已打开，以便控制台代理、数据分类、Active Directory 和数据源之间进行通信。

连接类型	端口	描述
控制台代理<>数据分类	8080 (TCP)、443 (TCP) 和 80。9000	控制台代理的防火墙或路由规则必须允许通过端口 443 进出数据分类实例的入站和出站流量。确保端口 8080 已打开，以便您可以在控制台中看到安装进度。如果 Linux 主机上使用防火墙，则 Ubuntu 服务器内的内部进程需要端口 9000。
控制台代理<> ONTAP 集群 (NAS)	443 (TCP)	控制台使用 HTTPS 发现ONTAP集群。如果您使用自定义防火墙策略，控制台代理主机必须允许通过端口 443 进行出站 HTTPS 访问。如果控制台代理位于云中，则预定义的防火墙或路由规则允许所有出站通信。

运行数据分类先决条件脚本

按照以下步骤运行数据分类先决条件脚本。

"[观看此视频](#)"了解如何运行先决条件脚本并解释结果。

开始之前

- 验证您的 Linux 系统是否满足[主机要求](#)。
- 验证系统是否安装了两个必备软件包（Docker Engine 或 Podman 和 Python 3）。
- 确保您在 Linux 系统上拥有 root 权限。

步骤

1. 从下载数据分类先决条件脚本 "[NetApp 支持站点](#)"。您应该选择的文件名为 **standalone-pre-requisite-tester-<version>**。
2. 将文件复制到您计划使用的 Linux 主机（使用 `scp` 或其他方法）。
3. 分配运行脚本的权限。

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

4. 使用以下命令运行脚本。

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

仅当您在没有互联网访问的主机上运行脚本时才添加选项“--darksite”。当主机未连接到互联网时，某些先决条件测试将被跳过。

5. 该脚本会提示您输入数据分类主机的 IP 地址。

- 输入 IP 地址或主机名。

6. 该脚本会提示您是否安装了控制台代理。

- 如果您没有安装控制台代理，请输入 **N**。
- 如果您确实安装了控制台代理，请输入 **Y**。然后输入控制台代理的 IP 地址或主机名，以便测试脚本可以测试此连接。

7. 该脚本在系统上运行各种测试，并在运行过程中显示结果。完成后，它会将会话日志写入名为 `prerequisites-test-<timestamp>.log` 在目录中 `/opt/netapp/install_logs`。

结果

如果所有先决条件测试均成功运行，您可以在准备就绪后在主机上安装数据分类。

如果发现任何问题，则将其归类为“建议”或“需要”修复。推荐的问题通常是会使数据分类扫描和分类任务运行速度变慢的项目。这些项目不需要更正 - 但您可能需要解决它们。

如果您有任何“必需”问题，您应该修复这些问题并再次运行先决条件测试脚本。

版权信息

版权所有 © 2026 NetApp, Inc.。保留所有权利。中国印刷。未经版权所有者事先书面许可，本文档中受版权保护的任何部分不得以任何形式或通过任何手段（图片、电子或机械方式，包括影印、录音、录像或存储在电子检索系统中）进行复制。

从受版权保护的 NetApp 资料派生的软件受以下许可和免责声明的约束：

本软件由 NetApp 按“原样”提供，不含任何明示或暗示担保，包括但不限于适销性以及针对特定用途的适用性的隐含担保，特此声明不承担任何责任。在任何情况下，对于因使用本软件而以任何方式造成的任何直接性、间接性、偶然性、特殊性、惩罚性或后果性损失（包括但不限于购买替代商品或服务；使用、数据或利润方面的损失；或者业务中断），无论原因如何以及基于何种责任理论，无论出于合同、严格责任或侵权行为（包括疏忽或其他行为），NetApp 均不承担责任，即使已被告知存在上述损失的可能性。

NetApp 保留在不另行通知的情况下随时对本文档所述的任何产品进行更改的权利。除非 NetApp 以书面形式明确同意，否则 NetApp 不承担因使用本文档所述产品而产生的任何责任或义务。使用或购买本产品不表示获得 NetApp 的任何专利权、商标权或任何其他知识产权许可。

本手册中描述的产品可能受一项或多项美国专利、外国专利或正在申请的专利的保护。

有限权利说明：政府使用、复制或公开本文档受 DFARS 252.227-7013（2014 年 2 月）和 FAR 52.227-19（2007 年 12 月）中“技术数据权利 — 非商用”条款第 (b)(3) 条规定的限制条件的约束。

本文档中所含数据与商业产品和/或商业服务（定义见 FAR 2.101）相关，属于 NetApp, Inc. 的专有信息。根据本协议提供的所有 NetApp 技术数据和计算机软件具有商业性质，并完全由私人出资开发。美国政府对这些数据的使用权具有非排他性、全球性、受限且不可撤销的许可，该许可既不可转让，也不可再许可，但仅限在与交付数据所依据的美国政府合同有关且受合同支持的情况下使用。除本文档规定的情形外，未经 NetApp, Inc. 事先书面批准，不得使用、披露、复制、修改、操作或显示这些数据。美国政府对国防部的授权仅限于 DFARS 的第 252.227-7015(b)（2014 年 2 月）条款中明确的权利。

商标信息

NetApp、NetApp 标识和 <http://www.netapp.com/TM> 上所列的商标是 NetApp, Inc. 的商标。其他公司和产品名称可能是其各自所有者的商标。