



## 部署 **Cloud Data sense**

### Cloud Data Sense

NetApp  
July 19, 2022

# 目录

- 部署 Cloud Data sense ..... 1
  - 在云中部署 Cloud Data sense ..... 1
  - 在可访问 Internet 的 Linux 主机上部署 Cloud Data sense ..... 5
  - 在内部部署 Cloud Data sense ，而无需访问 Internet ..... 12

# 部署 Cloud Data sense

## 在云中部署 Cloud Data sense

完成几个步骤，在云中部署 Cloud Data sense。

请注意，您也可以 ["在可访问 Internet 的 Linux 主机上部署 Data sense"](#)。如果您希望使用也位于内部的数据感知实例扫描内部 ONTAP 系统，则安装类型可能是一个不错的选择，但这并不是一项要求。无论您选择哪种安装方法，软件的工作方式都完全相同。

### 快速入门

按照以下步骤快速入门，或者向下滚动到其余部分以了解完整详细信息。

如果您还没有 Connector，请立即创建一个 Connector。请参见 ["在 AWS 中创建连接器"](#)，["在 Azure 中创建连接器"](#)或 ["在 GCP 中创建连接器"](#)。

您也可以 ["在内部部署 Connector"](#) 在网络或云中的 Linux 主机上。

确保您的环境可以满足前提条件。其中包括实例的出站 Internet 访问，通过端口 443 在 Connector 和 Cloud Data sense 之间建立连接等。 [请参见完整列表](#)。

默认配置要求云数据感知实例使用 16 个 vCPU。请参见 ["有关实例类型的更多信息"](#)。

启动安装向导以在云中部署 Cloud Data sense 实例。

Cloud Manager 中 Cloud Data 感知扫描的前 1 TB 数据是免费的。要在这之后继续扫描数据，需要通过云提供商 Marketplace 订阅 Cloud Manager 或从 NetApp 获得 BYOL 许可证。

### 创建连接器

如果您还没有 Connector，请在云提供商中创建一个 Connector。请参见 ["在 AWS 中创建连接器"](#) 或 ["在 Azure 中创建连接器"](#)或 ["在 GCP 中创建连接器"](#)。在大多数情况下，您可能在尝试激活 Cloud Data sense 之前设置 Connector，因为大多数情况下都是这样 ["Cloud Manager 功能需要使用 Connector"](#)但在某些情况下，您需要立即设置一个。

在某些情况下，您必须使用部署在特定云提供商中的 Connector：

- 在 AWS 中的 Cloud Volumes ONTAP，适用于 ONTAP 的 Amazon FSx 或 AWS S3 存储分段中扫描数据时，您可以使用 AWS 中的连接器。
- 在 Azure 或 Azure NetApp Files 中的 Cloud Volumes ONTAP 中扫描数据时，您可以使用 Azure 中的连接器。
  - 对于 Azure NetApp Files，必须将其部署在与要扫描的卷相同的区域中。
- 在 GCP 的 Cloud Volumes ONTAP 中扫描数据时，您可以在 GCP 中使用连接器。

使用上述任一 Cloud Connector 时，可以扫描内部 ONTAP 系统、非 NetApp 文件共享、通用 S3 对象存储、数据库、OneDrive 文件夹、SharePoint 帐户和 Google Drive 帐户。

请注意，您也可以 ["在内部部署 Connector"](#) 在网络或云中的 Linux 主机上。某些计划在内部安装 Data sense 的

用户也可以选择在内网安装 Connector。

如您所见，在某些情况下可能需要使用 "多个连接器"。

## 政府区域支持

如果在政府区域(AWS GovCloud、Azure Gov或Azure DoD)中部署了Connector、则支持Cloud Data sense。以这种方式部署时、Data sense具有以下限制：

- 无法扫描OneDrive帐户、SharePoint帐户和Google Drive帐户。
- 无法集成Microsoft Azure信息保护(AIP)标签功能。

## 查看前提条件

在云中部署 Cloud Data sense 之前，请查看以下前提条件以确保您的配置受支持。

### 从 Cloud Data sense 启用出站 Internet 访问

云数据感知需要出站 Internet 访问。如果虚拟或物理网络使用代理服务器访问 Internet，请确保 Data sense 实例具有出站 Internet 访问权限，以便与以下端点联系。在云中部署 Data sense 时，它与 Connector 位于同一子网中。

根据您是在 AWS，Azure 还是 GCP 中部署 Cloud Data sense，查看下表。

- AWS 部署所需的端点： \*

端点	目的
<a href="https://cloudmanager.cloud.netapp.com">https://cloudmanager.cloud.netapp.com</a>	与 Cloud Manager 服务进行通信，其中包括 NetApp 帐户。
<a href="https://netapp-cloud-account.auth0.com">https://netapp-cloud-account.auth0.com</a> <a href="https://auth0.com">https://auth0.com</a>	与 NetApp Cloud Central 进行通信以实现集中式用户身份验证。
<a href="https://cloud-compliance-support-netapp.s3.us-west-2.amazonaws.com">https://cloud-compliance-support-netapp.s3.us-west-2.amazonaws.com</a> <a href="https://hub.docker.com">https://hub.docker.com</a> <a href="https://auth.docker.io">https://auth.docker.io</a> <a href="https://registry-1.docker.io">https://registry-1.docker.io</a> <a href="https://index.docker.io/">https://index.docker.io/</a> <a href="https://dseasb33srrn.cloudfront.net/">https://dseasb33srrn.cloudfront.net/</a> <a href="https://production.cloudflare.docker.com/">https://production.cloudflare.docker.com/</a>	提供对软件映像、清单和模板的访问。
<a href="https://kinesis.us-east-1.amazonaws.com">https://kinesis.us-east-1.amazonaws.com</a>	使 NetApp 能够从审计记录流化数据。
<a href="https://cognito-idp.us-east-1.amazonaws.com">https://cognito-idp.us-east-1.amazonaws.com</a> <a href="https://cognito-identity.us-east-1.amazonaws.com">https://cognito-identity.us-east-1.amazonaws.com</a> <a href="https://user-feedback-store-prod.s3.us-west-2.amazonaws.com">https://user-feedback-store-prod.s3.us-west-2.amazonaws.com</a> <a href="https://customer-data-production.s3.us-west-2.amazonaws.com">https://customer-data-production.s3.us-west-2.amazonaws.com</a>	支持 Cloud Data sense 访问和下载清单和模板，以及发送日志和指标。

- Azure 和 GCP 部署所需的端点： \*

端点	目的
<a href="https://cloudmanager.cloud.netapp.com">https://cloudmanager.cloud.netapp.com</a>	与 Cloud Manager 服务进行通信，其中包括 NetApp 帐户。

端点	目的
<a href="https://netapp-cloud-account.auth0.com">https://netapp-cloud-account.auth0.com</a> <a href="https://auth0.com">https://auth0.com</a>	与 NetApp Cloud Central 进行通信以实现集中式用户身份验证。
<a href="https://support.compliance.cloudmanager.cloud.netapp.com/">https://support.compliance.cloudmanager.cloud.netapp.com/</a> <a href="https://hub.docker.com">https://hub.docker.com</a> <a href="https://auth.docker.io">https://auth.docker.io</a> <a href="https://registry-1.docker.io">https://registry-1.docker.io</a> <a href="https://index.docker.io/">https://index.docker.io/</a> <a href="https://dseasb33snnrn.cloudfront.net/">https://dseasb33snnrn.cloudfront.net/</a> <a href="https://production.cloudflare.docker.com/">https://production.cloudflare.docker.com/</a>	可用于访问软件映像，清单，模板以及发送日志和指标。
<a href="https://support.compliance.cloudmanager.cloud.netapp.com/">https://support.compliance.cloudmanager.cloud.netapp.com/</a>	使 NetApp 能够从审计记录流化数据。

### 确保 **Cloud Manager** 具有所需权限

确保 Cloud Manager 有权为 Cloud Data sense 实例部署资源并创建安全组。您可以在中找到最新的 Cloud Manager 权限 "[NetApp 提供的策略](#)"。

### 检查 **vCPU** 限制

确保云提供商的 vCPU 限制允许部署包含 16 个核心的实例。您需要验证运行 Cloud Manager 的区域中相关实例系列的 vCPU 限制。"[请参见所需的实例类型](#)"。

有关 vCPU 限制的详细信息，请参见以下链接：

- "[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](#)"

请注意，您可以在 CPU 较少且 RAM 较少的系统上部署 Data sense，但使用这些系统时会有一些限制。请参见 "[使用较小的实例类型](#)" 了解详细信息。

### 确保 **Cloud Manager Connector** 可以访问 **Cloud Data sense**

确保 Connector 与 Cloud Data sense 实例之间的连接。Connector 的安全组必须允许通过端口 443 与 Data sense 实例之间的入站和出站流量。通过此连接，可以部署 Data sense 实例，并可在合规性和监管选项卡中查看信息。AWS 和 Azure 中的政府区域支持云数据感知。

AWS 和 AWS GovCloud 部署还需要其他入站和出站规则。请参见 "[AWS 中连接器的规则](#)" 了解详细信息。

Azure 和 Azure 政府部署还需要其他入站和出站规则。请参见 "[Azure 中连接器的规则](#)" 了解详细信息。

### 确保您可以保持 **Cloud Data sense** 正常运行

云数据感知实例需要保持运行状态才能持续扫描数据。

### 确保 **Web** 浏览器连接到 **Cloud Data sense**

启用 Cloud Data sense 后，请确保用户从连接到 Data sense 实例的主机访问 Cloud Manager 界面。

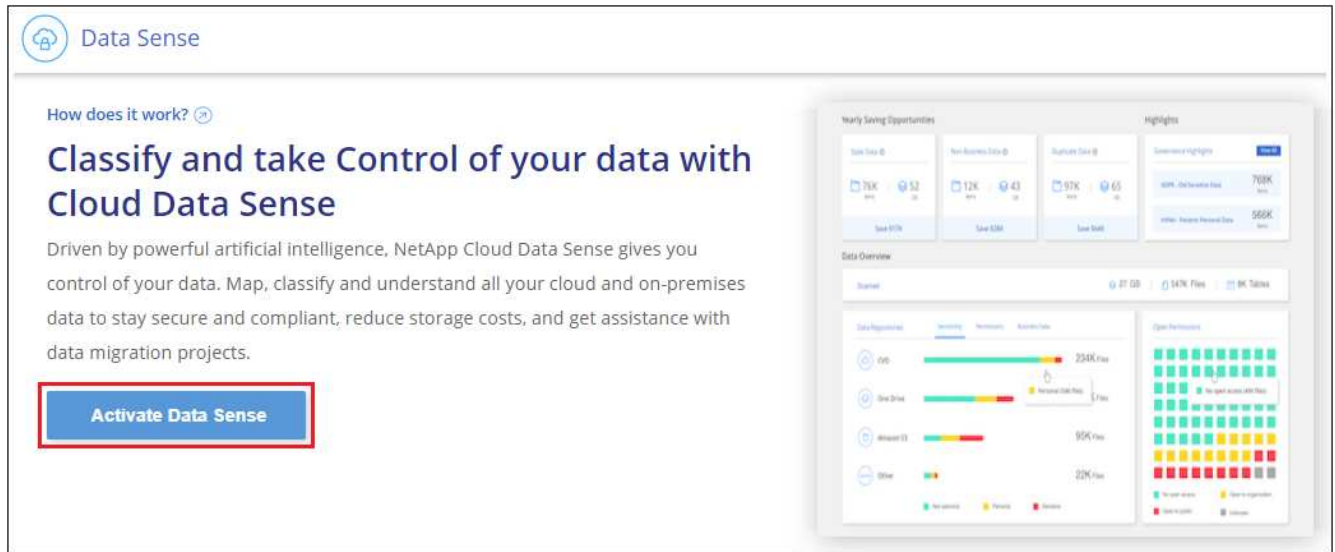
数据感知实例使用专用 IP 地址来确保索引数据无法通过 Internet 访问。因此，用于访问 Cloud Manager 的 Web 浏览器必须连接到该专用 IP 地址。此连接可以来自与云提供商的直接连接（例如 VPN），也可以来自与 Data sense 实例位于同一网络中的主机。

## 在云中部署 Data sense

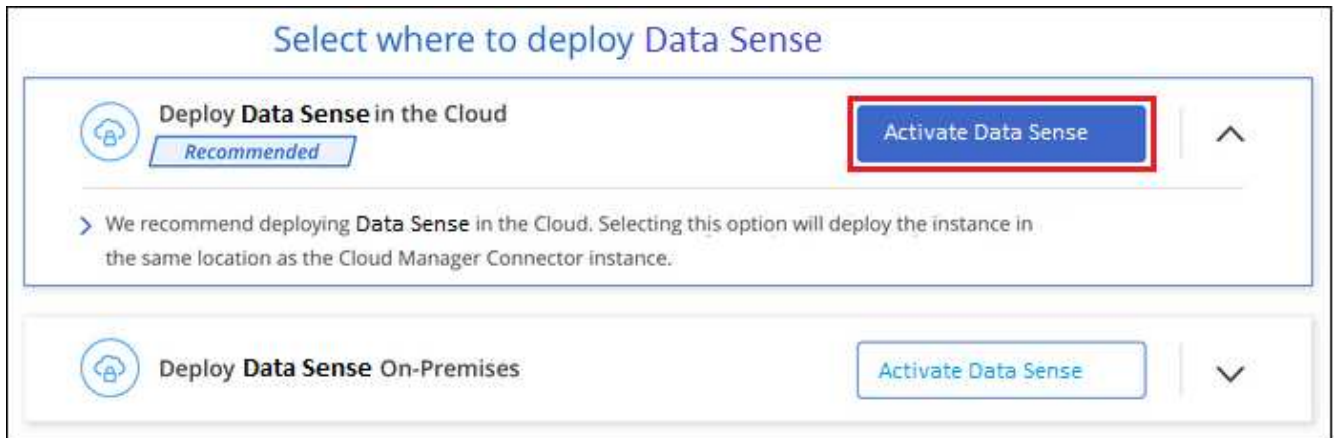
按照以下步骤在云中部署 Cloud Data sense 实例。

### 步骤

1. 从Cloud Manager左侧导航菜单中、单击\*数据感知\*。
2. 单击 \* 激活数据感知 \*。



3. 单击 \* 激活数据感知 \* 以启动云部署向导。



4. 向导将在完成部署步骤时显示进度。如果遇到任何问题，它将停止并请求输入。



5. 部署实例后，单击 \* 继续配置 \* 以转到 *Configuration* 页面。

Cloud Manager 在云提供商中部署 Cloud Data sense 实例。

在配置页面中，您可以选择要扫描的数据源。

您也可以 ["为 Cloud Data sense 设置许可"](#) 目前。在数据量超过 1 TB 之前，不会向您收取任何费用。

## 在可访问 Internet 的 Linux 主机上部署 Cloud Data sense

完成以下几个步骤，在您的网络或云中可访问 Internet 的 Linux 主机上部署 Cloud Data sense 。

如果您希望使用也位于内部的数据感知实例扫描内部 ONTAP 系统，则内部安装可能是一个不错的选择，但这并不是一项要求。无论您选择哪种安装方法，软件的工作方式都完全相同。

请注意，您也可以 ["在无法访问 Internet 的内部站点中部署 Data sense"](#) 适用于完全安全的站点。

### 快速入门

按照以下步骤快速入门，或者向下滚动到其余部分以了解完整详细信息。

如果您还没有 Connector，请立即创建一个 Connector。请参见 ["在 AWS 中创建连接器"](#)，["在 Azure 中创建连接器"](#)或 ["在 GCP 中创建连接器"](#)。

您也可以 ["在内部部署 Connector"](#) 在网络或云中的 Linux 主机上。

确保您的环境可以满足前提条件。其中包括实例的出站 Internet 访问，通过端口 443 在 Connector 和 Cloud Data sense 之间建立连接等。 [请参见完整列表](#)。

您还需要满足的 Linux 系统 [以下要求](#)。

从 NetApp 支持站点下载 Cloud Data sense 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 Data sense 实例。

Cloud Manager 中 Cloud Data 感知扫描的前 1 TB 数据是免费的。要在这之后继续扫描数据，需要订阅云提供商 Marketplace 或获得 NetApp 的 BYOL 许可证。

## 创建连接器

如果您还没有 Connector，请在云提供商中创建一个 Connector。请参见 ["在 AWS 中创建连接器"](#) 或 ["在 Azure 中创建连接器"](#) 或 ["在 GCP 中创建连接器"](#)。在大多数情况下，您可能在尝试激活 Cloud Data sense 之前设置 Connector，因为大多数情况下都是这样 ["Cloud Manager 功能需要使用 Connector"](#)但在某些情况下，您需要立即设置一个。

在某些情况下，您必须使用部署在特定云提供商中的 Connector：

- 在 AWS 中的 Cloud Volumes ONTAP，适用于 ONTAP 的 Amazon FSx 或 AWS S3 存储分段中扫描数据时，您可以使用 AWS 中的连接器。
- 在 Azure 或 Azure NetApp Files 中的 Cloud Volumes ONTAP 中扫描数据时，您可以使用 Azure 中的连接器。
  - 对于 Azure NetApp Files，必须将其部署在与要扫描的卷相同的区域中。
- 在 GCP 的 Cloud Volumes ONTAP 中扫描数据时，您可以在 GCP 中使用连接器。

内部 ONTAP 系统、非 NetApp 文件共享、通用 S3 对象存储、数据库、OneDrive 文件夹、SharePoint 帐户和 Google Drive 帐户均可使用这些 Cloud Connector 中的任何一种进行扫描。

请注意，您也可以 ["在内部部署 Connector"](#) 在网络或云中的 Linux 主机上。某些计划在内部安装 Data sense 的用户也可以选择在内网安装 Connector。

如您所见，在某些情况下可能需要使用 ["多个连接器"](#)。

## 准备 Linux 主机系统

数据感知软件必须在满足特定操作系统要求，RAM 要求，软件要求等要求的主机上运行。与其他应用程序共享的主机不支持数据感知 - 此主机必须是专用主机。

- 操作系统：Red Hat Enterprise Linux 或 CentOS 8.0、8.1 或 8.4 版
  - 可以使用 7.8 或 7.9 版，但 Linux 内核版本必须为 4.14 或更高版本
  - 操作系统必须能够安装 Docker 引擎（例如，根据需要禁用 *firewalld* 服务）
- 磁盘：SSD，500 GiB 可在 /，或上使用
  - /opt 上提供 100 GiB
  - /var 上提供 400 GiB
  - /tmp 上 5 GiB
- RAM：64 GB（必须在主机上禁用交换内存）
- CPU：16 个核心

请注意，您可以在 CPU 较少且 RAM 较少的系统上部署 Data sense，但使用这些系统时会有一些限制。请参见 ["使用较小的实例类型"](#) 了解详细信息。

- Red Hat Enterprise Linux 系统必须在 Red Hat 订购管理中注册。如果未注册，则系统无法在安装期间访问存储库以更新所需的第三方软件。
- 主机上必须安装以下软件。如果此主机上尚不存在此软件，则安装程序将为您安装此软件：
  - Docker 引擎版本 19 或更高版本。 ["查看安装说明"](#)。



- Python 3 3.6 或更高版本。 ["查看安装说明"](#)。

## 验证 **Cloud Manager** 和 **Data sense** 前提条件

在 Linux 系统上部署 Cloud Data sense 之前，请查看以下前提条件，以确保您的配置受支持。

### 从 **Cloud Data sense** 启用出站 **Internet** 访问

云数据感知需要出站 Internet 访问。如果虚拟或物理网络使用代理服务器访问 Internet，请确保 Data sense 实例具有出站 Internet 访问权限，以便与以下端点联系。

端点	目的
<a href="https://cloudmanager.cloud.netapp.com">https://cloudmanager.cloud.netapp.com</a>	与 Cloud Manager 服务进行通信，其中包括 NetApp 帐户。
<a href="https://netapp-cloud-account.auth0.com">https://netapp-cloud-account.auth0.com</a> <a href="https://auth0.com">https://auth0.com</a>	与 NetApp Cloud Central 进行通信以实现集中式用户身份验证。
<a href="https://support.compliance.cloudmanager.cloud.netapp.com/">https://support.compliance.cloudmanager.cloud.netapp.com/</a> <a href="https://hub.docker.com">https://hub.docker.com</a> <a href="https://auth.docker.io">https://auth.docker.io</a> <a href="https://registry-1.docker.io">https://registry-1.docker.io</a> <a href="https://index.docker.io/">https://index.docker.io/</a> <a href="https://dseasb33srrn.cloudfront.net/">https://dseasb33srrn.cloudfront.net/</a> <a href="https://production.cloudflare.docker.com/">https://production.cloudflare.docker.com/</a>	可用于访问软件映像，清单，模板以及发送日志和指标。
<a href="https://support.compliance.cloudmanager.cloud.netapp.com/">https://support.compliance.cloudmanager.cloud.netapp.com/</a>	使 NetApp 能够从审计记录流化数据。
<a href="https://github.com/docker">https://github.com/docker</a> <a href="https://download.docker.com">https://download.docker.com</a> <a href="http://mirror.centos.org">http://mirror.centos.org</a> <a href="http://mirrorlist.centos.org">http://mirrorlist.centos.org</a> <a href="http://mirror.centos.org/centos/7/extras/x86_64/Packages/container-selinux-2.107-3.el7.noarch.rpm">http://mirror.centos.org/centos/7/extras/x86_64/Packages/container-selinux-2.107-3.el7.noarch.rpm</a>	提供安装必备软件包。

### 确保 **Cloud Manager** 具有所需权限

确保 Cloud Manager 有权为 Cloud Data sense 实例部署资源并创建安全组。您可以在中找到最新的 Cloud Manager 权限 ["NetApp 提供的策略"](#)。

### 确保 **Cloud Manager Connector** 可以访问 **Cloud Data sense**

确保 Connector 与 Cloud Data sense 实例之间的连接。Connector 的安全组必须允许通过端口 443 与 Data sense 实例之间的入站和出站流量。

通过此连接，可以部署 Data sense 实例，并可在合规性和监管选项卡中查看信息。

确保端口 8080 已打开，以便您可以在 Cloud Manager 中查看安装进度。

### 确保您可以保持 **Cloud Data sense** 正常运行

云数据感知实例需要保持运行状态才能持续扫描数据。

### 确保 **Web** 浏览器连接到 **Cloud Data sense**

启用 Cloud Data sense 后，请确保用户从连接到 Data sense 实例的主机访问 Cloud Manager 界面。

数据感知实例使用专用 IP 地址来确保索引数据无法通过 Internet 访问。因此，用于访问 Cloud Manager 的 Web 浏览器必须连接到该专用 IP 地址。此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 Data sense 实例位于同一网络中的主机。

## 在内部部署 Data sense

对于典型配置，您将在一个主机系统上安装该软件。 [请在此处查看这些步骤](#)。

对于需要扫描数 PB 数据的大型配置，您可以使用多个主机来提供额外的处理能力。 [请在此处查看这些步骤](#)。

请参见 [准备 Linux 主机系统](#) 和 [查看前提条件](#) 了解部署 Cloud Data sense 之前的完整要求列表。

只要实例具有 Internet 连接，就会自动升级到 Data sense 软件。



如果软件安装在内部环境中，则 Cloud Data sense 当前无法扫描 S3 存储分段，Azure NetApp Files 或 FSX for ONTAP。在这种情况下，您需要在云和中部署单独的 Connector 和 Data sense 实例 ["在连接器之间切换"](#) 不同的数据源。

### 典型配置的单主机安装

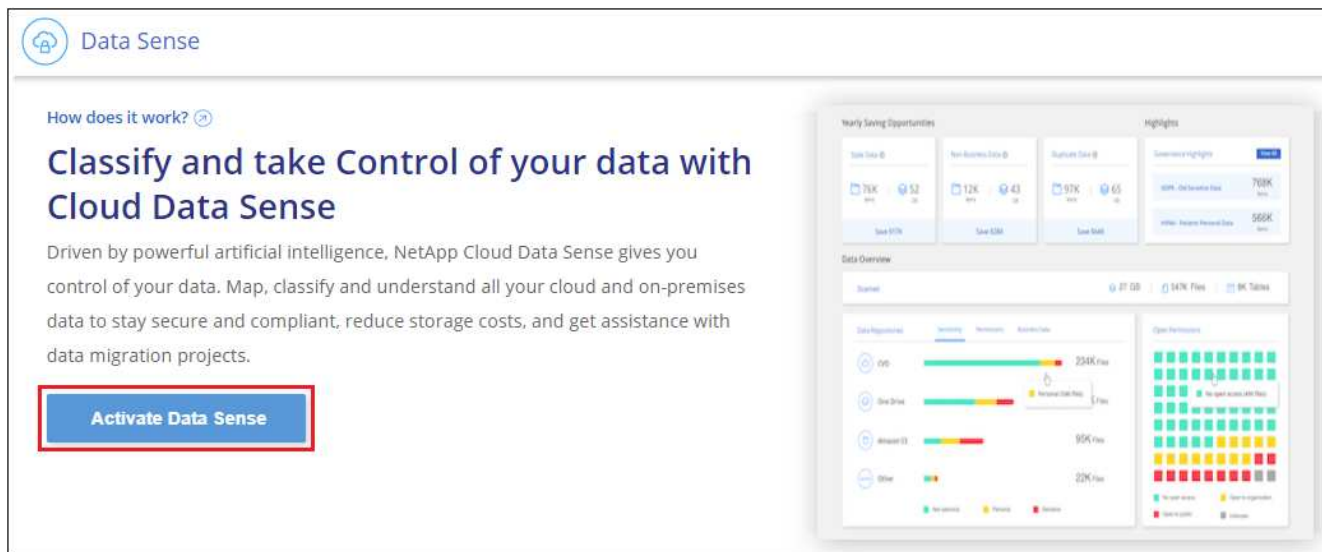
在单个内部主机上安装 Data sense 软件时，请按照以下步骤进行操作。

您需要什么？ **#8217** ；将需要什么

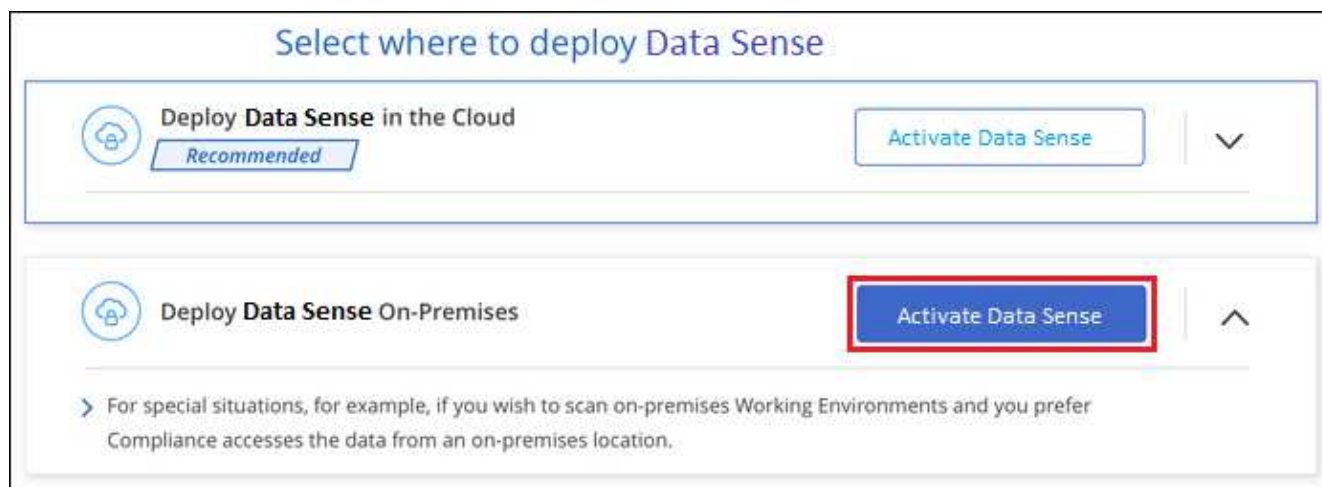
- 验证您的 Linux 系统是否满足 [主机要求](#)。
- （可选）验证系统是否已安装两个必备软件包（ Docker 引擎和 Python 3 ）。如果此软件尚未安装在系统上，安装程序将安装此软件。
- 确保您在 Linux 系统上具有 root 权限。
- 如果您使用的是代理，并且代理正在执行 TLS 截获，则需要了解 Data sense Linux 系统上用于存储 TLS CA 证书的路径。
- 验证脱机环境是否满足要求 [权限和连接](#)。

### 步骤

1. 从下载 Cloud Data sense 软件 ["NetApp 支持站点"](#)。您应选择的文件名为 \* cc\_onprem\_installer\_<版本>.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 Cloud Manager 中，单击 \* 数据感知 \*。
4. 单击 \* 激活数据感知 \*。



5. 单击 \* 激活数据感知 \* 以启动内部部署向导。



6. 在 *Deploy Data sense on premises* 对话框中，复制提供的命令并将其粘贴到文本文件中，以便稍后使用，然后单击 \* 关闭 \*。例如：

```
sudo ./install.sh -a 12345 -c 27ag75 -t 2198qq
```

7. 解压缩主机上的安装程序文件，例如：

```
tar -xzf cc_onprem_installer_1.13.1.tar.gz
```

8. 安装程序提示时，您可以在一系列提示中输入所需值，也可以将所需参数作为命令行参数提供给安装程序：

根据提示输入参数：	输入完整命令：
<p>a. 粘贴您从第 6 步复制的信息： <code>sUdo</code>  <code>./install.sh -a &lt;account_id&gt; -c</code>  <code>&lt;agent_id&gt; -t &lt;token&gt;</code></p> <p>b. 输入 Data sense 主机的 IP 地址或主机名，以便 Connector 实例可以访问它。</p> <p>c. 输入 Cloud Manager Connector 主机的 IP 地址或主机名，以便 Data sense 实例可以访问它。</p> <p>d. 根据提示输入代理详细信息。如果您的 Cloud Manager 已使用代理，则无需在此重新输入此信息，因为 Data sense 将自动使用 Cloud Manager 使用的代理。</p>	<p>或者，您也可以预先创建整个命令，并提供必要的主机和代理参数： <code>s udo ./install.sh -a &lt;account_id&gt; -c</code>  <code>&lt;agent_id&gt; -t &lt;token&gt; -host &lt;ds_host&gt; -manager</code>  <code>-host &lt;cm_host&gt; -proxy-host &lt;proxy_host&gt; -proxy</code>  <code>-port &lt;proxy_port&gt; -proxy-user-proxy_name&gt;</code>  <code>&lt;proxy_password&gt; -proxy_proxy_proxy_name&gt;</code>  <code>-proxy_proxy_proxy_proxy_name&gt;</code>  <code>-&lt;proxy_user&gt;</code></p>

变量值：

- `account_id` = NetApp 帐户 ID
- `agent_id` = 连接器 ID
- `token` = JWT 用户令牌
- `ds_host` = Data sense Linux 系统的 IP 地址或主机名。
- `cm_host` = Cloud Manager Connector 系统的 IP 地址或主机名。
- `proxy_host` = 代理服务器的 IP 或主机名（如果主机位于代理服务器之后）。
- `proxy_port` = 用于连接到代理服务器的端口（默认值为 80）。
- `proxy_scheme` = 连接方案：HTTPS 或 http（默认为 http）。
- `proxy_user` = 已通过身份验证的用户，用于连接到代理服务器（如果需要基本身份验证）。
- `proxy_password` = 指定用户名的密码。
- `ca_ct_dir` = 包含其他 TLS CA 证书包的 Data sense Linux 系统上的路径。仅当代理正在执行 TLS 截获时才需要。

Cloud Data sense 安装程序可安装软件包，安装 Docker，注册安装以及安装 Data sense。安装可能需要 10 到 20 分钟。

如果主机和 Connector 实例之间通过端口 8080 建立连接，则您将在 Cloud Manager 的 Data sense 选项卡中看到安装进度。

在配置页面中，您可以选择要扫描的数据源。

您也可以 ["为 Cloud Data sense 设置许可"](#) 目前。在数据量超过 1 TB 之前，不会向您收取任何费用。

### 适用于大型配置的多主机安装

对于需要扫描数 PB 数据的大型配置，您可以使用多个主机来提供额外的处理能力。使用多个主机系统时，主系统称为 *Manager node*，提供额外处理能力的其他系统称为 扫描 程序 *nodes*。

在多个内部主机上安装 Data sense 软件时，请按照以下步骤进行操作。

您需要什么？ #8217 ；将需要什么

- 验证管理器和扫描程序节点的所有 Linux 系统是否都符合 [主机要求](#)。
- （可选）验证系统是否已安装两个必备软件包（ Docker 引擎和 Python 3 ）。如果此软件尚未安装在系统上，安装程序将安装此软件。
- 确保您在 Linux 系统上具有 root 权限。
- 验证您的环境是否满足要求 [权限和连接](#)。
- 您必须具有计划使用的扫描程序节点主机的 IP 地址。
- 必须在所有主机上启用以下端口和协议：

Port	协议	Description
2377	TCP	集群管理通信
7946	TCP ， UDP	节点间通信
4789	UDP	覆盖网络流量
50	电子服务	加密的 IPsec 覆盖网络（ ESP ） 流量
111.	TCP ， UDP	用于在主机之间共享文件的 NFS 服务器（需要从每个扫描程序节点到管理器节点）
2049.	TCP ， UDP	用于在主机之间共享文件的 NFS 服务器（需要从每个扫描程序节点到管理器节点）

## 步骤

1. 按照中的步骤 1 至 7 进行操作 [单主机安装](#) 在管理器节点上。
2. 如步骤 8 所示，在安装程序提示时，您可以在一系列提示中输入所需值，也可以将所需参数作为命令行参数提供给安装程序。

除了可用于单主机安装的变量之外，还会使用一个新选项 \* -n <node\_IP>\* 来指定扫描程序节点的 IP 地址。多个扫描程序节点 IP 以逗号分隔。

例如，此命令会添加 3 个扫描程序节点： `sudo ./install.sh -a <account_id> -c <agent_id> -t <token> -host <ds_host> -manager-host <cm_host> * -n <node_ip1> , <node_ip2> , <node_ip3>* -host <proxy-host-proxy-user-proxy-port-<proxy-proxy_password>`

3. 在管理器节点安装完成之前，将显示一个对话框，其中显示了扫描程序节点所需的安装命令。复制命令并将其保存在文本文件中。例如：

```
sudo ./node_install.sh -m 10.11.12.13 -t ABCDEF-1-3u69m1-1s35212
```

4. 在 \* 每个 \* 扫描程序节点主机上：
  - a. 将 Data sense 安装程序文件（ \* cc\_onprem\_installer\_<version>.tar.gz\* ）复制到主机（使用 scp 或其他方法）。
  - b. 解压缩安装程序文件。
  - c. 粘贴并执行步骤 3 中复制的命令。

在所有扫描程序节点上完成安装且这些节点已加入管理器节点后，管理器节点安装也会完成。

Cloud Data sense 安装程序将完成软件包， Docker 的安装并注册安装。安装可能需要 10 到 20 分钟。

在配置页面中，您可以选择要扫描的数据源。

您也可以 ["为 Cloud Data sense 设置许可"](#) 目前。在数据量超过 1 TB 之前，不会向您收取任何费用。

## 在内部部署 Cloud Data sense ，而无需访问 Internet

完成一些步骤，在无法访问 Internet 的内部站点中的主机上部署 Cloud Data sense 。此类安装非常适合您的安全站点。

请注意，您也可以 ["在可访问 Internet 的内部站点中部署 Data sense"](#)。

### 支持的数据源

以这种方式安装（有时称为 " 脱机 " 或 " 暗 " 站点）时， Data sense 只能扫描内部站点本地数据源中的数据。此时， Data sense 可以扫描以下本地数据源：

- 内部部署 ONTAP 系统
- 数据库架构
- 非 NetApp NFS 或 CIFS 文件共享
- 使用简单存储服务（ S3 ）协议的对象存储

如果您需要非常安全的 Cloud Manager 安装，但也希望从 OneDrive 帐户或 SharePoint 帐户扫描本地数据，则可以使用 Data sense 脱机安装程序并提供对一些选定端点的 Internet 访问。请参见 [SharePoint 和 OneDrive 的特殊要求](#) 了解详细信息。

当前不支持在非公开站点中部署Data sense时扫描Cloud Volumes ONTAP 、 Azure NetApp Files 、 FSX for ONTAP 、 AWS S3或Google Drive帐户。

### 限制

大多数 Data sense 功能都适用于部署在无法访问 Internet 的站点上的情况。但是，不支持某些需要访问 Internet 的功能，例如：

- 管理 Microsoft Azure 信息保护（ AIP ）标签
- 在某些关键策略返回结果时向 Cloud Manager 用户发送电子邮件警报
- 为不同用户设置 Cloud Manager 角色（例如，帐户管理员或合规性查看器）
- 使用 Cloud Sync 复制和同步源文件
- 接收用户反馈
- 从 Cloud Manager 自动升级软件

Cloud Manager Connector 和 Data sense 都需要定期手动升级才能启用新功能。您可以在 Data sense UI 页面底部看到 Data sense 版本。检查 ["《云数据感知发行说明》"](#) 以查看每个版本中的新功能以及是否需要这些功能。然后，您可以按照以下步骤进行操作 [升级 Data sense 软件](#)。



## 快速入门

按照以下步骤快速入门，或者向下滚动到其余部分以了解完整详细信息。

如果您尚未在脱机内部站点上安装 Connector，["部署连接器"](#) 现在在 Linux 主机上。

确保您的 Linux 系统满足 [主机要求](#)，安装了所有必需的软件，并且脱机环境满足所需的要求 [权限和连接](#)。

从 NetApp 支持站点下载 Cloud Data sense 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 Cloud Data sense 实例。

Cloud Manager 中 Cloud Data 感知扫描的前 1 TB 数据是免费的。要在此之后继续扫描数据，需要获得 NetApp 的 BYOL 许可证。

## 安装 Cloud Manager Connector

如果您尚未在脱机内部站点上安装 Cloud Manager Connector，["部署连接器"](#) 在脱机站点的 Linux 主机上。

## 准备 Linux 主机系统

数据感知软件必须在满足特定操作系统要求，RAM 要求，软件要求等要求的主机上运行。与其他应用程序共享的主机不支持数据感知 - 此主机必须是专用主机。

- 操作系统：Red Hat Enterprise Linux或CentOS 8.0、8.1或8.4版
  - 可以使用7.8或7.9版、但Linux内核版本必须为4.14或更高版本
  - 操作系统必须能够安装 Docker 引擎（例如，根据需要禁用 *firewalld* 服务）
- 磁盘：SSD，500 GiB 可在 /，或上使用
  - /opt 上提供 100 GiB
  - /var 上提供 400 GiB
  - /tmp 上 5 GiB
- RAM：64 GB（必须在主机上禁用交换内存）
- CPU：16 个核心

请注意，您可以在 CPU 较少且 RAM 较少的系统上部署 Data sense，但使用这些系统时会有一些限制。请参见 ["使用较小的实例类型"](#) 了解详细信息。

在安装 Data sense 之前，必须在主机上安装以下软件：

- Docker 引擎版本 19 或更高版本。["查看安装说明"](#)。
- Python 3 3.6 或更高版本。["查看安装说明"](#)。

## 验证 Cloud Manager 和 Data sense 前提条件

在部署 Cloud Data sense 之前，请查看以下前提条件，以确保您的配置受支持。

- 确保 Cloud Manager 有权为 Cloud Data sense 实例部署资源并创建安全组。

- 确保 Cloud Manager Connector 可以访问 Data sense 实例。Connector 的安全组必须允许通过端口 443 与 Data sense 实例之间的入站和出站流量。

通过此连接可以部署 Data sense 实例，并可查看合规性和监管信息。

确保端口 8080 已打开，以便您可以在 Cloud Manager 中查看安装进度。

- 确保您可以保持 Cloud Data sense 正常运行云数据感知实例需要保持运行状态才能持续扫描数据。
- 确保 Web 浏览器连接到 Cloud Data sense 启用 Cloud Data sense 后，请确保用户从连接到 Data sense 实例的主机访问 Cloud Manager 界面。

Data sense 实例使用专用 IP 地址来确保索引数据不可供他人访问。因此，用于访问 Cloud Manager 的 Web 浏览器必须连接到该专用 IP 地址。此连接可以来自与 Data sense 实例位于同一网络中的主机。

## SharePoint 和 OneDrive 的特殊要求

如果 Cloud Manager 和 Data sense 部署在无法访问 Internet 的站点中，则可以通过为一些选定端点提供 Internet 访问来扫描 SharePoint 和 OneDrive 帐户中的本地文件。

端点	目的
login.microsoft.com \graph.microsoft.com	与 Microsoft 服务器通信以登录到选定的联机服务。
<a href="https://cloudmanager.cloud.netapp.com">https://cloudmanager.cloud.netapp.com</a>	与 Cloud Manager 服务进行通信，其中包括 NetApp 帐户。

只有在首次连接到这些外部服务期间，才需要访问 *cloudmanager.cloud.netapp.com*。

## 部署 Data sense

对于典型配置，您将在一个主机系统上安装该软件。"请在此处查看这些步骤"。

对于需要扫描数 PB 数据的大型配置，您可以使用多个主机来提供额外的处理能力。"请在此处查看这些步骤"。

### 典型配置的单主机安装

在脱机环境中的单个内部主机上安装 Data sense 软件时，请按照以下步骤进行操作。

您需要什么？ #8217 ；将需要什么

- 验证您的 Linux 系统是否满足 [主机要求](#)。
- 确认已安装两个必备软件包（ Docker 引擎和 Python 3 ）。
- 确保您在 Linux 系统上具有 root 权限。
- 验证脱机环境是否满足要求 [权限和连接](#)。

### 步骤

1. 在已配置 Internet 的系统上，从下载 Cloud Data sense 软件 "[NetApp 支持站点](#)"。您应选择的文件名为 \* Datisis-offline-bundle-<version>.tar.gz\*。
2. 将安装程序包复制到计划在非公开站点中使用的 Linux 主机。

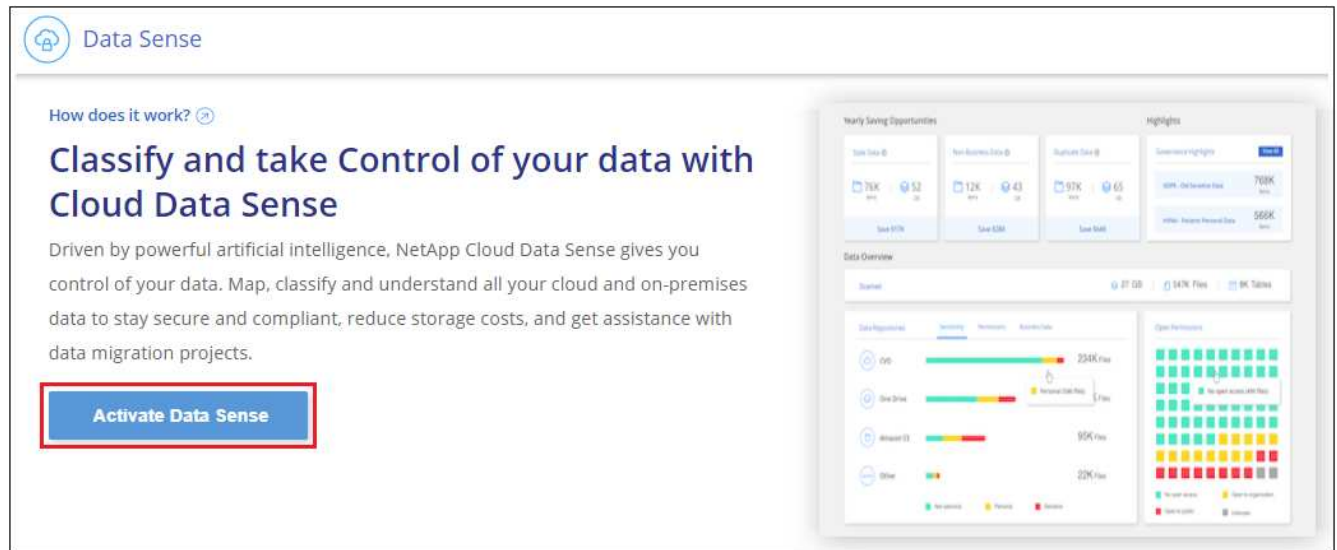


3. 解压缩主机上的安装程序包，例如：

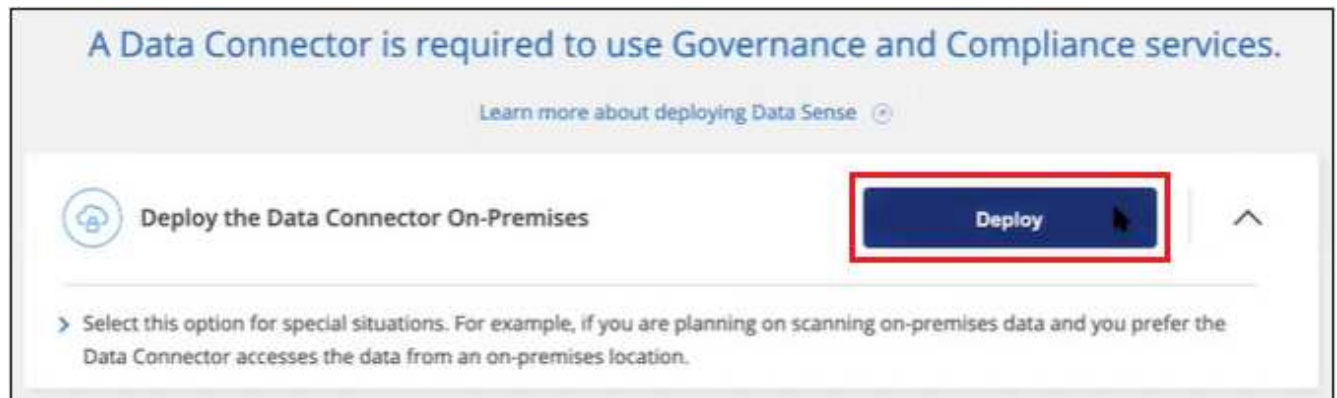
```
tar -xzf DataSense-offline-bundle-v1.13.1.tar.gz
```

此操作将提取所需的软件 and 实际安装文件 \* cc\_onprem\_installer\_< 版本 >.tar.gz\*。

4. 启动 Cloud Manager 并单击 \* 数据感知 \* 选项卡。
5. 单击 \* 激活数据感知 \*。



6. 单击 \* 部署 \* 以启动内部部署向导。



7. 在 *Deploy Data sense on premises* 对话框中，复制提供的命令并将其粘贴到文本文件中，以便稍后使用，然后单击 \* 关闭 \*。例如：

```
sudo ./install.sh -a 12345 -c 27ag75 -t 2198qq -dredestinm
```

8. 解压缩主机上的安装文件，例如：

```
tar -xzf cc_onprem_installer_1.13.1.tar.gz
```

9. 安装程序提示时，您可以在一系列提示中输入所需值，也可以将所需参数作为命令行参数提供给安装程序：

根据提示输入参数：	输入完整命令：
<p>a. 粘贴您从第 7 步复制的信息： <code>sUdo ./install.sh -a &lt;account_id&gt; -c &lt;agent_id&gt; -t &lt;token&gt; -drestrsite</code></p> <p>b. 输入 Data sense 主机的 IP 地址或主机名，以便 Connector 实例可以访问它。</p> <p>c. 输入 Cloud Manager Connector 主机的 IP 地址或主机名，以便 Data sense 实例可以访问它。</p>	<p>或者，您也可以预先创建整个命令，并提供必要的主机参数： <code>sUdo ./install.sh -a &lt;account_id&gt; -c &lt;agent_id&gt; -t &lt;token&gt; -host &lt;ds_host&gt; -manager-host &lt;cm_host&gt; -no-proxy -drestrsite</code></p>

变量值：

- `account_id` = NetApp 帐户 ID
- `agent_id` = 连接器 ID
- `token` = JWT 用户令牌
- `ds_host` = Data sense Linux 系统的 IP 地址或主机名。
- `cm_host` = Cloud Manager Connector 系统的 IP 地址或主机名。

Data sense 安装程序将安装软件包，注册安装并安装 Data sense 。安装可能需要 10 到 20 分钟。

如果主机和 Connector 实例之间通过端口 8080 建立连接，则您将在 Cloud Manager 的 Data sense 选项卡中看到安装进度。

在配置页面中，您可以选择本地 **"内部 ONTAP 集群"** 和 **"数据库"** 要扫描的。

您也可以 **"为 Cloud Data sense 设置 BYOL 许可"** 目前的数字电子钱包页面。在数据量超过 1 TB 之前，不会向您收取任何费用。

### 适用于大型配置的多主机安装

对于需要扫描数 PB 数据的大型配置，您可以使用多个主机来提供额外的处理能力。使用多个主机系统时，主系统称为 *Manager node* ，提供额外处理能力的其他系统称为 扫描 程序 *nodes* 。

在脱机环境中的多个内部主机上安装 Data sense 软件时，请按照以下步骤进行操作。

您需要什么？ **#8217** ；将需要什么

- 验证管理器和扫描程序节点的所有 Linux 系统是否都符合 **主机要求**。
- 确认已安装两个必备软件包（ Docker 引擎和 Python 3 ）。
- 确保您在 Linux 系统上具有 root 权限。
- 验证脱机环境是否满足要求 **权限和连接**。
- 您必须具有计划使用的扫描程序节点主机的 IP 地址。
- 必须在所有主机上启用以下端口和协议：

Port	协议	Description
2377	TCP	集群管理通信
7946	TCP , UDP	节点间通信
4789	UDP	覆盖网络流量
50	电子服务	加密的 IPsec 覆盖网络（ESP）流量
111.	TCP , UDP	用于在主机之间共享文件的 NFS 服务器（需要从每个扫描程序节点到管理器节点）
2049.	TCP , UDP	用于在主机之间共享文件的 NFS 服务器（需要从每个扫描程序节点到管理器节点）

## 步骤

1. 按照中的步骤 1 至 8 进行操作 **"单主机安装"** 在管理器节点上。
2. 如步骤 9 所示，在安装程序提示时，您可以在一系列提示中输入所需值，也可以将所需参数作为命令行参数提供给安装程序。

除了可用于单主机安装的变量之外，还会使用一个新选项 `* -n <node_ip>*` 来指定扫描程序节点的 IP 地址。多个节点 IP 以逗号分隔。

例如，此命令会添加 3 个扫描程序节点：`sudo ./install.sh -a <account_id> -c <agent_id> -t <token> -host <ds_host> -manager-host <cm_host> * -n <node_ip1> , <node_ip2> , <node_ip3>*` **-无代理站点**

3. 在管理器节点安装完成之前，将显示一个对话框，其中显示了扫描程序节点所需的安装命令。复制命令并将其保存在文本文件中。例如：

```
sudo ./node_install.sh -m 10.11.12.13 -t ABCDEF-1-3u69m1-1s35212
```

4. 在 `* 每个 *` 扫描程序节点主机上：
  - a. 将 Data sense 安装程序文件（`* cc_onprem_installer_<version>.tar.gz*`）复制到主机。
  - b. 解压缩安装程序文件。
  - c. 粘贴并运行在步骤 3 中复制的命令。

在所有扫描程序节点上完成安装且这些节点已加入管理器节点后，管理器节点安装也会完成。

Cloud Data sense 安装程序将完成软件包安装，并注册安装。安装可能需要 15 到 25 分钟。

在配置页面中，您可以选择本地 **"内部 ONTAP 集群"** 和本地 **"数据库"** 要扫描的。

您也可以 **"为 Cloud Data sense 设置 BYOL 许可"** 目前的数字电子钱包页面。在数据量超过 1 TB 之前，不会向您收取任何费用。

## 升级 Data sense 软件

由于 Data sense 软件会定期更新新功能，因此您应按照例行程序定期检查新版本，以确保您使用的是最新的软件和功能。您需要手动升级 Data sense 软件，因为没有 Internet 连接，无法自动执行升级。

## 开始之前

- 数据感知软件一次可升级一个主要版本。例如、如果您安装了1.1.1.x版、则只能升级到1.12.x如果您有几个主要版本，则需要多次升级此软件。
- 确认您的内部连接器软件已升级到最新可用版本。"请参见 [Connector 升级步骤](#)"。

## 步骤

1. 在已配置 Internet 的系统上，从下载 Cloud Data sense 软件 "[NetApp 支持站点](#)"。您应选择的文件名为 \* Datasense-offline-bundle-<version>.tar.gz\* 。
2. 将软件包复制到非公开站点中安装了 Data sense 的 Linux 主机。
3. 解压缩主机上的软件包，例如：

```
tar -xvf DataSense-offline-bundle-v1.13.1.tar.gz
```

此操作将提取安装文件 \* cc\_onprem\_installer\_< 版本 >.tar.gz\* 。

4. 解压缩主机上的安装文件，例如：

```
tar -xzf cc_onprem_installer_1.13.1.tar.gz
```

此操作将提取升级脚本 \* 启动 \_didssite\_upgrade.sh\* 以及任何所需的第三方软件。

5. 在主机上运行升级脚本，例如：

```
start_darksite_upgrade.sh
```

Data sense 软件将在主机上进行升级。更新可能需要 5 到 10 分钟。

请注意，如果您在多个主机系统上部署了 Data sense 来扫描非常大的配置，则扫描程序节点不需要升级。

您可以通过检查 Data sense UI 页面底部的版本来验证软件是否已更新。

## 版权信息

版权所有©2022 NetApp、Inc.。保留所有权利。Printed in the U.S.版权所涵盖的本文档的任何部分不得以任何形式或任何手段复制、包括影印、录制、磁带或存储在电子检索系统中—未经版权所有者事先书面许可。

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

本软件由NetApp按"原样"提供、不含任何明示或默示担保、包括但不限于适销性和特定用途适用性的默示担保、特此声明不承担任何任何责任。IN NO EVENT SHALL NETAPP BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

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

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

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

## 商标信息

NetApp、NetApp标识和中列出的标记 <http://www.netapp.com/TM> 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。