



入门 Cloud Data Sense

NetApp
July 13, 2022

目录

- 入门 1
 - 了解 Cloud Data sense 1
 - 部署 Cloud Data sense 6
 - 激活对数据源的扫描 23
 - 将 Active Directory 与 Cloud Data sense 集成 60
 - 为 Cloud Data sense 设置许可 63
 - 有关 Cloud Data sense 的常见问题 68

入门

了解 Cloud Data sense

Cloud Data sense 是 Cloud Manager 的一项数据管理服务，可对企业内部和云数据源和工作环境进行扫描，以映射和分类数据并识别私有信息。这有助于降低安全性和合规性风险，降低存储成本，并有助于您的数据迁移项目。

["了解 Cloud Data sense 的用例"](#)。

功能

Cloud Data sense 提供了多种工具，可帮助您完成合规性工作。您可以使用 Data sense：

- 识别个人信息（PII）
- 根据 GDPR，CCPA，PCI 和 HIPAA 隐私法规的要求确定广泛的敏感信息
- 响应数据主体访问请求（DSAr）
- 当文件包含特定的 PIE 时，通过电子邮件通知 Cloud Manager 用户（您可以使用定义此条件）["策略"](#)）
- 查看和修改 ["Azure 信息保护（AIP）标签"](#) 在文件中
- 向文件添加自定义标记（例如 "需要移动"），并分配 Cloud Manager 用户，以便用户可以拥有文件更新
- 复制，移动和删除文件

Cloud Data sense 还提供了一些工具，可以帮助您进行监管工作。您可以使用 Cloud Data sense：

- 确定系统中的陈旧数据，非业务数据，重复文件，具有打开权限的文件以及非常大的文件。
您可以使用此信息来确定是要将某些文件移动，删除，还是将其分层到成本较低的对象存储。
- 在移动数据之前，查看数据的大小以及任何数据是否包含敏感信息。

如果您计划将数据从内部位置迁移到云，则此功能非常有用。

支持的工作环境和数据源

Cloud Data sense 可以扫描以下类型的工作环境和数据源中的数据：

- 工作环境： *
- Cloud Volumes ONTAP（部署在 AWS，Azure 或 GCP 中）
- 内部 ONTAP 集群
- Azure NetApp Files
- 适用于 ONTAP 的 Amazon FSX
- Amazon S3
- 数据源： *

- 非 NetApp 文件共享
- 对象存储（使用 S3 协议）
- 数据库
- OneDrive 帐户
- SharePoint 帐户
- Google Drive 帐户

Data sense 支持 NFS 3.x，4.0 和 4.1 以及 CIFS 1.x，2.0，2.1 和 3.0 版。

成本

- 使用 Cloud Data sense 的成本取决于您要扫描的数据量。Data sense 在 Cloud Manager 工作空间中扫描的前 1 TB 数据是免费的。这包括所有工作环境和数据源中的所有数据。要在这之后继续扫描数据，需要订阅 AWS，Azure 或 GCP Marketplace 或 NetApp 提供的 BYOL 许可证。请参见 ["定价"](#) 了解详细信息。

["了解如何获得 Cloud Data sense 的许可"](#)。

- 要在云中安装 Cloud Data sense，需要部署云实例，这会导致部署该实例的云提供商收取费用。请参见 [为每个云提供商部署的实例类型](#)。如果您在内部系统上安装 Data sense，则无需任何成本。
- Cloud Data sense 要求您已部署 Cloud Manager Connector。在许多情况下，由于您在 Cloud Manager 中使用的其他存储和服务，您已经有了 Connector。Connector 实例会从部署该实例的云提供商处收取费用。请参见 ["为每个云提供商部署的实例类型"](#)。如果在内部部署系统上安装 Connector，则不需要任何成本。

数据传输成本

数据传输成本取决于您的设置。如果云数据感知实例和数据源位于同一可用性区域和区域，则不会产生数据传输成本。但是，如果数据源（例如 Cloud Volumes ONTAP 系统或 S3 存储分段）位于 *Different* 可用性区域或区域，则云提供商会向您收取数据传输成本。有关详细信息，请参见以下链接：

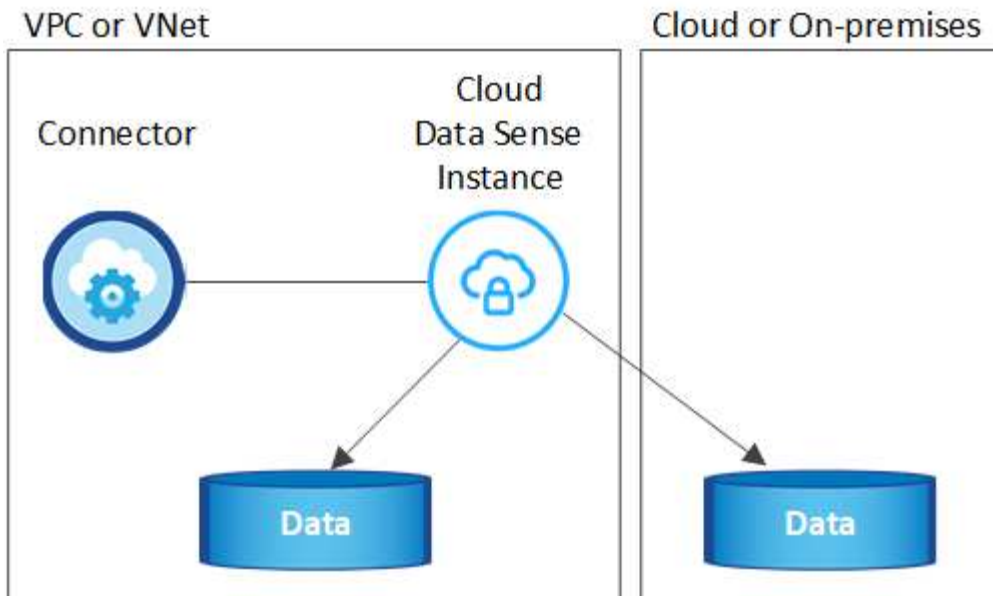
- ["AWS：Amazon EC2 定价"](#)
- ["Microsoft Azure：带宽定价详细信息"](#)
- ["Google Cloud：存储传输服务定价"](#)

云数据感知实例

在云中部署 Data sense 时，Cloud Manager 会将实例部署在与 Connector 相同的子网中。 ["了解有关连接器的更多信息。"](#)



如果 Connector 安装在内部，则它会将云数据感知实例部署在与请求中的第一个 Cloud Volumes ONTAP 系统相同的 VPC 或 vNet 中。您也可以在内部安装 Data sense。



请注意以下有关默认实例的信息：

- 在 AWS 中，Cloud Data sense 在上运行 "m5.4xlarge 实例" 使用 500 GB GP2 磁盘。操作系统映像为 Amazon Linux 2 （ Red Hat 7.3.1 ）。

在 m5.4xlarge 不可用的区域中，Data sense 会在 m4.4xlarge 实例上运行。

- 在 Azure 中，Cloud Data sense 在上运行 "标准的 D16s_v3 VM" 使用 512 GB 磁盘。操作系统映像为 CentOS 7.8 。
- 在 GCP 中，Cloud Data sense 在上运行 "n2-standard-16 虚拟机" 使用 512 GB 标准持久性磁盘。操作系统映像为 CentOS 7.9 。

在 n2-standard-16 不可用的区域中，Data sense 运行在 n2D-standard-16 或 n1-standard-16 VM 上。

- 此实例名为 *CloudCompliance* ，并与生成的哈希（UUID）串联在一起。例如： *CloudCompliance" — 16bb6564-38AD-4080-9a92 — 36f5fd2f71c7*
- 每个连接器只部署一个数据感知实例。
- 只要实例可以访问 Internet ，就会自动升级 Data sense 软件。



此实例应始终保持运行状态，因为 Cloud Data sense 会持续扫描数据。

使用较小的实例类型

您可以在 CPU 较少且 RAM 较少的系统上部署 Data sense ，但使用这些功能较差的系统时会存在一些限制。

系统大小	规格	限制
超大（默认）	16 个 CPU ， 64 GB RAM ， 500 GB SSD	无
中等	8 个 CPU ， 32 GB RAM ， 200 GB SSD	扫描速度较慢，最多只能扫描 100 万个文件。

系统大小	规格	限制
小型	8 个 CPU ， 16 GB RAM ， 100 GB SSD	限制与 " 中等 " 相同，并且还可以识别 "数据主题名称" 已禁用内部文件。

在云中部署 Data sense 时，如果您要使用其中一个较小的系统，请发送电子邮件至 ng-contact-data-sense@netapp.com 以获得帮助。我们需要与您合作来部署这些较小的云配置。

在内部部署 Data sense 时，只需使用规格较小的 Linux 主机即可。您无需联系 NetApp 以获得帮助。

云数据感知的工作原理

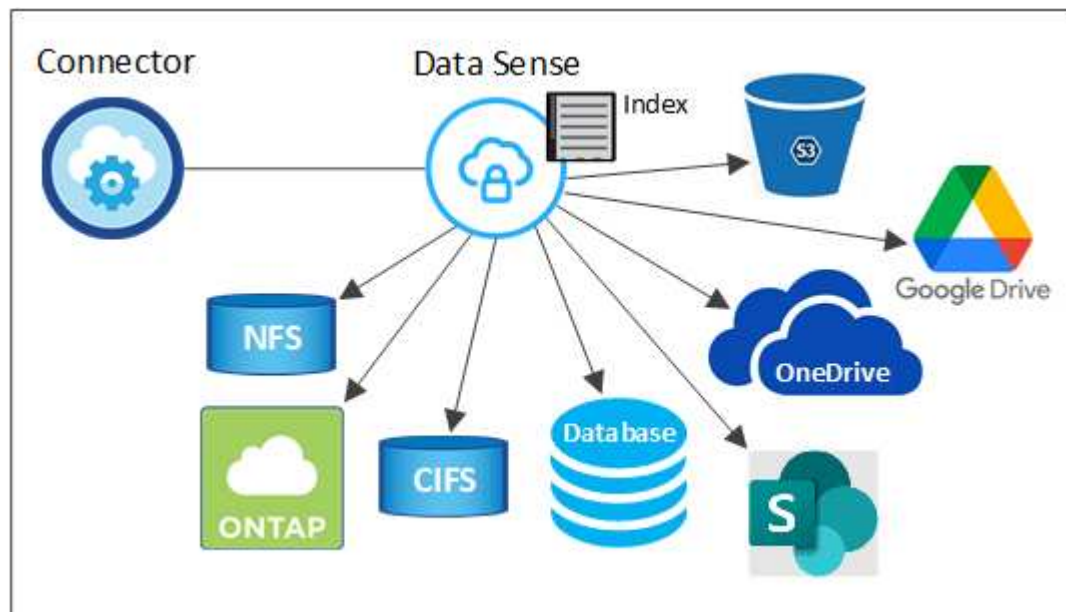
概括地说，Cloud Data sense 的工作原理如下：

1. 您可以在 Cloud Manager 中部署 Data sense 实例。
2. 您可以在一个或多个工作环境或数据源上启用高级别映射或深度扫描。
3. Data sense 使用 AI 学习过程扫描数据。
4. 您可以使用提供的信息板和报告工具帮助您开展合规和监管工作。

扫描的工作原理

启用 Cloud Data sense 并选择要扫描的卷，分段，数据库架构或 OneDrive 或 SharePoint 用户数据后，它将立即开始扫描数据以确定个人和敏感数据。它会映射您的组织数据，对每个文件进行分类，并标识和提取数据中的实体和预定义模式。扫描的结果是个人信息，敏感个人信息，数据类别和文件类型的索引。

Data sense 通过挂载 NFS 和 CIFS 卷与任何其他客户端一样连接到数据。NFS 卷会自动以只读方式访问，而您需要提供 Active Directory 凭据来扫描 CIFS 卷。



在初始扫描之后，Data sense 会持续扫描数据以检测增量更改（因此，保持实例正常运行非常重要）。

您可以在卷级别，存储分段级别，数据库架构级别，OneDrive 用户级别和 SharePoint 站点级别启用和禁用扫描。

映射扫描与分类扫描有何区别

您可以通过 Cloud Data sense 对选定工作环境和数据源运行常规 "映射" 扫描。映射仅提供数据的概览，而 "分类" 则提供数据的深度扫描。由于无法访问文件以查看数据源中的数据，因此可以非常快速地对数据源进行映射。

许多用户喜欢此功能，因为他们希望快速扫描其数据以确定需要更多研究的数据源，然后只能对所需的数据源或卷启用分类扫描。

下表显示了一些差异：

功能	分类	映射
扫描速度	速度较慢	快速
文件类型和已用容量的列表	是的。	是的。
文件数和已用容量	是的。	是的。
文件的期限和大小	是的。	是的。
能够运行 "数据映射报告"	是的。	是的。
数据调查页面以查看文件详细信息	是的。	否
搜索文件中的名称	是的。	否
创建 "策略" 可提供自定义搜索结果	是的。	否
使用 AIP 标签和状态标记对数据进行分类	是的。	否
复制，删除和移动源文件	是的。	否
能够运行其他报告	是的。	否

云数据感知的索引信息

Data sense 收集数据（文件）并为其创建索引和分配类别。Data sense 索引的数据包括以下内容：

标准元数据

Cloud Data sense 收集有关文件的标准元数据：文件类型，大小，创建和修改日期等。

个人数据

个人身份信息，例如电子邮件地址，标识号或信用卡号。 ["了解有关个人数据的更多信息"](#)。

敏感的个人数据

GDPR 和其他隐私法规定义的特殊类型的敏感信息，例如健康数据，种族或政治观点。 ["了解有关敏感个人数据的更多信息"](#)。

类别

Cloud Data sense 会将扫描的数据划分为不同类型的类别。类别是基于 AI 对每个文件的内容和元数据的分析而得出的主题。 ["了解有关类别的更多信息"](#)。

类型

Cloud Data sense 会提取所扫描的数据，并按文件类型对其进行细分。 ["了解有关类型的更多信息"](#)。

名称实体识别

Cloud Data sense 使用 AI 从文档中提取自然人的姓名。 ["了解如何响应数据主体访问请求"](#)。

网络概述

Cloud Manager 将云数据感知实例部署到一个安全组中，该安全组可从 Connector 实例启用入站 HTTP 连接。

在 SaaS 模式下使用 Cloud Manager 时，将通过 HTTPS 提供与 Cloud Manager 的连接，并通过端到端加密保护浏览器与 Data sense 实例之间发送的私有数据，这意味着 NetApp 和第三方无法读取这些数据。

出站规则完全开放。要安装和升级 Data sense 软件以及发送使用量指标，需要访问 Internet 。

如果您有严格的网络连接要求， ["了解 Cloud Data 感知所接触的端点"](#)。

用户访问合规性信息

为每个用户分配的角色可在 Cloud Manager 和 Cloud Data sense 中提供不同的功能：

- * 帐户管理员 * 可以管理所有工作环境的合规性设置并查看合规性信息。
- 只有当系统具有访问权限时，* 工作空间管理员 * 才能管理合规性设置并查看合规性信息。如果 Workspace 管理员无法在 Cloud Manager 中访问工作环境，则他们无法在 "数据感知" 选项卡中查看工作环境的任何合规性信息。
- 具有 * 合规性查看器 * 角色的用户只能查看其有权访问的系统的合规性信息并生成报告。这些用户无法启用 / 禁用卷，分段或数据库架构的扫描。这些用户也无法复制，移动或删除文件。

["了解有关 Cloud Manager 角色的更多信息"](#) 以及操作方法 ["添加具有特定角色的用户"](#)。

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详细信息"](#)。

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 的用户也可以选择内部安装 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
https://cloudmanager.cloud.netapp.com	与 Cloud Manager 服务进行通信，其中包括 NetApp 帐户。
https://netapp-cloud-account.auth0.com https://auth0.com	与 NetApp Cloud Central 进行通信以实现集中式用户身份验证。
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://dseasb33srrn.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 能够从审计记录流化数据。
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	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
https://cloudmanager.cloud.netapp.com	与 Cloud Manager 服务进行通信，其中包括 NetApp 帐户。
https://netapp-cloud-account.auth0.com https://auth0.com	与 NetApp Cloud Central 进行通信以实现集中式用户身份验证。
https://support.compliance.cloudmanager.cloud.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/	可用于访问软件映像，清单，模板以及发送日志和指标。
https://support.compliance.cloudmanager.cloud.netapp.com/	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。

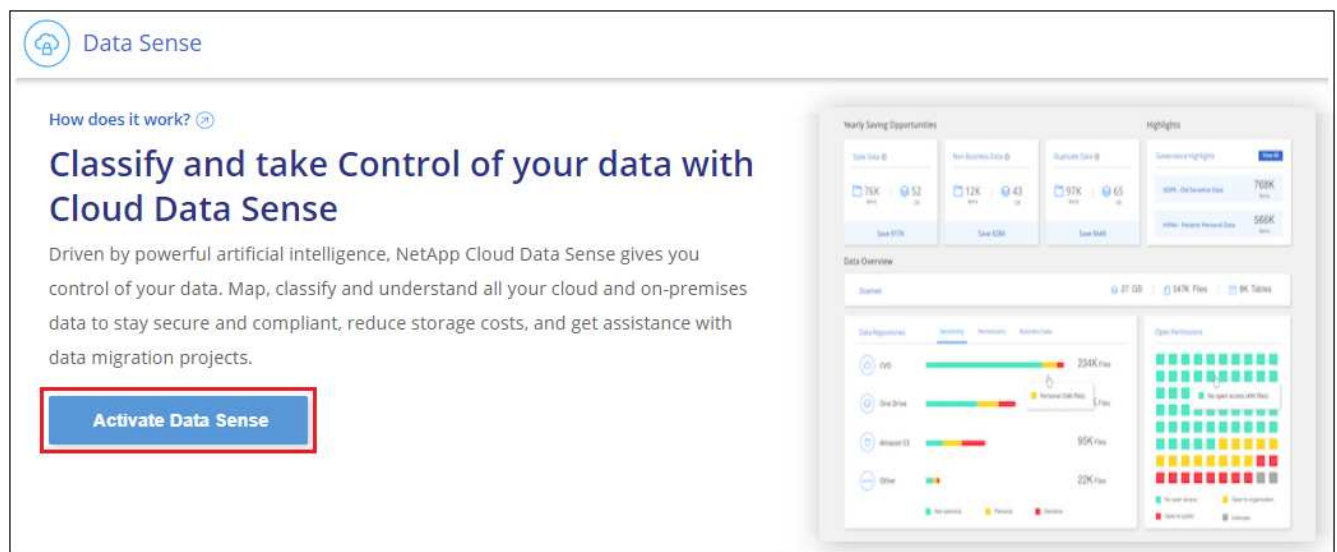
数据感知实例使用专用 IP 地址来确保索引数据无法通过 Internet 访问。因此，用于访问 Cloud Manager 的 Web 浏览器必须连接到该专用 IP 地址。此连接可以来自与云提供商的直接连接（例如 VPN），也可以来自与 Data sense 实例位于同一网络中的主机。

在云中部署 Data sense

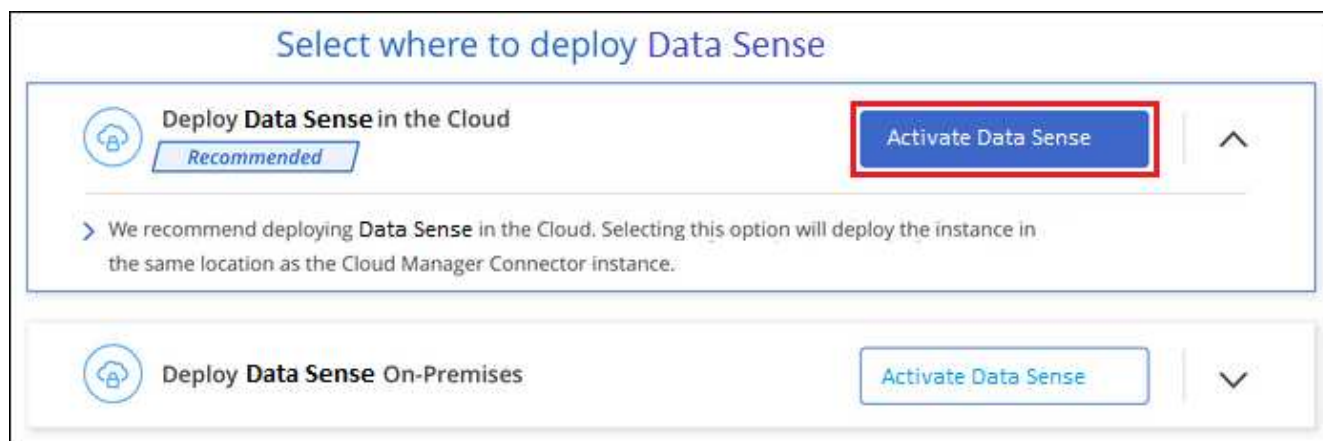
按照以下步骤在云中部署 Cloud Data sense 实例。

步骤

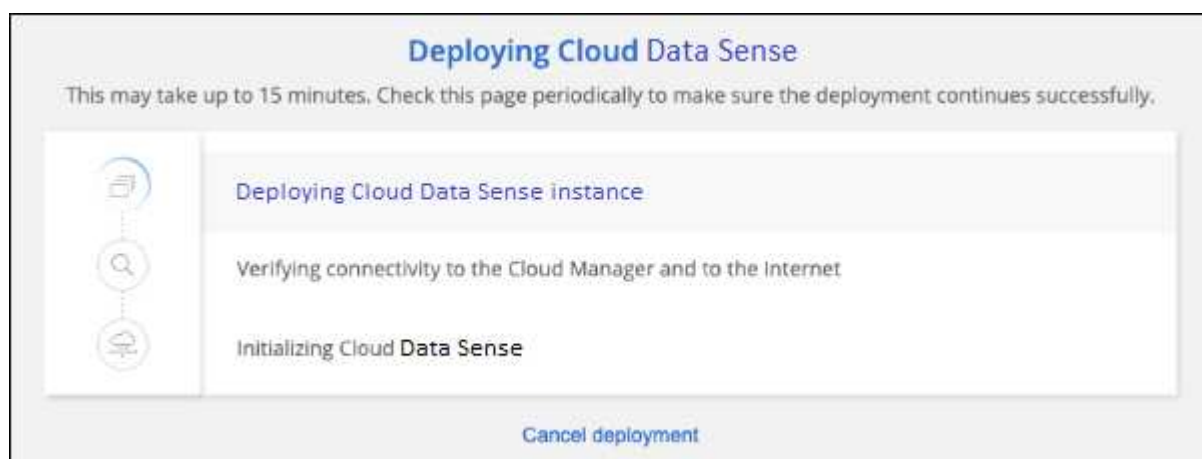
1. 从 Cloud Manager 左侧导航菜单中、单击*数据感知*。
2. 单击 * 激活数据感知 *。



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
https://cloudmanager.cloud.netapp.com	与 Cloud Manager 服务进行通信，其中包括 NetApp 帐户。
https://netapp-cloud-account.auth0.com https://auth0.com	与 NetApp Cloud Central 进行通信以实现集中式用户身份验证。
https://support.compliance.cloudmanager.cloud.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/	可用于访问软件映像，清单，模板以及发送日志和指标。
https://support.compliance.cloudmanager.cloud.netapp.com/	使 NetApp 能够从审计记录流化数据。
https://github.com/docker https://download.docker.com http://mirror.centos.org http://mirrorlist.centos.org http://mirror.centos.org/centos/7/extras/x86_64/Packages/container-selinux-2.107-3.el7.noarch.rpm	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

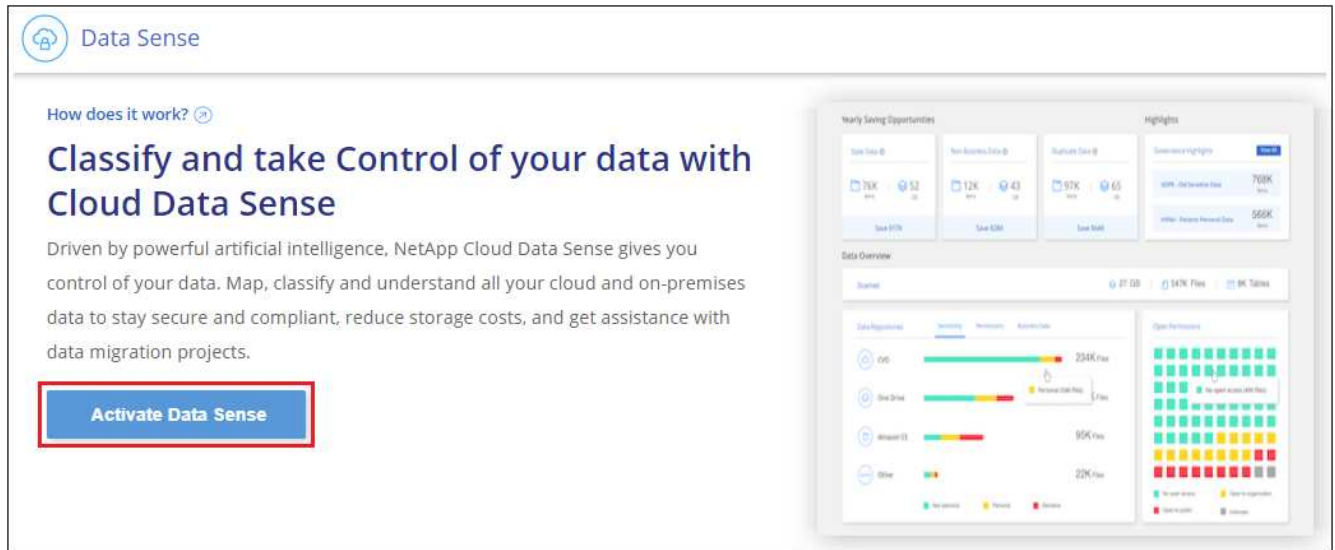
在单个内部主机上安装 Data sense 软件时，请按照以下步骤进行操作。

您需要什么？ **#8217** ；将需要什么

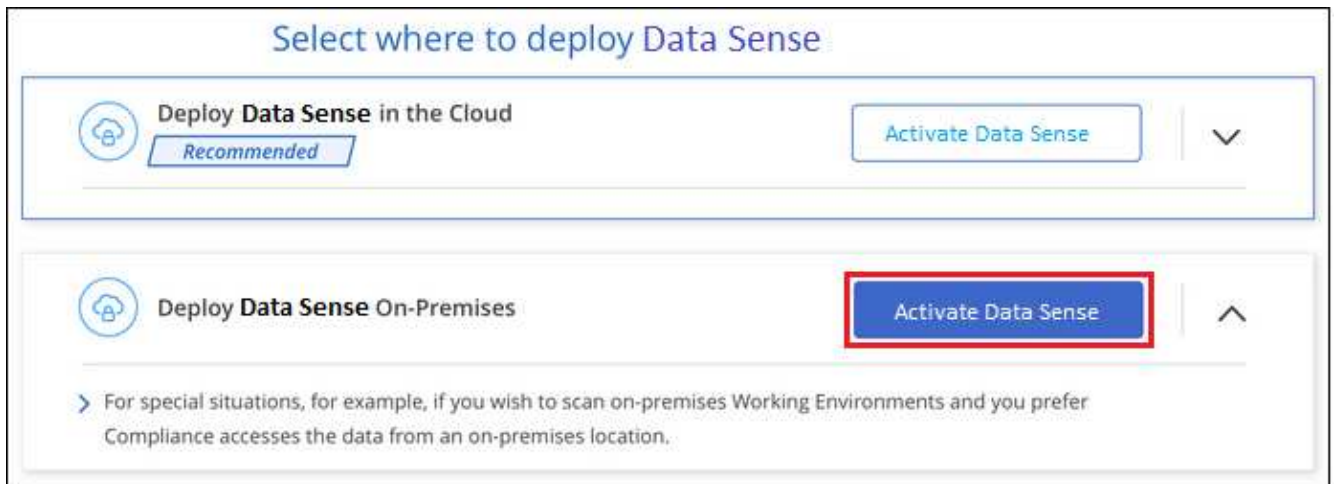
- 验证您的 Linux 系统是否满足 [主机要求](#)。
- （可选）验证系统是否已安装两个必备软件包（ Docker 引擎和 Python 3 ）。如果此软件尚未安装在系统上，安装程序将安装此软件。
- 确保您在 Linux 系统上具有 root 权限。
- 如果您使用的是代理，并且代理正在执行 TLS 截获，则需要了解 Data sense Linux 系统上用于存储 TLS CA 证书的路径。
- 验证脱机环境是否满足要求 [权限和连接](#)。

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 <account_id> -c</code> <code><agent_id> -t <token></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>sudo ./install.sh -a <account_id> -c</code> <code><agent_id> -t <token> -host <ds_host> -manager</code> <code>-host <cm_host> -proxy-host <proxy_host> -proxy</code> <code>-port <proxy_port> -proxy-user <proxy_name></code> <code><proxy_password> -proxy-proxy <proxy_name></code> <code>-proxy-proxy-proxy <proxy_name></code> <code>-<proxy_user></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。
https://cloudmanager.cloud.netapp.com	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：

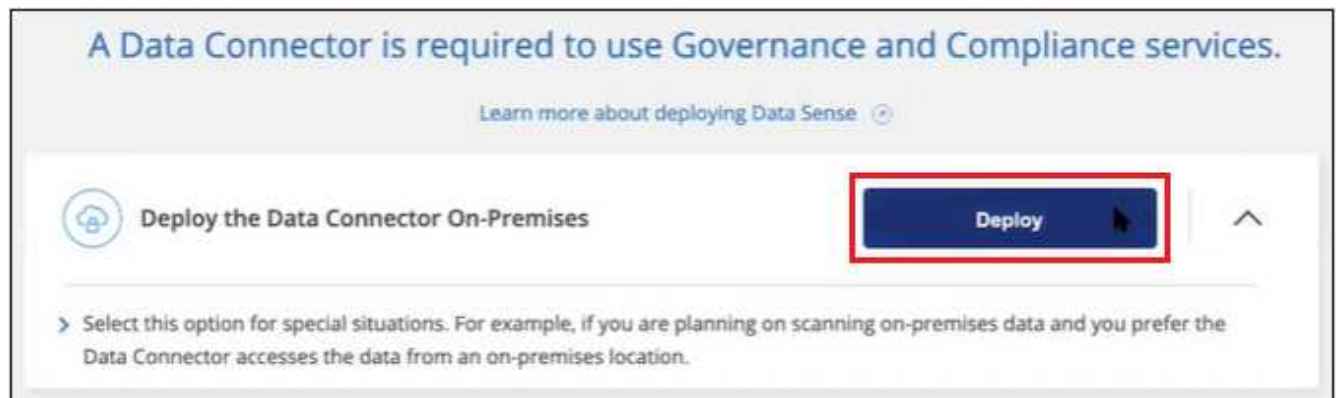
```
tar -xzf DataSense-offline-bundle-v1.13.1.tar.gz
```

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： sUdo ./install.sh -a <account_id> -c <agent_id> -t <token> -drestrsite</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： sUdo ./install.sh -a <account_id> -c <agent_id> -t <token> -host <ds_host> -manager-host <cm_host> -no-proxy -drestrsite</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

Port	协议	Description
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。

激活对数据源的扫描

开始使用适用于 **Cloud Volumes ONTAP** 和内部 **ONTAP** 的云数据感知

完成几个步骤，开始使用 Cloud Data sense 扫描 Cloud Volumes ONTAP 和内部 ONTAP 卷。

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。

在扫描卷之前，必须在 Cloud Manager 中将系统添加为工作环境：

- 对于 Cloud Volumes ONTAP 系统，这些工作环境应已在 Cloud Manager 中可用
- 对于内部 ONTAP 系统，"[Cloud Manager 必须发现 ONTAP 集群](#)"

"部署 [Cloud Data sense](#)" 如果尚未部署实例。

单击 * 数据感知 *，选择 * 配置 * 选项卡，然后为特定工作环境中的卷激活合规性扫描。

启用 Cloud Data sense 后，请确保它可以访问所有卷。

- 云数据感知实例需要与每个 Cloud Volumes ONTAP 子网或内部 ONTAP 系统建立网络连接。
- Cloud Volumes ONTAP 的安全组必须允许来自数据感知实例的入站连接。
- 确保这些端口对 Data sense 实例开放：
 - 对于 NFS —端口 111 和 2049。
 - 对于 CIFS —端口 139 和 445。
- NFS 卷导出策略必须允许从 Data sense 实例进行访问。
- Data sense 需要 Active Directory 凭据才能扫描 CIFS 卷。

单击 * 合规性 * > * 配置 * > * 编辑 CIFS 凭据 * 并提供凭据。

选择或取消选择要扫描的卷，Cloud Data sense 将开始或停止扫描这些卷。

发现要扫描的数据源

如果您要扫描的数据源尚未位于 Cloud Manager 环境中，则可以此时将其添加到画布中。

您的 Cloud Volumes ONTAP 系统应已在 Cloud Manager 的 "画布" 中可用。对于内部部署的 ONTAP 系统，您需要拥有 "[Cloud Manager 将发现这些集群](#)"。

部署 **Cloud Data sense** 实例

如果尚未部署实例，请部署 Cloud Data sense。

如果您要扫描可通过 Internet 访问的 Cloud Volumes ONTAP 和内部 ONTAP 系统，则可以 "[在云中部署 Cloud Data sense](#)" 或 "[位于可访问 Internet 的内部位置](#)"。

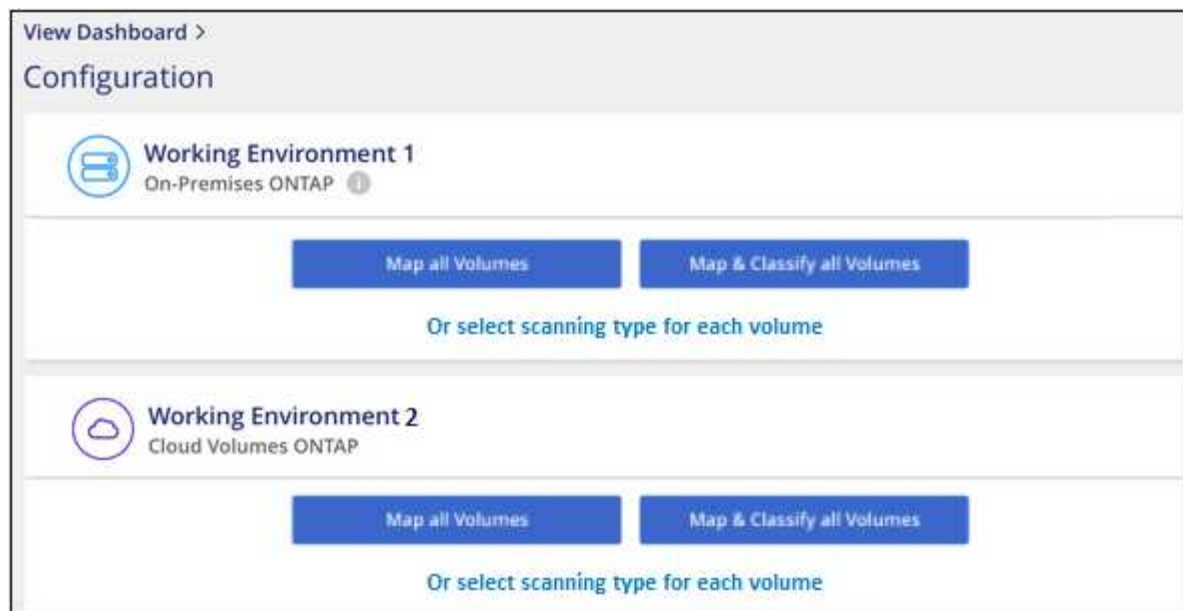
如果要扫描安装在无法访问 Internet 的非公开站点上的内部 ONTAP 系统，则需要执行以下操作 "[在无法访问 Internet 的同一内部位置部署 Cloud Data sense](#)"。这还要求 Cloud Manager Connector 部署在同一内部位置。

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

在您的工作环境中实现云数据感知

您可以在任何受支持的云提供商的 Cloud Volumes ONTAP 系统和内部 ONTAP 集群上启用云数据感知。

1. 从 Cloud Manager 左侧导航菜单中、单击 * 数据感知 *、然后选择 * 配置 * 选项卡。



2. 选择要如何扫描每个工作环境中的卷。 ["了解映射和分类扫描"](#):

- 要映射所有卷，请单击 * 映射所有卷 *。
- 要映射所有卷并对其进行分类，请单击 * 映射并分类所有卷 *。
- 要自定义每个卷的扫描，请单击 * 或选择每个卷的扫描类型 *，然后选择要映射和 / 或分类的卷。

请参见 [在卷上启用和禁用合规性扫描](#) 了解详细信息。

3. 在确认对话框中，单击 * 批准 * 以使 Data sense 开始扫描卷。

Cloud Data sense 开始扫描您在工作环境中选择的卷。一旦 Cloud Data sense 完成初始扫描，" 合规性 " 信息板将显示结果。所需时间取决于数据量—可能需要几分钟或几小时。

验证 **Cloud Data sense** 是否有权访问卷

通过检查网络，安全组和导出策略，确保 Cloud Data sense 可以访问卷。您需要为 Data sense 提供 CIFS 凭据，以便它可以访问 CIFS 卷。

步骤

1. 确保云数据感知实例与包含 Cloud Volumes ONTAP 或内部 ONTAP 集群卷的每个网络之间存在网络连接。
2. 确保 Cloud Volumes ONTAP 的安全组允许来自数据感知实例的入站流量。

您可以从 Data sense 实例的 IP 地址打开流量安全组，也可以从虚拟网络内部打开所有流量的安全组。

3. 确保以下端口对 Data sense 实例开放：

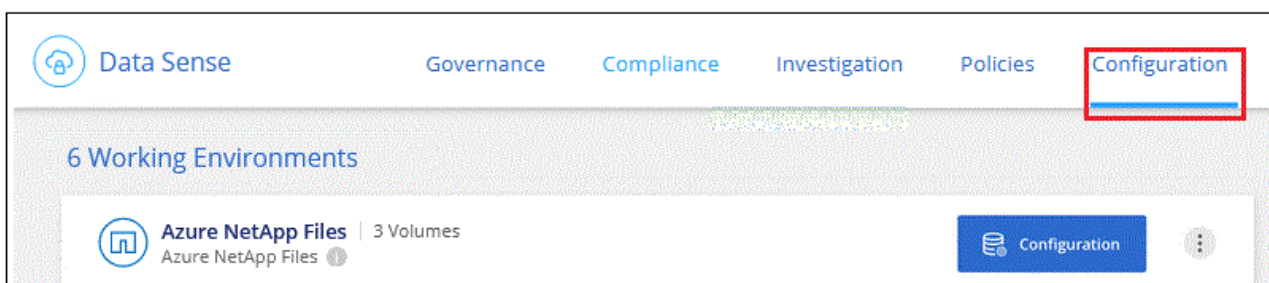
- 对于 NFS —端口 111 和 2049。
- 对于 CIFS —端口 139 和 445。

4. 确保 NFS 卷导出策略包含 Data sense 实例的 IP 地址，以便它可以访问每个卷上的数据。

5. 如果使用 CIFS，请为 Data sense 提供 Active Directory 凭据，以便它可以扫描 CIFS 卷。

- a. 在 Cloud Manager 顶部，单击 * 数据感知 *。

b. 单击 * 配置 * 选项卡。

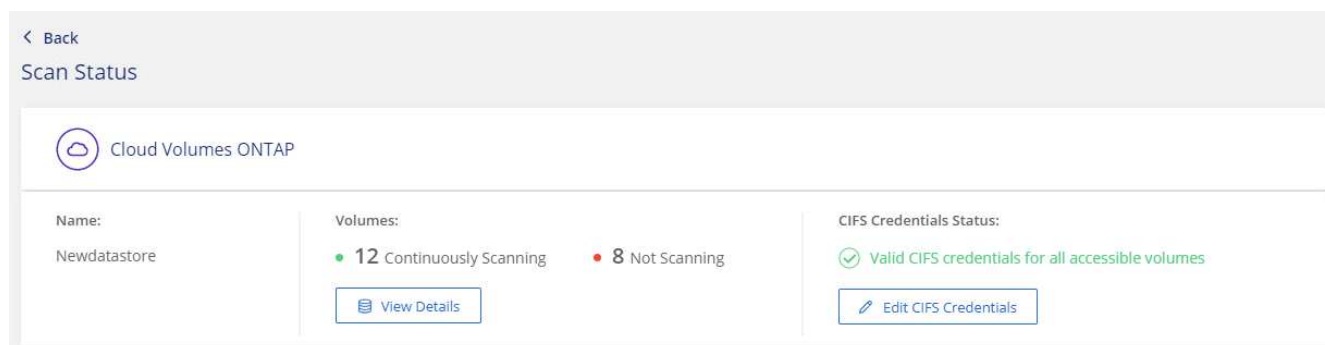


c. 对于每个工作环境，单击 * 编辑 CIFS 凭据 *，然后输入 Data sense 访问系统上 CIFS 卷所需的用户名和密码。

这些凭据可以是只读的，但提供管理员凭据可确保 Data sense 可以读取任何需要提升权限的数据。这些凭据存储在 Cloud Data sense 实例上。

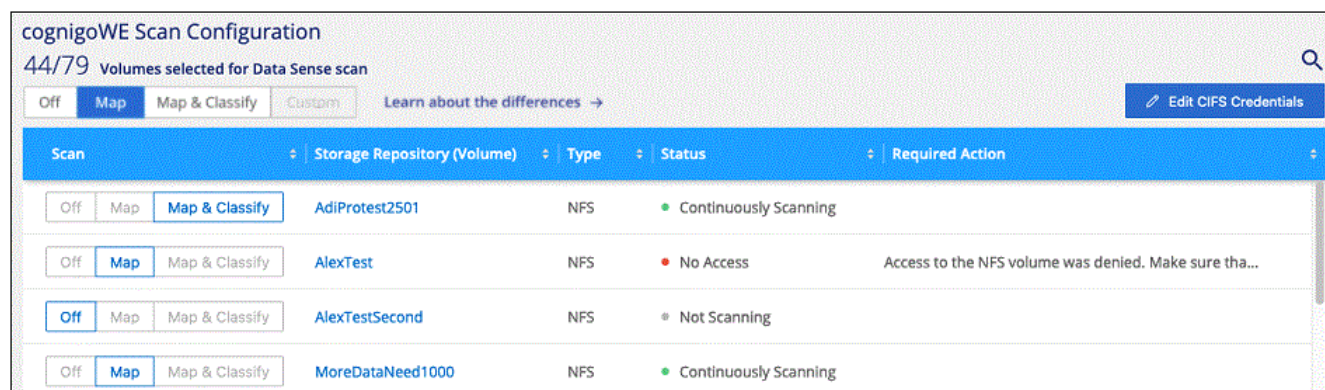
如果要确保数据感知分类扫描不会更改文件的"上次访问时间"、我们建议用户具有"写入属性"权限。如果可能、我们建议将Active Directory配置的用户设置为组织中有权访问所有文件的父组的一部分。

输入凭据后，您应看到一条消息，指出所有 CIFS 卷均已成功通过身份验证。



6. 在 *Configuration* 页面上，单击 * 查看详细信息 * 以查看每个 CIFS 和 NFS 卷的状态并更正任何错误。

例如，下图显示了四个卷；其中一个卷由于 Data sense 实例与卷之间的网络连接问题而无法扫描。



在卷上启用和禁用合规性扫描

您可以随时从 " 配置 " 页面在工作环境中启动或停止仅映射扫描或映射和分类扫描。您也可以从仅映射扫描更改

为映射和分类扫描，反之亦然。建议您扫描所有卷。

cognigoWE Scan Configuration

44/79 Volumes selected for Data Sense scan

OffMapMap & ClassifyCustom

Learn about the differences →

Edit CIFS Credentials

Scan	Storage Repository (Volume)	Type	Status	Required Action
<div>OffMapMap & Classify</div>	AdiNFSVol_copy	NFS	No Access	Access to the NFS volume was denied. Make sure tha...
<div>OffMapMap & Classify</div>	AdiProtest2501	NFS	Continuously Scanning	
<div>OffMapMap & Classify</div>	AlexTest	NFS	No Access	Access to the NFS volume was denied. Make sure tha...
<div>OffMapMap & Classify</div>	AlexTestSecond	NFS	Not Scanning	
<div>OffMapMap & Classify</div>	MoreDataNeed1000	NFS	Continuously Scanning	

收件人：	执行以下操作：
在卷上启用仅映射扫描	在卷区域中，单击 * 映射 *
对卷启用完全扫描	在卷区域中，单击 * 映射和分类 *
禁用对卷的扫描	在卷区域中，单击 * 关闭 *
在所有卷上启用仅映射扫描	在标题区域中，单击 * 映射 *
对所有卷启用完全扫描	在标题区域中，单击 * 映射和分类 *
禁用对所有卷的扫描	在标题区域中，单击 * 关闭 *



只有在标题区域中设置了 * 映射 * 或 * 映射和分类 * 设置后，才会自动扫描添加到工作环境中的新卷。如果在标题区域中设置为 * 自定义 * 或 * 关闭 *，则需要在工作环境中添加的每个新卷上激活映射和 / 或完全扫描。

扫描数据保护卷

默认情况下，不会扫描数据保护（DP）卷，因为它们不会公开在外部，并且 Cloud Data sense 无法访问它们。这些卷是从内部 ONTAP 系统或 Cloud Volumes ONTAP 系统执行 SnapMirror 操作的目标卷。

最初，卷列表会将这些卷标识为 *Type* * dp*，并显示 *Status* * 未扫描 * 和 *Required Action* * Enable Access to DP volumes*。

'Working Environment Name' Configuration

22/28 Volumes selected for compliance scan

OffMapMap & ClassifyCustom

Learn about the differences →

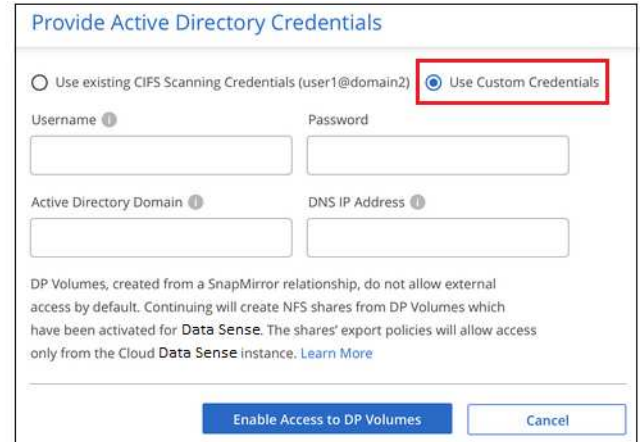
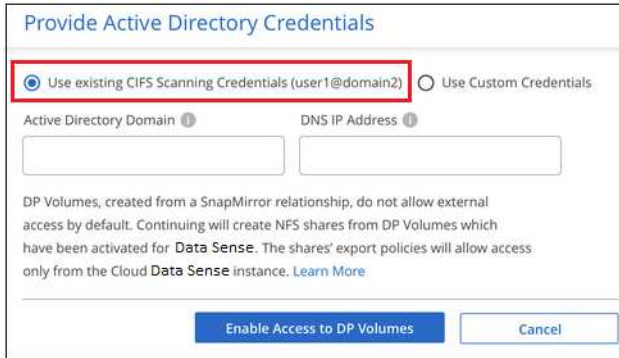
Enable Access to DP Volumes

Edit CIFS Credentials

Scan	Storage Repository (Volume)	Type	Status	Required Action
<div>OffMapMap & Classify</div>	VolumeName1	DP	Not Scanning	Enable access to DP Volumes ⓘ
<div>OffMapMap & Classify</div>	VolumeName2	NFS	Continuously Scanning	
<div>OffMapMap & Classify</div>	VolumeName3	CIFS	Not Scanning	

如果要扫描这些数据保护卷：

1. 单击页面顶部的 * 启用对 DP 卷的访问 *。
2. 查看确认消息，然后再次单击 * 启用对 DP 卷的访问 *。
 - 系统会启用最初在源 ONTAP 系统中创建为 NFS 卷的卷。
 - 最初在源 ONTAP 系统中创建为 CIFS 卷的卷需要输入 CIFS 凭据才能扫描这些 DP 卷。如果您已输入 Active Directory 凭据，以便 Cloud Data sense 可以扫描 CIFS 卷，则可以使用这些凭据，也可以指定一组不同的管理员凭据。



3. 激活要扫描的每个 DP 卷 [与启用其他卷的方式相同](#)。

启用后，Cloud Data sense 会从已激活进行扫描的每个 DP 卷创建一个 NFS 共享。共享导出策略仅允许从 Data sense 实例进行访问。

- 注意：* 如果在最初启用对 DP 卷的访问时没有 CIFS 数据保护卷，稍后再添加一些，则配置页面顶部会显示 * 启用对 CIFS DP* 的访问。单击此按钮并添加 CIFS 凭据，以便能够访问这些 CIFS DP 卷。



Active Directory 凭据仅在第一个 CIFS DP 卷的 Storage VM 中注册，因此将扫描该 SVM 上的所有 DP 卷。驻留在其他 SVM 上的任何卷都不会注册 Active Directory 凭据，因此不会扫描这些 DP 卷。

开始使用适用于 **Azure NetApp Files** 的云数据感知

完成几个步骤，开始使用适用于 Azure NetApp Files 的云数据感知。

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。

扫描 Azure NetApp Files 卷之前，["必须设置 Cloud Manager 才能发现配置"](#)。

["在 Cloud Manager 中部署 Cloud Data sense"](#) 如果尚未部署实例。

单击 * 合规性 *，选择 * 配置 * 选项卡，然后为特定工作环境中的卷激活合规性扫描。

启用 Cloud Data sense 后，请确保它可以访问所有卷。

- 云数据感知实例需要与每个 Azure NetApp Files 子网建立网络连接。
- 确保这些端口对 Data sense 实例开放：
 - 对于 NFS —端口 111 和 2049 。
 - 对于 CIFS —端口 139 和 445 。
- NFS 卷导出策略必须允许从 Data sense 实例进行访问。
- Data sense 需要 Active Directory 凭据才能扫描 CIFS 卷。

单击 * 合规性 * > * 配置 * > * 编辑 CIFS 凭据 * 并提供凭据。

选择或取消选择要扫描的卷， Cloud Data sense 将开始或停止扫描这些卷。

正在发现要扫描的 **Azure NetApp Files** 系统

如果您要扫描的 Azure NetApp Files 系统尚未作为工作环境在 Cloud Manager 中，您可以此时将其添加到画布中。

["了解如何在 Cloud Manager 中发现 Azure NetApp Files 系统"](#)。

部署 **Cloud Data sense** 实例

["部署 Cloud Data sense"](#) 如果尚未部署实例。

扫描 Azure NetApp Files 卷时，必须在云中部署数据感知，并且数据感知必须与要扫描的卷部署在同一区域。

- 注意： * 扫描 Azure NetApp Files 卷时，当前不支持在内部位置部署云数据感知。

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

在您的工作环境中实现云数据感知

您可以在 Azure NetApp Files 卷上启用云数据感知。

1. 从Cloud Manager左侧导航菜单中、单击*数据感知*、然后选择*配置*选项卡。



2. 选择要如何扫描每个工作环境中的卷。 ["了解映射和分类扫描"](#)：
 - 要映射所有卷，请单击 * 映射所有卷 *。
 - 要映射所有卷并对其进行分类，请单击 * 映射并分类所有卷 *。

- 要自定义每个卷的扫描，请单击 * 或选择每个卷的扫描类型 *，然后选择要映射和 / 或分类的卷。

请参见 [在卷上启用和禁用合规性扫描](#) 了解详细信息。

3. 在确认对话框中，单击 * 批准 * 以使 Data sense 开始扫描卷。

Cloud Data sense 开始扫描您在工作环境中选择的卷。一旦 Cloud Data sense 完成初始扫描，" 合规性 " 信息板将显示结果。所需时间取决于数据量—可能需要几分钟或几小时。

验证 **Cloud Data sense** 是否有权访问卷

通过检查网络，安全组和导出策略，确保 Cloud Data sense 可以访问卷。您需要为 Data sense 提供 CIFS 凭据，以便它可以访问 CIFS 卷。

步骤

1. 确保云数据感知实例与包含 Azure NetApp Files 卷的每个网络之间存在网络连接。



对于 Azure NetApp Files，Cloud Data sense 只能扫描与 Cloud Manager 位于同一区域的卷。

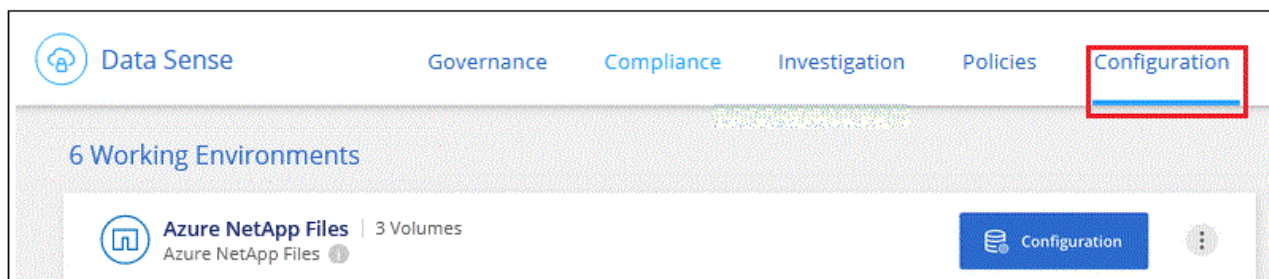
2. 确保以下端口对 Data sense 实例开放：

- 对于 NFS —端口 111 和 2049。
- 对于 CIFS —端口 139 和 445。

3. 确保 NFS 卷导出策略包含 Data sense 实例的 IP 地址，以便它可以访问每个卷上的数据。

4. 如果使用 CIFS，请为 Data sense 提供 Active Directory 凭据，以便它可以扫描 CIFS 卷。

- a. 在 Cloud Manager 顶部，单击 * 数据感知 *。
- b. 单击 * 配置 * 选项卡。

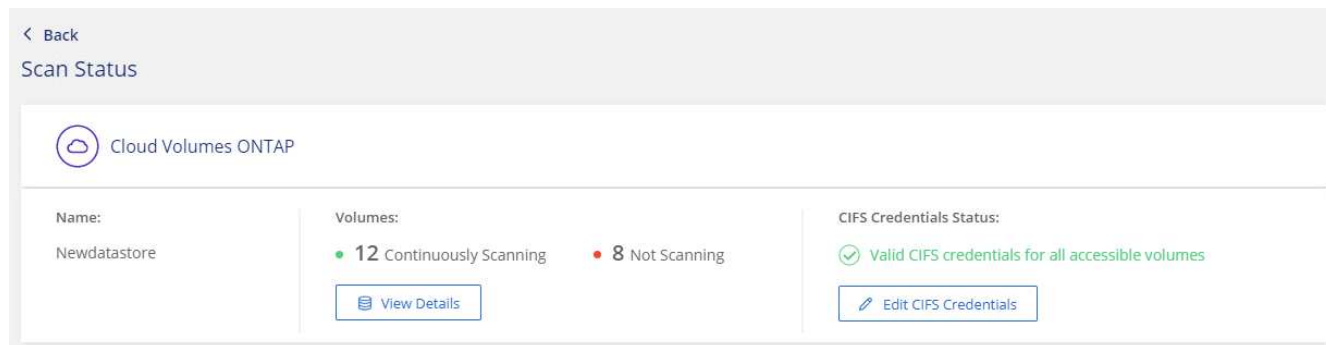


- c. 对于每个工作环境，单击 * 编辑 CIFS 凭据 *，然后输入 Data sense 访问系统上 CIFS 卷所需的用户名和密码。

这些凭据可以是只读的，但提供管理员凭据可确保 Data sense 可以读取任何需要提升权限的数据。这些凭据存储在 Cloud Data sense 实例上。

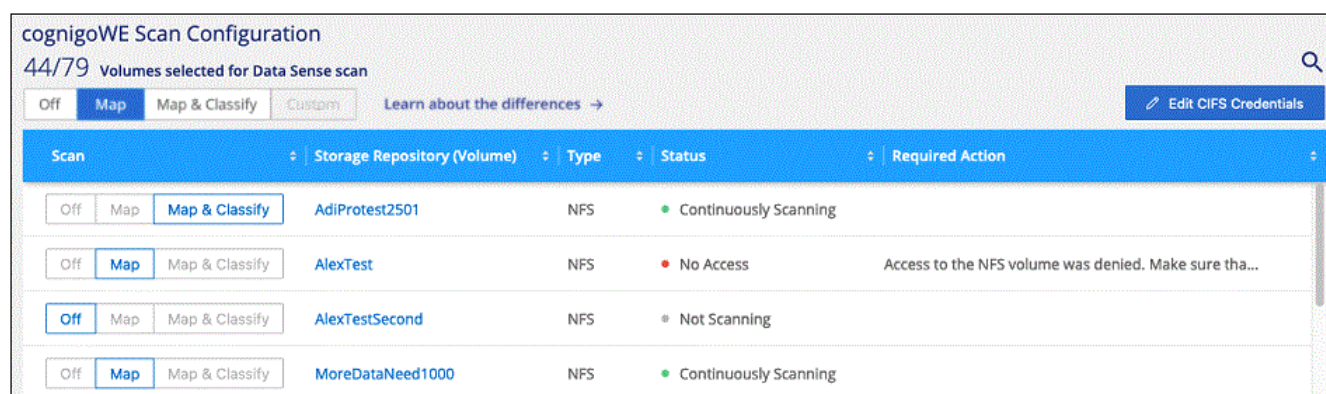
如果要确保数据感知分类扫描不会更改文件的"上次访问时间"、我们建议用户具有"写入属性"权限。如果可能、我们建议将Active Directory配置的用户设置为组织中有权访问所有文件的父组的一部分。

输入凭据后，您应看到一条消息，指出所有 CIFS 卷均已成功通过身份验证。



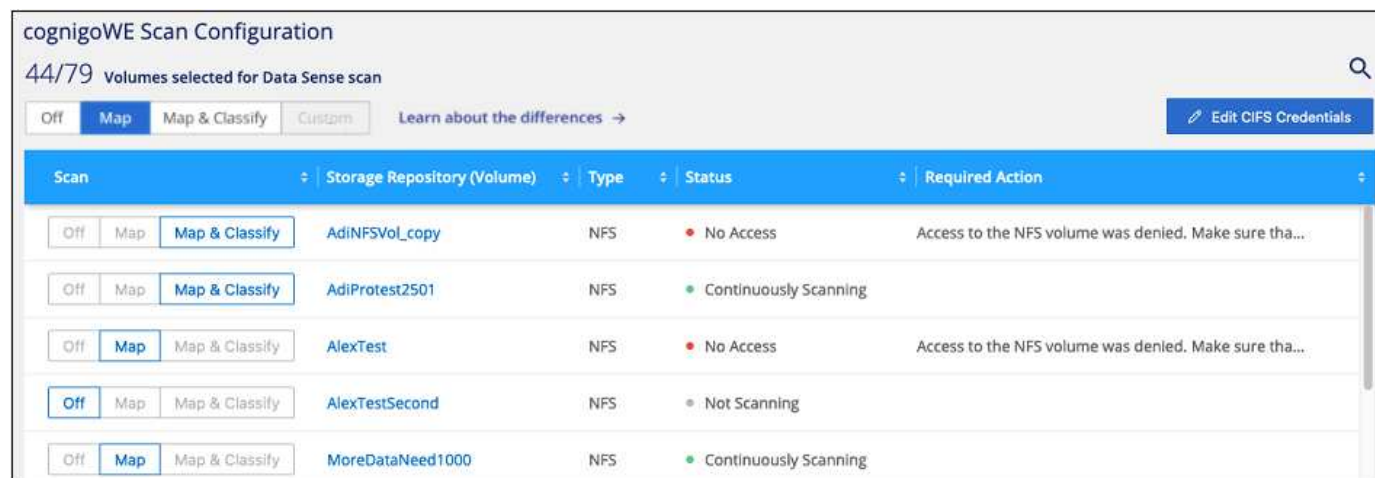
5. 在 *Configuration* 页面上，单击 * 查看详细信息 * 以查看每个 CIFS 和 NFS 卷的状态并更正任何错误。

例如，下图显示了四个卷；其中一个卷由于 Data sense 实例与卷之间的网络连接问题而无法扫描。



在卷上启用和禁用合规性扫描

您可以随时从 " 配置 " 页面在工作环境中启动或停止仅映射扫描或映射和分类扫描。您也可以从仅映射扫描更改为映射和分类扫描，反之亦然。建议您扫描所有卷。



收件人：	执行以下操作：
在卷上启用仅映射扫描	在卷区域中，单击 * 映射 *
对卷启用完全扫描	在卷区域中，单击 * 映射和分类 *
禁用对卷的扫描	在卷区域中，单击 * 关闭 *

收件人：	执行以下操作：
在所有卷上启用仅映射扫描	在标题区域中，单击 * 映射 *
对所有卷启用完全扫描	在标题区域中，单击 * 映射和分类 *
禁用对所有卷的扫描	在标题区域中，单击 * 关闭 *



只有在标题区域中设置了 * 映射 * 或 * 映射和分类 * 设置后，才会自动扫描添加到工作环境中的新卷。如果在标题区域中设置为 * 自定义 * 或 * 关闭 *，则需要在工作环境中添加的每个新卷上激活映射和 / 或完全扫描。

开始使用适用于 Amazon FSX for ONTAP 的 Cloud Data sense

请完成几个步骤，开始使用 Cloud Data sense 扫描 Amazon FSX for ONTAP 卷。

开始之前

- 您需要在 AWS 中使用主动连接器来部署和管理 Data sense。
- 您在创建工作环境时选择的安全组必须允许来自云数据感知实例的流量。您可以使用连接到 FSX for ONTAP 文件系统的 ENI 来查找关联的安全组，并使用 AWS 管理控制台对其进行编辑。

["适用于 Linux 实例的 AWS 安全组"](#)

["适用于 Windows 实例的 AWS 安全组"](#)

["AWS 弹性网络接口（ENI）"](#)

快速入门

按照以下步骤快速入门，或者向下滚动以查看完整详细信息。

在扫描 ONTAP 卷的 FSX 之前，["您必须具有配置了卷的 FSX 工作环境"](#)。

["在 Cloud Manager 中部署 Cloud Data sense"](#) 如果尚未部署实例。

单击 * 数据感知 *，选择 * 配置 * 选项卡，然后为特定工作环境中的卷激活合规性扫描。

启用 Cloud Data sense 后，请确保它可以访问所有卷。

- 云数据感知实例需要与 ONTAP 子网的每个 FSX 建立网络连接。
- 确保以下端口已对 Data sense 实例开放。
 - 对于 NFS —端口 111 和 2049。
 - 对于 CIFS —端口 139 和 445。
- NFS 卷导出策略必须允许从 Data sense 实例进行访问。
- Data sense 需要 Active Directory 凭据才能扫描 CIFS 卷。+ 单击 * 合规性 * > * 配置 * > * 编辑 CIFS 凭据 * 并提供凭据。

选择或取消选择要扫描的卷， Cloud Data sense 将开始或停止扫描这些卷。

正在发现要扫描的 **FSX for ONTAP** 文件系统

如果您要扫描的适用于 ONTAP 的 FSX 文件系统尚未作为工作环境在 Cloud Manager 中，则可以此时将其添加到画布中。

["了解如何在 Cloud Manager 中发现或创建适用于 ONTAP 的 FSX 文件系统"](#)。

部署 **Cloud Data sense** 实例

["部署 Cloud Data sense"](#) 如果尚未部署实例。

您应将 Data sense 部署在与 Connector for AWS 和要扫描的 FSX 卷相同的 AWS 网络中。

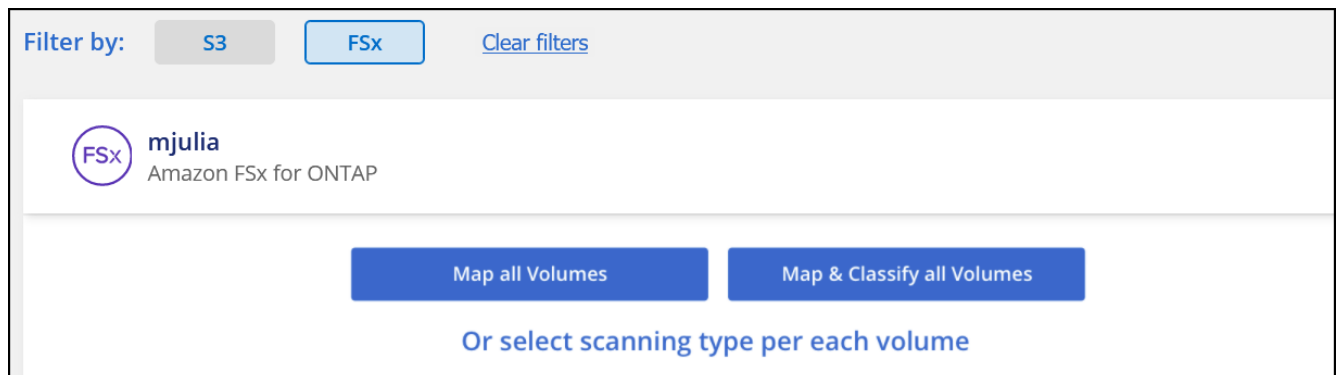
- 注意： * 扫描 FSX 卷时，当前不支持在内部位置部署 Cloud Data sense 。

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

在您的工作环境中实现云数据感知

您可以为 ONTAP 卷的 FSX 启用云数据感知。

1. 从Cloud Manager左侧导航菜单中、单击*数据感知*、然后选择*配置*选项卡。



2. 选择要如何扫描每个工作环境中的卷。 ["了解映射和分类扫描"](#):
 - 要映射所有卷，请单击 * 映射所有卷 *。
 - 要映射所有卷并对其进行分类，请单击 * 映射并分类所有卷 *。
 - 要自定义每个卷的扫描，请单击 * 或选择每个卷的扫描类型 *，然后选择要映射和 / 或分类的卷。

请参见 [在卷上启用和禁用合规性扫描](#) 了解详细信息。

3. 在确认对话框中，单击 * 批准 * 以使 Data sense 开始扫描卷。

Cloud Data sense 开始扫描您在工作环境中选择的卷。一旦 Cloud Data sense 完成初始扫描， " 合规性 " 信息板将显示结果。所需时间取决于数据量—可能需要几分钟或几小时。

验证 **Cloud Data sense** 是否有权访问卷

通过检查网络，安全组和导出策略，确保 Cloud Data sense 可以访问卷。

您需要为 Data sense 提供 CIFS 凭据，以便它可以访问 CIFS 卷。

步骤

1. 在 *Configuration* 页面上，单击 * 查看详细信息 * 以查看状态并更正任何错误。

例如，下图显示了由于 Data sense 实例与卷之间的网络连接问题，卷 Cloud Data sense 无法扫描。

Scan	Storage Repository (Volume)	Type	Status	Required Action
Off Map Map & Classify	jrmclone	NFS	No Access	Check network connectivity between the Data Sense ...

2. 确保云数据感知实例与包含适用于 ONTAP 的 FSX 卷的每个网络之间存在网络连接。



对于适用于 ONTAP 的 FSX，Cloud Data sense 只能扫描与 Cloud Manager 位于同一区域的卷。

3. 确保以下端口对 Data sense 实例开放：
 - 对于 NFS —端口 111 和 2049。
 - 对于 CIFS —端口 139 和 445。
4. 确保 NFS 卷导出策略包含 Data sense 实例的 IP 地址，以便它可以访问每个卷上的数据。
5. 如果使用 CIFS，请为 Data sense 提供 Active Directory 凭据，以便它可以扫描 CIFS 卷。
 - a. 在 Cloud Manager 顶部，单击 * 数据感知 *。
 - b. 单击 * 配置 * 选项卡。
 - c. 对于每个工作环境，单击 * 编辑 CIFS 凭据 *，然后输入 Data sense 访问系统上 CIFS 卷所需的用户名和密码。

这些凭据可以是只读的，但提供管理员凭据可确保 Data sense 可以读取任何需要提升权限的数据。这些凭据存储在 Cloud Data sense 实例上。

如果要确保数据感知分类扫描不会更改文件的"上次访问时间"、我们建议用户具有"写入属性"权限。如果可能、我们建议将Active Directory配置的用户设置为组织中有权访问所有文件的父组的一部分。

输入凭据后，您应看到一条消息，指出所有 CIFS 卷均已成功通过身份验证。

在卷上启用和禁用合规性扫描

您可以随时从 " 配置 " 页面在工作环境中启动或停止仅映射扫描或映射和分类扫描。您也可以从仅映射扫描更改为映射和分类扫描，反之亦然。建议您扫描所有卷。

cognigoWE Scan Configuration					
44/79 Volumes selected for Data Sense scan					
<div> Off Map Map & Classify Custom Learn about the differences → </div> <div>Edit CIFS Credentials</div>					
Scan	Storage Repository (Volume)	Type	Status	Required Action	
Off Map Map & Classify	AdiNFSVol_copy	NFS	No Access	Access to the NFS volume was denied. Make sure tha...	
Off Map Map & Classify	AdiProtest2501	NFS	Continuously Scanning		
Off Map Map & Classify	AlexTest	NFS	No Access	Access to the NFS volume was denied. Make sure tha...	
Off Map Map & Classify	AlexTestSecond	NFS	Not Scanning		
Off Map Map & Classify	MoreDataNeed1000	NFS	Continuously Scanning		

收件人：	执行以下操作：
在卷上启用仅映射扫描	在卷区域中，单击 * 映射 *
对卷启用完全扫描	在卷区域中，单击 * 映射和分类 *
禁用对卷的扫描	在卷区域中，单击 * 关闭 *
在所有卷上启用仅映射扫描	在标题区域中，单击 * 映射 *
对所有卷启用完全扫描	在标题区域中，单击 * 映射和分类 *
禁用对所有卷的扫描	在标题区域中，单击 * 关闭 *



只有在标题区域中设置了 * 映射 * 或 * 映射和分类 * 设置后，才会自动扫描添加到工作环境中的新卷。如果在标题区域中设置为 * 自定义 * 或 * 关闭 * ，则需要在工作环境中添加的每个新卷上激活映射和 / 或完全扫描。

扫描数据保护卷

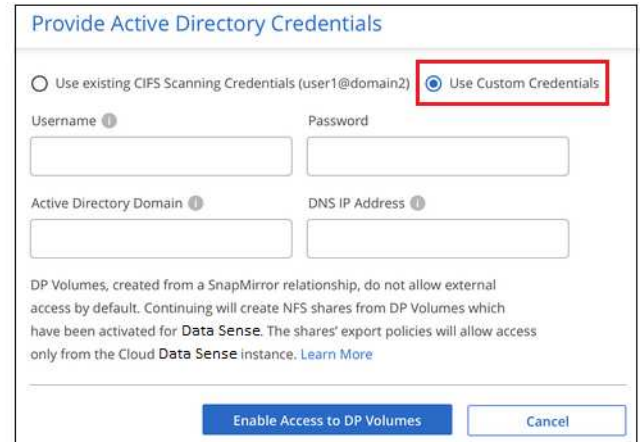
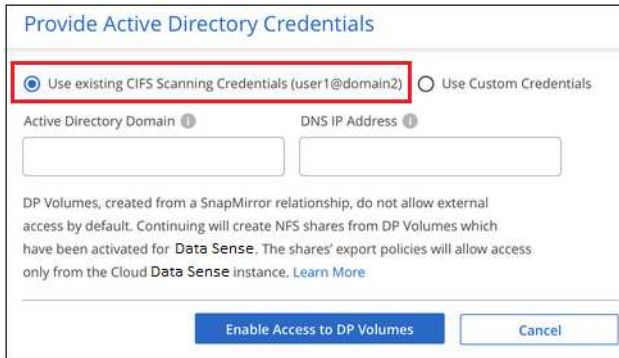
默认情况下，不会扫描数据保护（DP）卷，因为它们不会公开在外部，并且 Cloud Data sense 无法访问它们。这些卷是从适用于 ONTAP 的 FSX 文件系统执行 SnapMirror 操作的目标卷。

最初，卷列表会将这些卷标识为 *Type* * dp* ，并显示 *Status* * 未扫描 * 和 *Required Action* * Enable Access to DP volumes* 。

'Working Environment Name' Configuration					
22/28 Volumes selected for compliance scan					
<div> Off Map Map & Classify Custom Learn about the differences → </div> <div>Enable Access to DP Volumes</div> <div>Edit CIFS Credentials</div>					
Scan	Storage Repository (Volume)	Type	Status	Required Action	
Off Map Map & Classify	VolumeName1	DP	Not Scanning	Enable access to DP Volumes ⓘ	
Off Map Map & Classify	VolumeName2	NFS	Continuously Scanning		
Off Map Map & Classify	VolumeName3	CIFS	Not Scanning		

如果要扫描这些数据保护卷：

1. 单击页面顶部的 * 启用对 DP 卷的访问 *。
2. 查看确认消息，然后再次单击 * 启用对 DP 卷的访问 *。
 - 系统将启用最初在源 FSX for ONTAP 文件系统中创建为 NFS 卷的卷。
 - 最初在源 FSX for ONTAP 文件系统中创建为 CIFS 卷的卷需要输入 CIFS 凭据才能扫描这些 DP 卷。如果您已输入 Active Directory 凭据，以便 Cloud Data sense 可以扫描 CIFS 卷，则可以使用这些凭据，也可以指定一组不同的管理员凭据。



3. 激活要扫描的每个 DP 卷 与启用其他卷的方式相同。

启用后，Cloud Data sense 会从已激活进行扫描的每个 DP 卷创建一个 NFS 共享。共享导出策略仅允许从 Data sense 实例进行访问。

- 注意：* 如果在最初启用对 DP 卷的访问时没有 CIFS 数据保护卷，稍后再添加一些，则配置页面顶部会显示 * 启用对 CIFS DP* 的访问。单击此按钮并添加 CIFS 凭据，以便能够访问这些 CIFS DP 卷。



Active Directory 凭据仅在第一个 CIFS DP 卷的 Storage VM 中注册，因此将扫描该 SVM 上的所有 DP 卷。驻留在其他 SVM 上的任何卷都不会注册 Active Directory 凭据，因此不会扫描这些 DP 卷。

Amazon S3 云数据感知入门

Cloud Data sense 可以扫描 Amazon S3 存储分段，以确定 S3 对象存储中的个人和敏感数据。Cloud Data sense 可以扫描帐户中的任何存储分段，而不管它是否是为 NetApp 解决方案创建的。

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。

确保您的云环境能够满足 Cloud Data sense 的要求，包括准备 IAM 角色以及设置从 Data sense 到 S3 的连接。请参见完整列表。

"部署 Cloud Data sense" 如果尚未部署实例。

选择 Amazon S3 工作环境，单击 * 启用 *，然后选择包含所需权限的 IAM 角色。

选择要扫描的存储分段， Cloud Data sense 将开始扫描这些存储分段。

查看 **S3** 前提条件

以下要求特定于扫描 S3 存储分段。

为云数据感知实例设置 **IAM** 角色

Cloud Data sense 需要获得连接到您帐户中的 S3 存储分段并对其进行扫描的权限。设置一个包含以下权限的 IAM 角色。在 Amazon S3 工作环境中启用 Data sense 时， Cloud Manager 会提示您选择 IAM 角色。

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

提供从 **Cloud Data sense** 到 **Amazon S3** 的连接

Cloud Data sense 需要连接到 Amazon S3。提供此连接的最佳方式是通过 VPC 端点连接到 S3 服务。有关说明，请参见 ["AWS 文档：创建网关端点"](#)。

创建 VPC 端点时，请务必选择与云数据感知实例对应的区域，VPC 和路由表。您还必须修改安全组才能添加出站 HTTPS 规则、该规则允许通信到 S3 端点。否则，Data sense 将无法连接到 S3 服务。

如果遇到任何问题，请参见 ["AWS 支持知识中心：为什么我无法使用网关 VPC 端点连接到 S3 存储分段？"](#)

另一种方法是使用 NAT 网关提供连接。



您无法使用代理通过 Internet 访问 S3。

部署 Cloud Data sense 实例

"在 Cloud Manager 中部署 Cloud Data sense" 如果尚未部署实例。

您需要使用部署在 AWS 中的 Connector 部署此实例，以便 Cloud Manager 自动发现此 AWS 帐户中的 S3 存储分段并将其显示在 Amazon S3 工作环境中。

- 注意：* 扫描 S3 存储分段时，当前不支持在内部位置部署 Cloud Data sense。

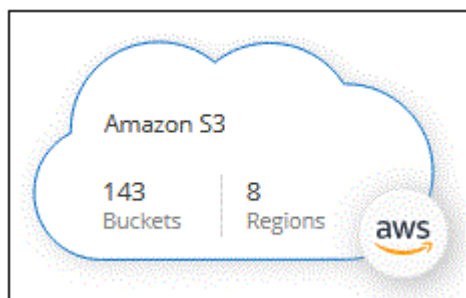
只要实例具有 Internet 连接，就会自动升级到 Data sense 软件。

在 S3 工作环境中激活 Data sense

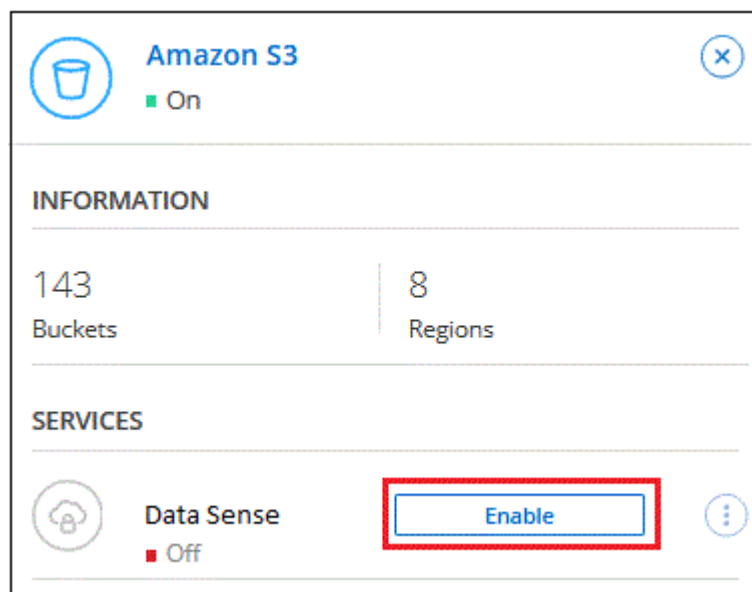
验证前提条件后，在 Amazon S3 上启用 Cloud Data sense。

步骤

1. 从 Cloud Manager 左侧导航菜单中、单击*画布*。
2. 选择 Amazon S3 工作环境。



3. 在右侧的数据感知窗格中，单击 * 启用 *。



4. 出现提示时，将 IAM 角色分配给具有 Cloud Data sense 实例 [所需权限](#)。

Assign an AWS IAM Role for Cloud Data Sense

To enable **Cloud Data Sense** on Amazon S3 buckets, select an existing IAM Role. Make sure that your AWS IAM Role has the permission defined in the [Policy Requirements](#).

Select IAM Role

occm

VPC Endpoint for Amazon S3 Required

A VPC endpoint to the Amazon S3 service is required so **Data Sense** can securely scan the data.

Alternatively, ensure that the **Data Sense** instance has direct access to the internet via a NAT Gateway or Internet Gateway.

Free for the 1st TB

Over 1 TB you pay only for what you use. [Learn more about pricing.](#)

Enable

Cancel

5. 单击 * 启用 *。



您也可以通过单击在配置页面中为工作环境启用合规性扫描  按钮并选择 * 激活数据感知 *。

Cloud Manager 将 IAM 角色分配给实例。

在 **S3** 存储分段上启用和禁用合规性扫描

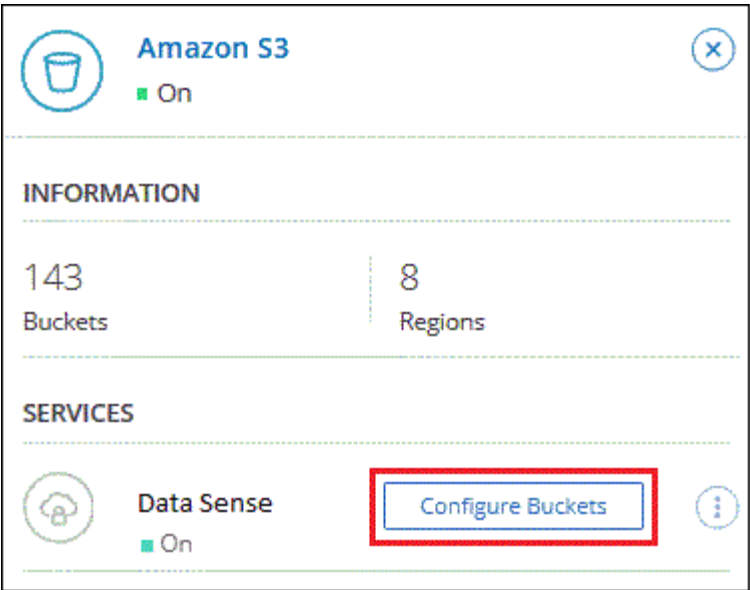
在 Cloud Manager 在 Amazon S3 上启用 Cloud Data sense 后，下一步是配置要扫描的分段。

如果 Cloud Manager 运行在包含要扫描的 S3 存储分段的 AWS 帐户中，则它会发现这些存储分段并将其显示在 Amazon S3 工作环境中。

云数据感知也可以 [扫描位于不同 AWS 帐户中的 S3 存储分段](#)。

步骤

1. 选择 Amazon S3 工作环境。
2. 在右侧窗格中，单击 * 配置分段 *。



3. 在存储分段上启用仅映射扫描或映射和分类扫描。

Amazon S3 Configuration			
15/28 Buckets in Scan Scope.			
Scan	Bucket Name	Status	Required Action
<div>OffMapMap & Classify</div>	BucketName1	● Not Scanning	Add Credentials
<div>OffMapMap & Classify</div>	BucketName2	● Continuously Scanning	
<div>OffMapMap & Classify</div>	BucketName3	● Not Scanning	

收件人：	执行以下操作：
在存储分段上启用仅映射扫描	单击 * 映射 *
对存储分段启用完全扫描	单击 * 映射和分类 *
禁用对存储分段的扫描	单击 * 关闭 *

Cloud Data sense 开始扫描您启用的 S3 存储分段。如果存在任何错误，它们将显示在状态列中，并显示修复此错误所需的操作。

从其他 **AWS** 帐户扫描存储分段

您可以通过从其他 AWS 帐户中分配角色来扫描此帐户下的 S3 存储分段，以访问现有 Cloud Data sense 实例。

步骤

1. 转到要扫描 S3 存储分段的目标 AWS 帐户，然后选择 * 其他 AWS 帐户 * 来创建 IAM 角色。

Create role





1

2

3

4


Select type of trusted entity

 AWS service EC2, Lambda and others	 Another AWS account Belonging to you or 3rd party	 Web identity Cognito or any OpenID provider	 SAML 2.0 federation Your corporate directory
--------------------------------------------------------------------------------------------------------------------------------	-----------------------------------------------------------------------------------------------------------------------------------------------	-----------------------------------------------------------------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------------------------------------------------------

Allows entities in other accounts to perform actions in this account. [Learn more](#)

Specify accounts that can use this role

Account ID*

- Options**
- ☐ Require external ID (Best practice when a third party will assume this role)
 - ☐ Require MFA 

请务必执行以下操作：

- 输入 Cloud Data sense 实例所在帐户的 ID。
- 将 * 最大 CLI/API 会话持续时间 * 从 1 小时更改为 12 小时，然后保存此更改。
- 附加云数据感知 IAM 策略。确保它具有所需的权限。

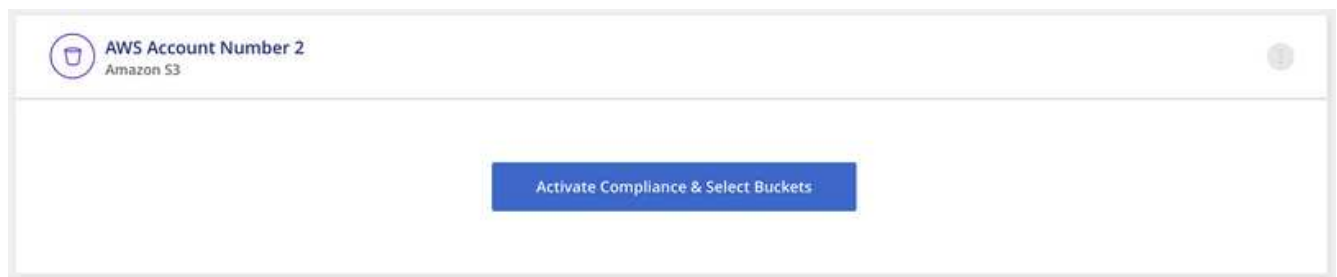
```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "s3:Get*",
        "s3:List*",
        "s3:PutObject"
      ],
      "Resource": "*"
    }
  ]
}
```

2. 转到 Data sense 实例所在的源 AWS 帐户，然后选择附加到该实例的 IAM 角色。
 - a. 将 * 最大 CLI/API 会话持续时间 * 从 1 小时更改为 12 小时，然后保存此更改。
 - b. 单击 * 附加策略 *，然后单击 * 创建策略 *。
 - c. 创建一个包含 "STS : AssumeRole" 操作的策略，并指定您在目标帐户中创建的角色 ARN。

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": "sts:AssumeRole",
      "Resource": "arn:aws:iam::<ADDITIONAL-ACCOUNT-ID>:role/<ADDITIONAL_ROLE_NAME>"
    },
    {
      "Effect": "Allow",
      "Action": [
        "iam:GetPolicyVersion",
        "iam:GetPolicy",
        "iam:ListAttachedRolePolicies"
      ],
      "Resource": [
        "arn:aws:iam::*:policy/*",
        "arn:aws:iam::*:role/*"
      ]
    }
  ]
}
```

Cloud Data sense 实例配置文件帐户现在可以访问其他 AWS 帐户。

3. 转到 * Amazon S3 Configuration* 页面，此时将显示新的 AWS 帐户。请注意， Cloud Data sense 可能需要几分钟时间来同步新帐户的工作环境并显示此信息。



4. 单击 * 激活数据感知并选择分段 *，然后选择要扫描的分段。

Cloud Data sense 将开始扫描您启用的新 S3 存储分段。

正在扫描数据库架构

完成几个步骤，开始使用 Cloud Data sense 扫描数据库架构。

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。

确保您的数据库受支持，并且您具有连接到数据库所需的信息。

"部署 Cloud Data sense" 如果尚未部署实例。

添加要访问的数据库服务器。

选择要扫描的模式。

查看前提条件

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

支持的数据库

Cloud Data sense 可以从以下数据库扫描架构：

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



数据库中必须启用 * 统计信息收集功能。

数据库要求

可以扫描连接到云数据感知实例的任何数据库，而不管其托管在何处。要连接到数据库，您只需提供以下信息：

- IP 地址或主机名
- Port
- 服务名称（仅用于访问 Oracle 数据库）
- 允许对模式进行读取访问的凭据

选择用户名和密码时，请务必选择对要扫描的所有架构和表具有完全读取权限的用户名和密码。建议您为 Cloud Data sense 系统创建一个具有所有所需权限的专用用户。

- 注：* 对于 MongoDB，需要只读管理员角色。

部署 Cloud Data sense 实例

如果尚未部署实例，请部署 Cloud Data sense。

如果要扫描可通过 Internet 访问的数据库架构，则可以 "在云中部署 Cloud Data sense" 或 "在可访问 Internet 的内部位置部署 Data sense"。

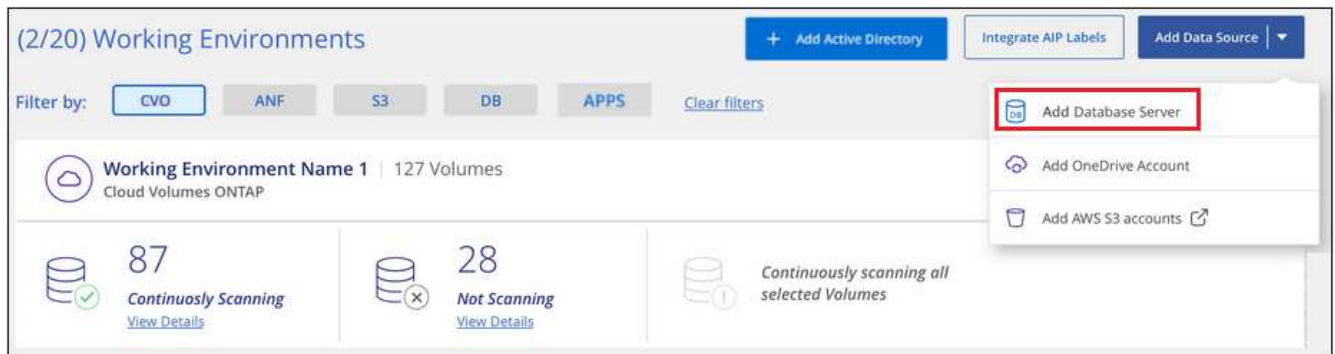
如果要扫描的数据库架构安装在无法访问 Internet 的非公开站点中，则需要执行 "在无法访问 Internet 的同一内部位置部署 Cloud Data sense"。这还要求 Cloud Manager Connector 部署在同一内部位置。

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

正在添加数据库服务器

添加架构所在的数据库服务器。

1. 在工作环境配置页面中，单击 * 添加数据源 * > * 添加数据库服务器 * 。



2. 输入所需信息以标识数据库服务器。
 - a. 选择数据库类型。
 - b. 输入要连接到数据库的端口和主机名或 IP 地址。
 - c. 对于 Oracle 数据库，输入服务名称。
 - d. 输入凭据，以便 Cloud Data sense 可以访问服务器。
 - e. 单击 * 添加数据库服务器 * 。

Add DB Server

To activate Compliance on Databases, first add a Database Server. After this step, you'll be able to select which Database Schemas you would like to activate Compliance for.

Database

Database Type	Host Name or IP Address
<input type="text"/>	<input type="text"/>
Port	Service Name
<input type="text"/>	<input type="text"/>

Credentials

Username	Password
<input type="text"/>	<input type="text"/>

Add DB ServerCancel

数据库将添加到工作环境列表中。

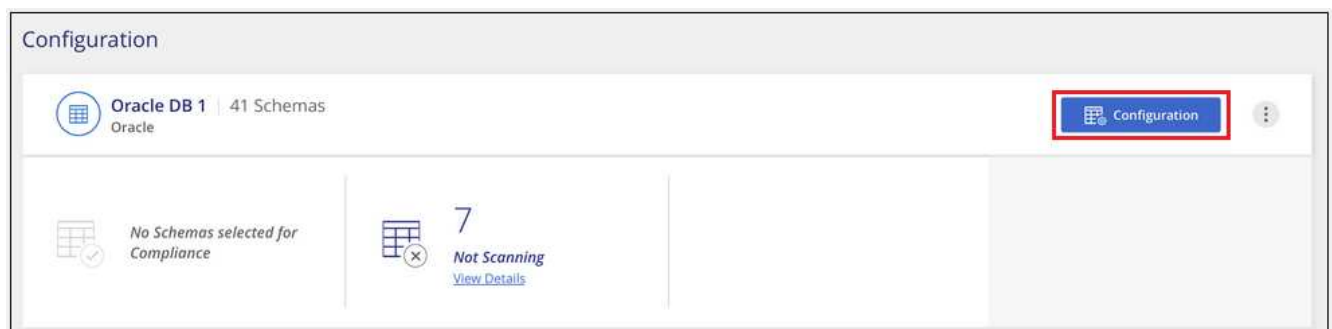
对数据库架构启用和禁用合规性扫描

您可以随时停止或开始对架构进行完全扫描。



没有为数据库架构选择仅映射扫描的选项。

1. 在 *Configuration* 页面中，单击要配置的数据库的 * 配置 * 按钮。



2. 向右移动滑块以选择要扫描的架构。

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

Cloud Data sense 将开始扫描您启用的数据库架构。如果存在任何错误，它们将显示在状态列中，并显示修复此错误所需的操作。

正在扫描 OneDrive 帐户

完成几个步骤，使用 Cloud Data sense 扫描用户的 OneDrive 文件夹中的文件。

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。

确保您拥有登录到 OneDrive 帐户所需的管理员凭据。

["部署 Cloud Data sense"](#) 如果尚未部署实例。

使用管理员用户凭据登录到要访问的 OneDrive 帐户，以便将其添加为新的工作环境。

从 OneDrive 帐户中添加要扫描的用户列表，然后选择扫描类型。一次最多可以添加 100 个用户。

查看 OneDrive 要求

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

- 您必须具有 OneDrive for Business 帐户的管理员登录凭据、该帐户可提供对用户文件的读取访问权限。
- 您需要列出要扫描其 OneDrive 文件夹的所有用户的电子邮件地址、并以行分隔。

部署 Cloud Data sense 实例

如果尚未部署实例，请部署 Cloud Data sense 。

数据感知可以是 ["部署在云中"](#) 或 ["位于可访问 Internet 的内部位置"](#)。

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

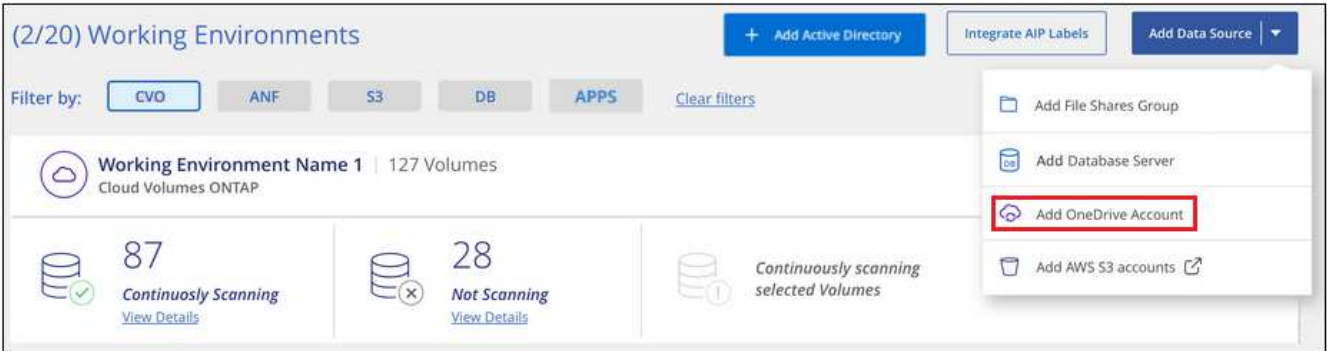
数据感知也可以是 ["部署在无法访问 Internet 的内部位置"](#)。但是，要扫描本地 OneDrive 文件，您需要为一些选定端点提供 Internet 访问。 ["请在此处查看所需端点的列表"](#)。

正在添加 **OneDrive** 帐户

添加用户文件所在的 OneDrive 帐户。

步骤

1. 在工作环境配置页面中，单击 * 添加数据源 * > * 添加 OneDrive 帐户 *。



2. 在添加 OneDrive 帐户对话框中，单击 * 登录到 OneDrive*。
3. 在显示的 Microsoft 页面中，选择 OneDrive 帐户并输入所需的管理员用户和密码，然后单击 * 接受 * 以允许 Cloud Data sense 从此帐户读取数据。

OneDrive 帐户将添加到工作环境列表中。

将 **OneDrive** 用户添加到合规性扫描

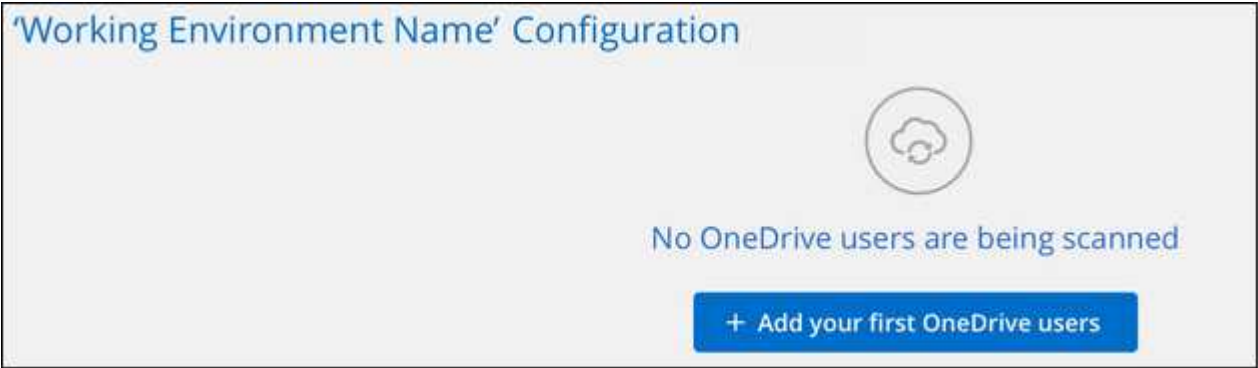
您可以添加单个 OneDrive 用户或所有 OneDrive 用户，以便 Cloud Data sense 对其文件进行扫描。

步骤

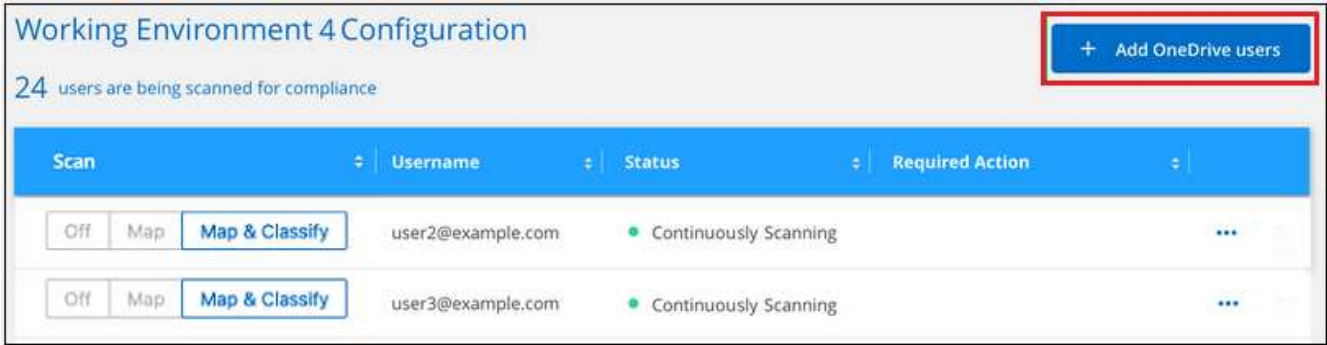
1. 在 *Configuration* 页面中，单击 OneDrive 帐户的 * 配置 * 按钮。



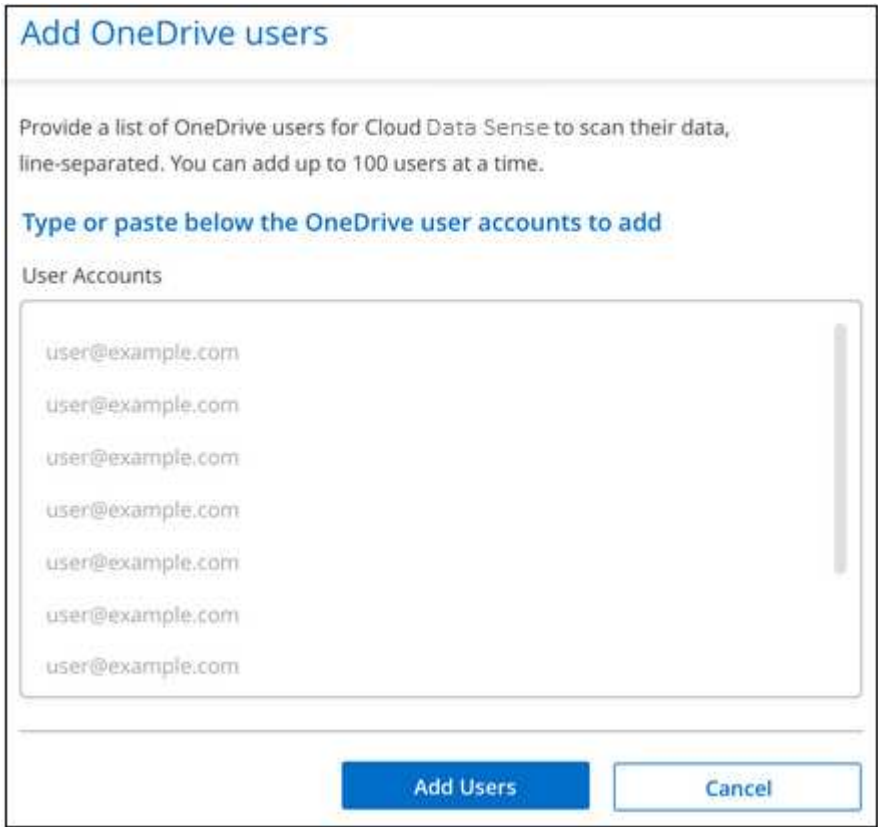
2. 如果这是首次为此 OneDrive 帐户添加用户，请单击 * 添加您的首个 OneDrive 用户 *。



如果要从 OneDrive 帐户添加其他用户，请单击 * 添加 OneDrive 用户 *。



3. 为要扫描其文件的用户添加电子邮件地址 - 每行一个电子邮件地址（每个会话最多 100 个） - 然后单击 * 添加用户 *。



确认对话框将显示已添加的用户数。

如果此对话框列出了任何无法添加的用户，请捕获此信息，以便解析问题描述。在某些情况下，您可以使用更正后的电子邮件地址重新添加用户。

4. 对用户文件启用仅映射扫描或映射和分类扫描。

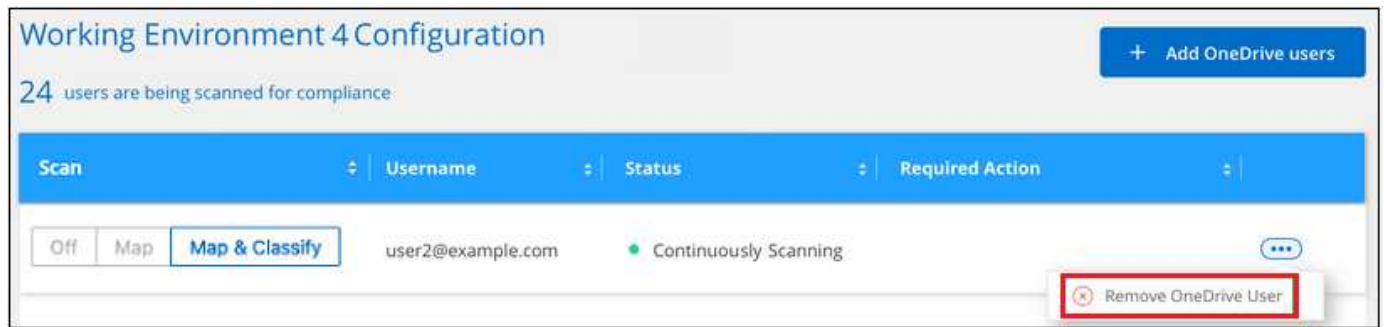
收件人：	执行以下操作：
对用户文件启用仅映射扫描	单击 * 映射 *
对用户文件启用完全扫描	单击 * 映射和分类 *

收件人：	执行以下操作：
禁用对用户文件的扫描	单击 * 关闭 *

Cloud Data sense 开始扫描您添加的用户文件，结果将显示在信息板和其他位置。

从合规性扫描中删除 OneDrive 用户

如果用户离开公司或其电子邮件地址发生变化，您可以随时删除单个 OneDrive 用户的文件扫描功能。只需从配置页面中单击 * 删除 OneDrive 用户 * 即可。



请注意，您可以 ["从Data sense中删除整个OneDrive帐户"](#) 如果您不想再扫描OneDrive帐户中的任何用户数据。

扫描 SharePoint 帐户

完成几个步骤，使用 Cloud Data sense 扫描 SharePoint 帐户中的文件。

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。

请确保您拥有用于登录到 SharePoint 帐户的管理员凭据，并且拥有要扫描的 SharePoint 站点的 URL。

["部署 Cloud Data sense"](#) 如果尚未部署实例。

使用管理员用户凭据登录到要访问的 SharePoint 帐户，以便将其添加为新的数据源 / 工作环境。

在 SharePoint 帐户中添加要扫描的 SharePoint 站点 URL 列表，然后选择扫描类型。一次最多可以添加 100 个 URL。

查看 SharePoint 要求

请查看以下前提条件，以确保您已准备好在 SharePoint 帐户上启用 Cloud Data sense。

- 您必须具有可对所有 SharePoint 站点进行读取访问的 SharePoint 帐户的管理员登录凭据。
- 对于要扫描的所有数据，您需要一个以行分隔的 SharePoint 站点 URL 列表。

部署 Cloud Data sense 实例

如果尚未部署实例，请部署 Cloud Data sense。

数据感知可以是 "部署在云中" 或 "位于可访问 Internet 的内部位置"。

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

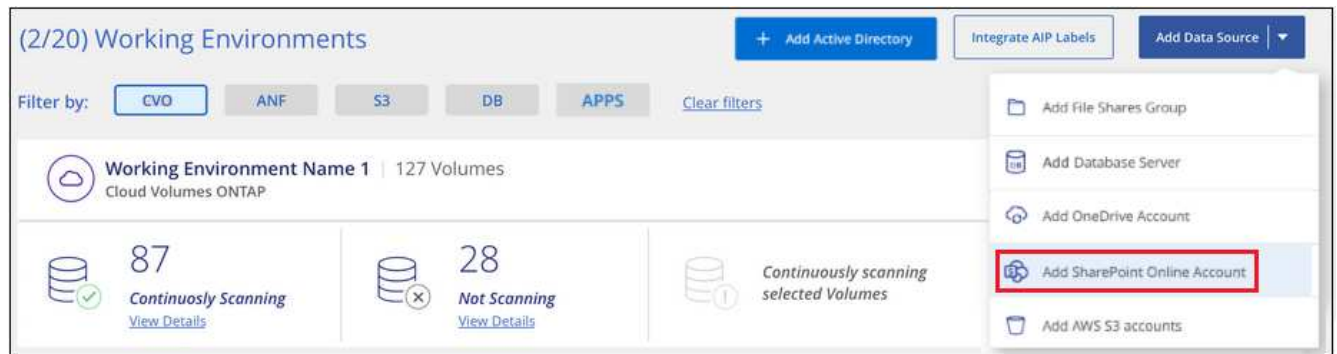
数据感知也可以是 "部署在无法访问 Internet 的内部位置"。但是，要扫描本地 SharePoint 文件，您需要为一些选定端点提供 Internet 访问。"请在此处查看所需端点的列表"。

正在添加 **SharePoint** 帐户

添加用户文件所在的 SharePoint 帐户。

步骤

1. 在工作环境配置页面中，单击 * 添加数据源 * > * 添加 SharePoint Online 帐户 *。



2. 在添加 SharePoint Online 帐户对话框中，单击 * 登录到 SharePoint*。
3. 在显示的 Microsoft 页面中，选择 SharePoint 帐户并输入所需的管理员用户和密码，然后单击 * 接受 * 以允许 Cloud Data sense 从此帐户读取数据。

SharePoint 帐户将添加到工作环境列表中。

将 **SharePoint** 站点添加到合规性扫描

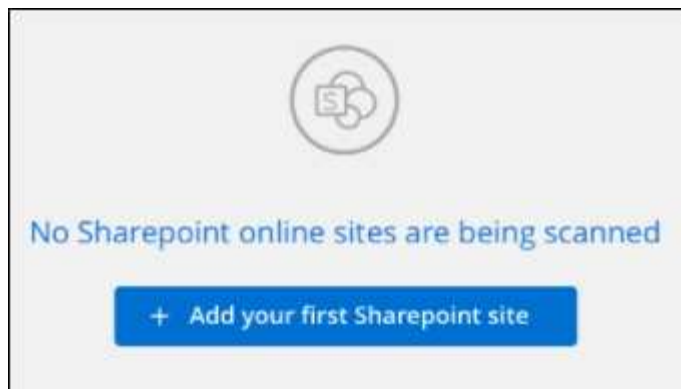
您可以在帐户中添加单个 SharePoint 站点或所有 SharePoint 站点，以便 Cloud Data sense 扫描关联的文件。

步骤

1. 在 *Configuration* 页面中，单击 SharePoint 帐户的 * 配置 * 按钮。



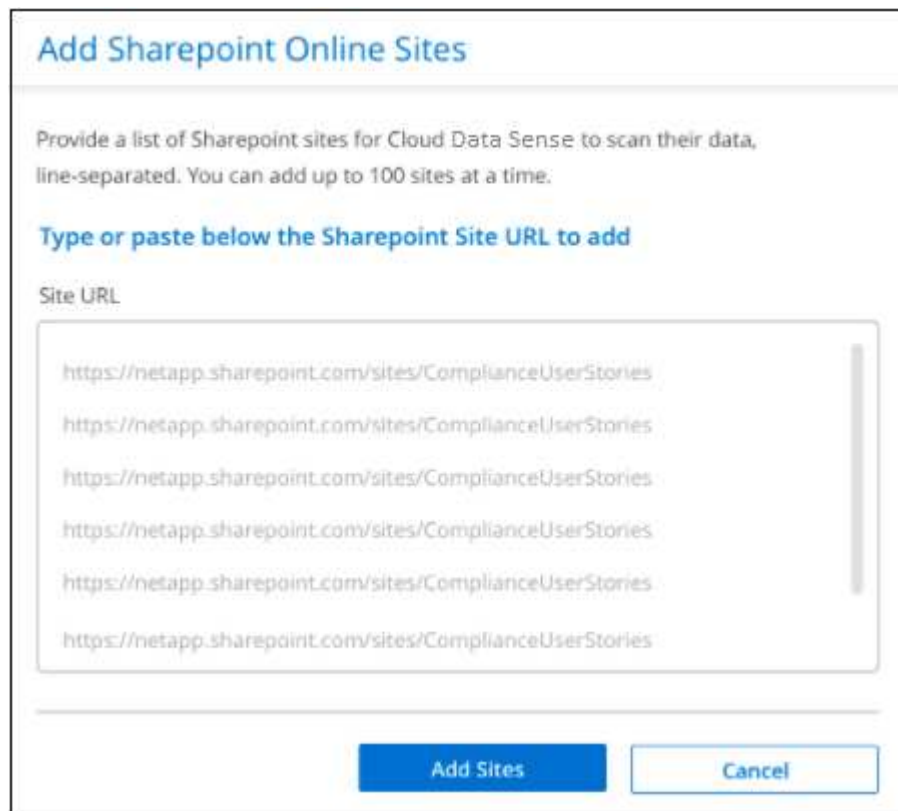
2. 如果这是首次为此 SharePoint 帐户添加站点，请单击 * 添加您的第一个 SharePoint 站点 *。



如果要从 SharePoint 帐户添加其他用户，请单击 * 添加 SharePoint 站点 *。



3. 为要扫描其文件的站点添加 URL - 每行一个 URL（每个会话最多 100 个 URL） - 然后单击 * 添加站点 *。



确认对话框将显示已添加的站点数量。

如果此对话框列出了任何无法添加的站点，请捕获此信息，以便您可以解析问题描述。在某些情况下，您可

以使用更正后的 URL 重新添加此站点。

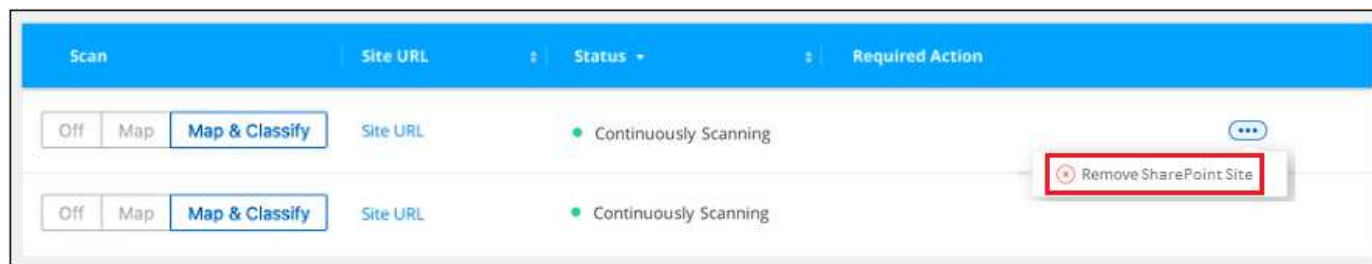
4. 对 SharePoint 站点中的文件启用仅映射扫描或映射和分类扫描。

收件人：	执行以下操作：
对文件启用仅映射扫描	单击 * 映射 *
对文件启用完全扫描	单击 * 映射和分类 *
禁用文件扫描	单击 * 关闭 *

Cloud Data sense 开始扫描您添加的 SharePoint 站点中的文件，结果将显示在信息板和其他位置。

从合规性扫描中删除 **SharePoint** 站点

如果您将来删除某个 SharePoint 站点，或者决定不扫描 SharePoint 站点中的文件，则可以随时删除各个 SharePoint 站点，使其无法扫描其文件。只需从配置页面中单击 * 删除 SharePoint 站点 * 即可。



请注意，您可以 ["从Data sense中删除整个SharePoint帐户"](#) 如果您不想再扫描SharePoint帐户中的任何用户数据。

扫描Google Drive帐户

请完成几个步骤、开始使用Cloud Data sense扫描Google Drive帐户中的用户文件。

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。

确保您拥有登录到Google Drive帐户所需的管理员凭据。

["部署 Cloud Data sense"](#) 如果尚未部署实例。

使用管理员用户凭据登录到要访问的Google Drive帐户、以便将其添加为新的数据源。

选择要对用户文件执行的扫描类型；映射或映射和分类。

查看Google Drive要求

请查看以下前提条件、以确保您已准备好在Google Drive帐户上启用Cloud Data sense。

- 您必须拥有Google Drive帐户的管理员登录凭据、该帐户可提供对用户文件的读取访问权限

当前限制

Google Drive文件当前不支持以下Data sense功能：

- 在"数据调查"页面中查看文件时、按钮栏中的操作未处于活动状态。您不能复制、移动、删除等任何文件。
- 无法在Google Drive中的文件中标识权限、因此调查页面中不会显示任何权限信息。

部署 Cloud Data sense

如果尚未部署实例，请部署 Cloud Data sense 。

数据感知可以是 "部署在云中" 或 "位于可访问 Internet 的内部位置"。

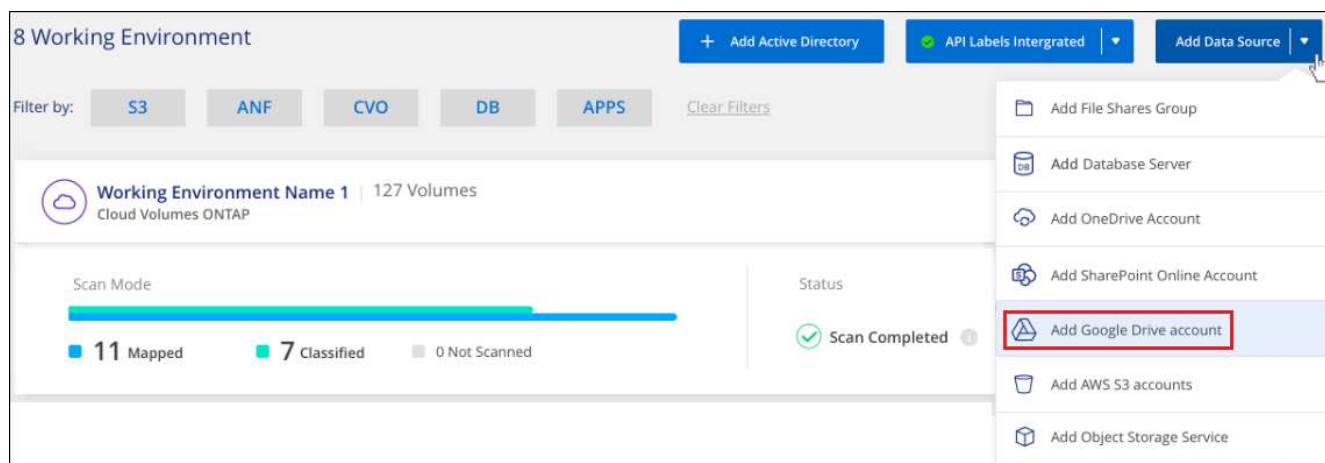
只要实例具有 Internet 连接，就会自动升级到 Data sense 软件。

正在添加Google Drive帐户

添加用户文件所在的Google Drive帐户。如果要扫描多个用户的文件、则需要对每个用户运行此步骤。

步骤

1. 在工作环境配置页面中、单击*添加数据源*>*添加Google Drive帐户*。



2. 在添加Google Drive帐户对话框中、单击*登录到Google Drive*。
3. 在显示的Google页面中、选择Google Drive帐户并输入所需的管理员用户和密码、然后单击*接受*以允许Cloud Data sense从此帐户读取数据。

Google Drive帐户将添加到工作环境列表中。

选择扫描用户数据的类型

选择Cloud Data sense将对用户数据执行的扫描类型。

步骤

1. 在_Configuration_页面中、单击Google Drive帐户的*配置*按钮。



2. 对Google Drive帐户中的文件启用仅映射扫描或映射和分类扫描。



收件人：	执行以下操作：
对文件启用仅映射扫描	单击 * 映射 *
对文件启用完全扫描	单击 * 映射和分类 *
禁用文件扫描	单击 * 关闭 *

Cloud Data sense开始扫描您添加的Google Drive帐户中的文件、结果将显示在信息板和其他位置。

从合规性扫描中删除Google Drive帐户

由于只有一个用户的Google Drive文件属于一个Google Drive帐户、因此、如果要停止扫描用户的Google Drive帐户中的文件、则应执行此操作 ["从Data sense中删除Google Drive帐户"](#)。

正在扫描文件共享

完成几个步骤，直接使用 Cloud Data sense 扫描非 NetApp NFS 或 CIFS 文件共享。这些文件共享可以驻留在内部或云中。

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。

对于 CIFS （ SMB ） 共享，请确保您具有访问这些共享的凭据。

["部署 Cloud Data sense"](#) 如果尚未部署实例。

组是要扫描的文件共享的容器，它用作这些文件共享的工作环境名称。

添加要扫描的文件共享列表并选择扫描类型。一次最多可以添加 100 个文件共享。

查看文件共享要求

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

- 共享可以托管在任何位置，包括云或内部。这些文件共享驻留在非 NetApp 存储系统上。
- Data sense 实例和共享之间需要有网络连接。
- 确保这些端口对 Data sense 实例开放：
 - 对于 NFS —端口 111 和 2049 。
 - 对于 CIFS —端口 139 和 445 。
- 您需要采用格式 ``<host_name> : /<share_path>`` 添加的共享列表。您可以单独输入共享，也可以提供要扫描的文件共享的行分隔列表。
- 对于 CIFS （ SMB ） 共享，请确保您具有 Active Directory 凭据来提供对共享的读取访问权限。如果 Cloud Data sense 需要扫描任何需要提升权限的数据，则最好使用管理员凭据。

如果要确保数据感知分类扫描不会更改文件的"上次访问时间"、我们建议用户具有"写入属性"权限。如果能、我们建议将Active Directory配置的用户设置为组织中有权访问所有文件的父组的一部分。

部署 Cloud Data sense 实例

如果尚未部署实例，请部署 Cloud Data sense 。

如果要扫描可通过 Internet 访问的非 NetApp NFS 或 CIFS 文件共享，则可以 ["在云中部署 Cloud Data sense"](#) 或 ["在可访问 Internet 的内部位置部署 Data sense"](#)。

如果要扫描的非 NetApp NFS 或 CIFS 文件共享安装在无法访问 Internet 的非公开站点中，则需要执行以下操作 ["在无法访问 Internet 的同一内部位置部署 Cloud Data sense"](#)。这还要求 Cloud Manager Connector 部署在同一内部位置。

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

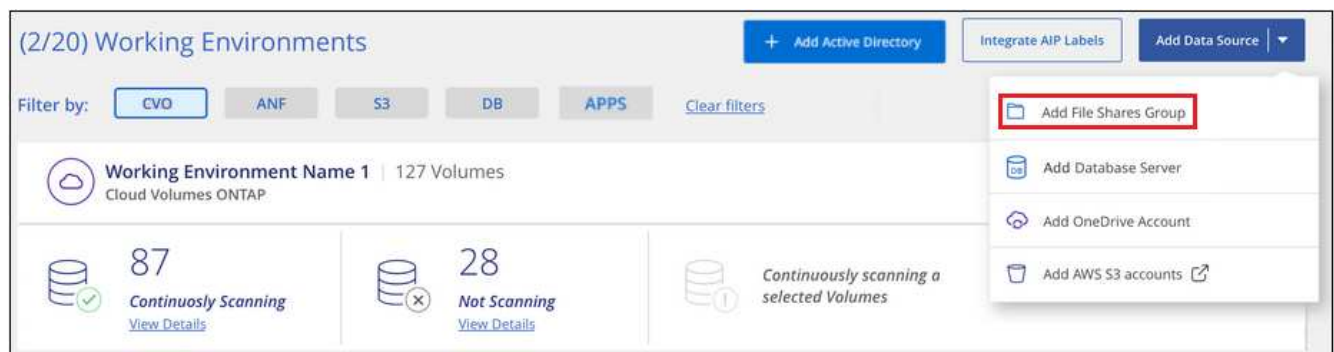
为文件共享创建组

您必须先添加文件共享 "group" ，然后才能添加文件共享。组是要扫描的文件共享的容器，组名称用作这些文件共享的工作环境名称。

您可以在同一个组中混用 NFS 和 CIFS 共享，但是，一个组中的所有 CIFS 文件共享都需要使用相同的 Active Directory 凭据。如果您计划添加使用不同凭据的 CIFS 共享，则必须为每组唯一的凭据创建一个单独的组。

步骤

1. 在工作环境配置页面中，单击 * 添加数据源 * > * 添加文件共享组 * 。



2. 在添加文件共享组对话框中，输入共享组的名称，然后单击 * 继续 *。

新的文件共享组将添加到工作环境列表中。

将文件共享添加到组

您可以将文件共享添加到文件共享组，以便 Cloud Data sense 扫描这些共享中的文件。添加的共享格式为 ``<host_name> : /<share_path>``。

您可以添加单个文件共享，也可以提供要扫描的文件共享的行分隔列表。一次最多可以添加 100 个共享。

在一个组中同时添加 NFS 和 CIFS 共享时，您需要运行此过程两次，一次是添加 NFS 共享，然后再次添加 CIFS 共享。

步骤

1. 在 *Working Environments* 页面中，单击文件共享组的 * 配置 * 按钮。



2. 如果这是首次为此文件共享组添加文件共享，请单击 * 添加您的第一个共享 *。



如果要向现有组添加文件共享，请单击 * 添加共享 *。



3. 选择要添加的文件共享的协议，添加要扫描的文件共享 - 每行一个文件共享 - 然后单击 * 继续 *。

添加 CIFS （ SMB ） 共享时，您需要输入 Active Directory 凭据，以提供对共享的读取访问权限。首选管理员凭据。

Adding Shares

Directly add any NFS or CIFS (SMB) File Shares, located in the cloud or on-premises.

Select Protocol

You'll be able to add additional shares from the other protocol later.

☒ NFS

☐ CIFS (SMB)

Type or paste below the Shares to add

Provide a list of shares, line-separated. You can add up to 100 at a time (you can add more later).

Hostname:/SHAREPATH

Hostname:/SHAREPATH

Hostname:/SHAREPATH

Continue

Cancel

☐ NFS

☒ CIFS (SMB)

Provide CIFS Credentials ⓘ

Username ⓘ

Password

确认对话框将显示已添加的共享数量。

如果此对话框列出了任何无法添加的共享，请捕获此信息，以便解析此问题描述。在某些情况下，您可以使用更正后的主机名或共享名称重新添加共享。

4. 在每个文件共享上启用仅映射扫描或映射和分类扫描。

收件人：	执行以下操作：
对文件共享启用仅映射扫描	单击 * 映射 *
对文件共享启用完全扫描	单击 * 映射和分类 *
禁用对文件共享的扫描	单击 * 关闭 *

Cloud Data sense 开始扫描您添加的文件共享中的文件，结果将显示在信息板和其他位置。

从合规性扫描中删除文件共享

如果您不再需要扫描某些文件共享，则可以随时从扫描其文件中删除各个文件共享。只需单击配置页面中的 * 删除共享 * 即可。



扫描使用 S3 协议的对象存储

完成几个步骤，直接使用 Cloud Data sense 扫描对象存储中的数据。数据感知功能可以扫描使用简单存储服务（ Simple Storage Service ， S3 ）协议的任何对象存储服务中的数据。其中包括 NetApp StorageGRID ， IBM 云对象存储， Azure Blob （使用 MinIO ）， Linode ， B2 云存储， Amazon S3 等。

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。

您需要具有端点 URL 才能连接到对象存储服务。

您需要从对象存储提供程序获取访问密钥和机密密钥，以便 Cloud Data sense 可以访问存储分段。

"部署 Cloud Data sense" 如果尚未部署实例。

将对象存储服务添加到 Cloud Data sense 。

选择要扫描的存储分段， Cloud Data sense 将开始扫描这些存储分段。

查看对象存储要求

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

- 您需要具有端点 URL 才能连接到对象存储服务。
- 您需要从对象存储提供程序获取访问密钥和机密密钥，以便 Data sense 可以访问存储分段。
- 要支持 Azure Blob ， 您需要使用 "MinIO 服务"。

部署 Cloud Data sense 实例

如果尚未部署实例，请部署 Cloud Data sense 。

如果要从可通过 Internet 访问的 S3 对象存储扫描数据，则可以 "在云中部署 Cloud Data sense" 或 "在可访问 Internet 的内部位置部署 Data sense"。

如果要从安装在无法访问 Internet 的非公开站点中的 S3 对象存储扫描数据，则需要执行以下操作 "在无法访问 Internet 的同一内部位置部署 Cloud Data sense"。这还要求 Cloud Manager Connector 部署在同一内部位置。

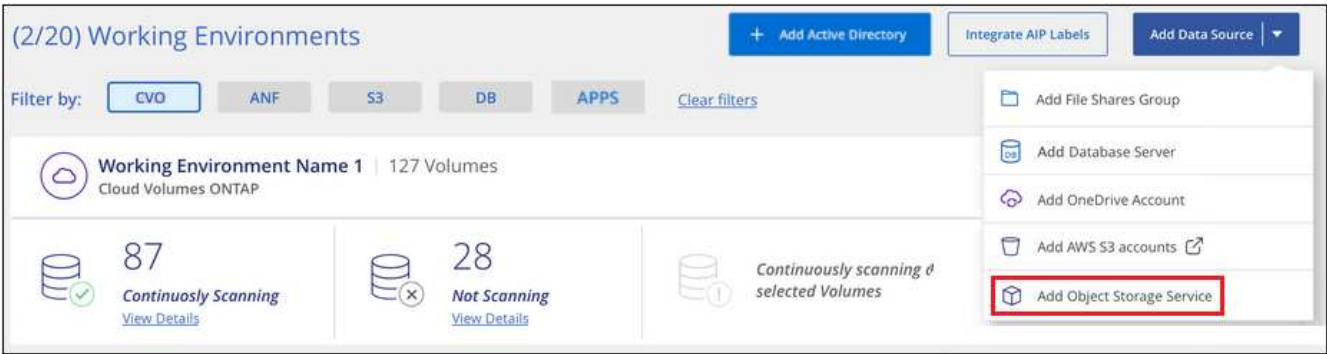
只要实例具有 Internet 连接，就会自动升级到 Data sense 软件。

将对象存储服务添加到 **Cloud Data sense**

添加对象存储服务。

步骤

- 1. 在工作环境配置页面中，单击 * 添加数据源 * > * 添加对象存储服务 *。



- 2. 在添加对象存储服务对话框中，输入对象存储服务的详细信息，然后单击 * 继续 *。
 - a. 输入要用于工作环境的名称。此名称应反映要连接到的对象存储服务的名称。
 - b. 输入端点 URL 以访问对象存储服务。
 - c. 输入访问密钥和机密密钥，以便 Cloud Data sense 可以访问对象存储中的存储分段。

Add Object Storage Service

Cloud Data Sense can scan data from any Object Storage service which uses the S3 protocol. This includes NetApp StorageGRID, IBM Object Store, and more.

To continue, enter the following information. In the next steps you'll need to select the buckets you want to scan.

Name the Working Environment

Endpoint URL

object_myIBM

http://my.endpoint.com

Access Key

Secret Key

AJUKDO574NDJG86795

.....

Continue

Cancel

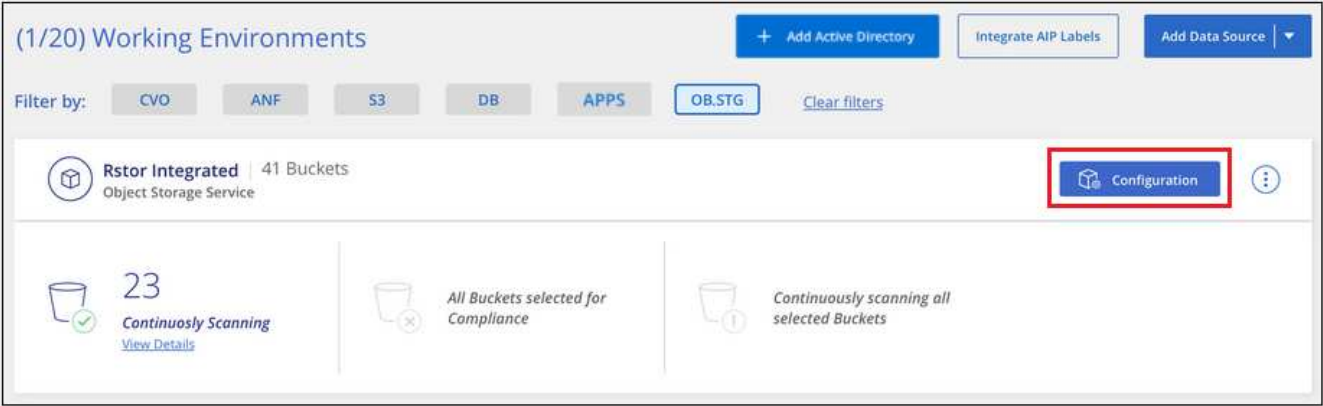
新的对象存储服务将添加到工作环境列表中。

启用和禁用对象存储分段上的合规性扫描

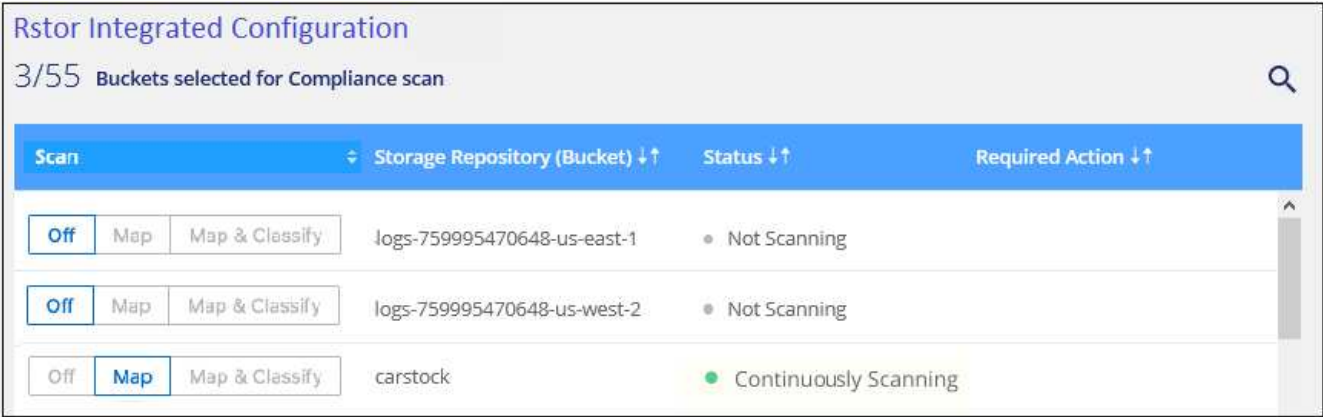
在对象存储服务上启用 Cloud Data sense 后，下一步是配置要扫描的分段。Data sense 会发现这些分段并将其显示在您创建的工作环境中。

步骤

1. 在配置页面中，单击对象存储服务工作环境中的 * 配置 *。



2. 在存储分段上启用仅映射扫描或映射和分类扫描。



收件人：	执行以下操作：
在存储分段上启用仅映射扫描	单击 * 映射 *
对存储分段启用完全扫描	单击 * 映射和分类 *
禁用对存储分段的扫描	单击 * 关闭 *

Cloud Data sense 开始扫描您启用的存储分段。如果存在任何错误，它们将显示在状态列中，并显示修复此错误所需的操作。

将 Active Directory 与 Cloud Data sense 集成

您可以将全局 Active Directory 与 Cloud Data sense 集成在一起，以增强 Data sense 报告有关文件所有者以及哪些用户和组有权访问您的文件的结果。

在设置某些数据源（如下所示）时，您需要输入 Active Directory 凭据才能使 Data sense 扫描 CIFS 卷。此集成可为 Data sense 提供这些数据源中数据的文件所有者和权限详细信息。为这些数据源输入的 Active Directory 可能与您在此输入的全局 Active Directory 凭据不同。Data sense 将在所有集成的 Active Directory 中查找用户和权限详细信息。

此集成可在 Data sense 的以下位置提供追加信息：

- 您可以使用 " 文件所有者 " ["筛选器"](#) 并在 " 调查 " 窗格中查看文件元数据的结果。此文件所有者不包含 SID（安全标识符），而是使用实际用户名进行填充。
- 您可以看到 ["完整文件权限"](#) 单击"查看所有权限"按钮时对每个文件和目录执行此操作。
- 在中 ["监管信息板"](#)下，打开权限面板将显示有关数据的更详细信息。



本地用户 SID 和未知域中的 SID 不会转换为实际用户名。

支持的数据源

Active Directory 与 Cloud Data sense 的集成可以识别以下数据源中的数据：

- 内部部署 ONTAP 系统
- Cloud Volumes ONTAP
- Azure NetApp Files
- 适用于 ONTAP 的 FSX
- 非NetApp CIFS文件共享(非NFS文件共享)

不支持从数据库架构、OneDrive帐户、SharePoint帐户、Google Drive帐户、Amazon S3帐户、或使用简单存储服务(S3)协议的对象存储。

正在连接到 **Active Directory** 服务器

在部署 Data sense 并对数据源激活扫描之后，您可以将 Data sense 与 Active Directory 集成。可以使用 DNS 服务器 IP 地址或 LDAP 服务器 IP 地址访问 Active Directory 。

Active Directory 凭据可以是只读凭据，但提供管理员凭据可确保 Data sense 可以读取任何需要提升权限的数据。这些凭据存储在 Cloud Data sense 实例上。

对于CIFS卷/文件共享、如果要确保数据感知分类扫描未更改文件的"上次访问时间"、建议用户具有写入属性权限。如果可能、我们建议将Active Directory配置的用户设置为组织中有权访问所有文件的父组的一部分。

要求

- 您必须已为公司中的用户设置 Active Directory 。
- 您必须具有 Active Directory 的信息：

- DNS 服务器 IP 地址或多个 IP 地址

或

LDAP 服务器 IP 地址或多个 IP 地址

- 用于访问服务器的用户名和密码
- 域名（ Active Directory 名称）
- 是否使用安全 LDAP （ LDAPS ）

◦ LDAP 服务器端口（对于 LDAP，通常为 389；对于安全 LDAP，通常为 636）

- 以下端口必须为 Data sense 实例的出站通信打开：

协议	Port	目标	目的
TCP 和 UDP	389.	Active Directory	LDAP
TCP	636	Active Directory	基于 SSL 的 LDAP
TCP	3268	Active Directory	全局目录
TCP	3369	Active Directory	基于 SSL 的全局目录

步骤

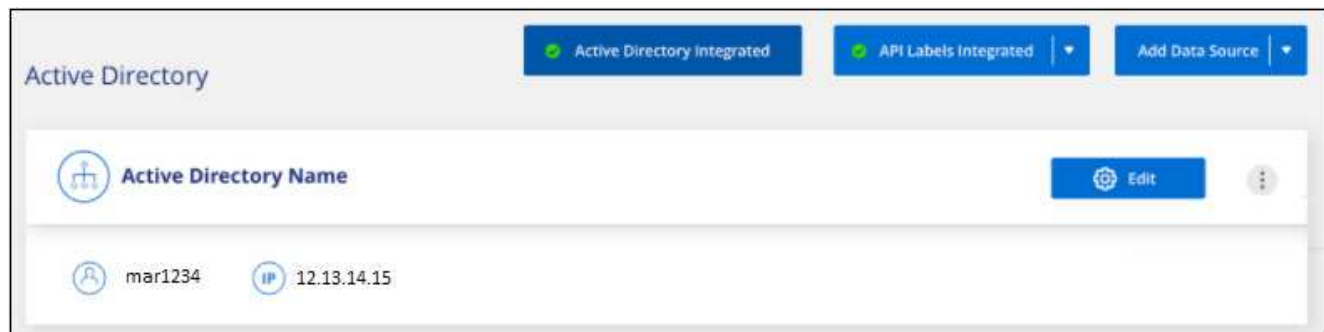
1. 在 "Cloud Data sense Configuration" 页面中，单击 * 添加 Active Directory* 。



2. 在连接到 Active Directory 对话框中，输入 Active Directory 详细信息，然后单击 * 连接 * 。

如果需要，可以单击 * 添加 IP* 来添加多个 IP 地址。

Data sense 可集成到 Active Directory，并在配置页面中添加了一个新部分。



管理 Active Directory 集成

如果需要修改 Active Directory 集成中的任何值，请单击 * 编辑 * 按钮并进行更改。

如果您不再需要集成，也可以通过单击来删除此集成  按钮，然后单击 * 删除 Active Directory*。

为 Cloud Data sense 设置许可

Cloud Data 感知在 Cloud Manager 工作空间中扫描的前 1 TB 数据是免费的。要在这之后继续扫描数据，需要 NetApp 提供的 BYOL 许可证或云提供商所在市场提供的 Cloud Manager 订阅。


在阅读其他内容之前，请先阅读一些注释：

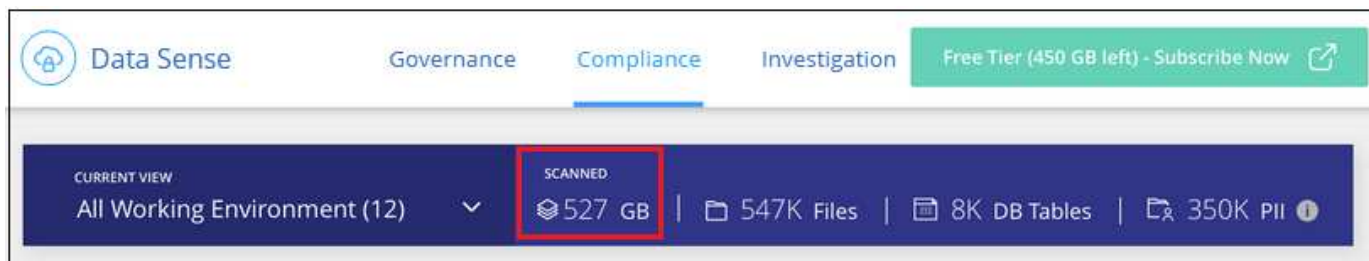
- 如果您已在云提供商的市场中订阅 Cloud Manager 按需购买（PAYGO）订阅，则您也会自动订阅 Cloud Data sense。您无需重新订阅。
- Cloud Data sense 自带许可证（BYOL）是一种 *float* 许可证，您可以在计划扫描的工作空间中的所有工作环境和数据源中使用。您将在数字电子钱包中看到有效订阅。

["详细了解与 Cloud Data sense 相关的许可和成本"](#)。

使用 Cloud Data sense PAYGO 订阅

您可以从云提供商的市场订阅按需购买，从而获得使用 Cloud Volumes ONTAP 系统和云数据服务（如云数据感知）的许可。

您可以随时订阅，在数据量超过 1 TB 之前，不会向您收取任何费用。您始终可以从数据感知信息板查看正在扫描的总数据量。现在订阅  按钮可以让您在准备就绪后轻松订阅。



这些步骤必须由具有 *Account Admin* 角色的用户完成。

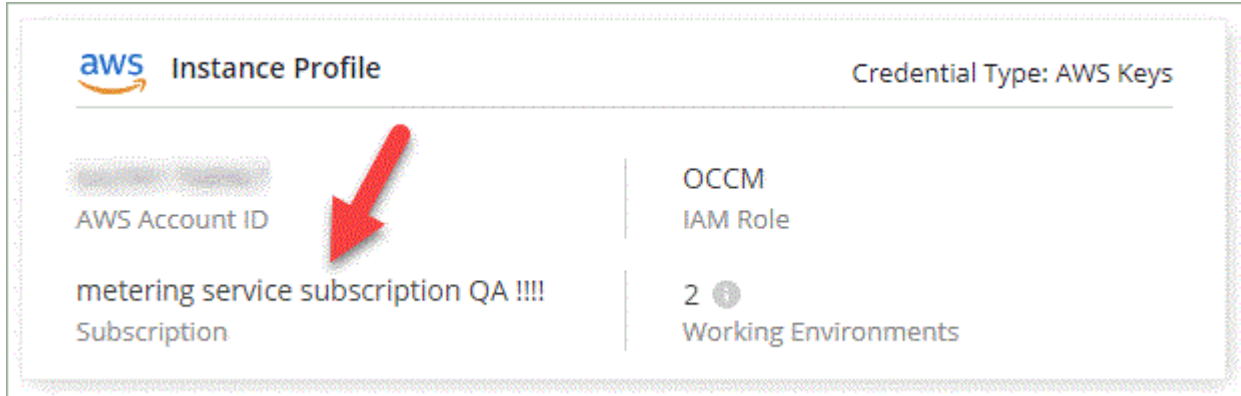
1. 在 Cloud Manager 控制台的右上角，单击设置图标，然后选择 * 凭据 *。



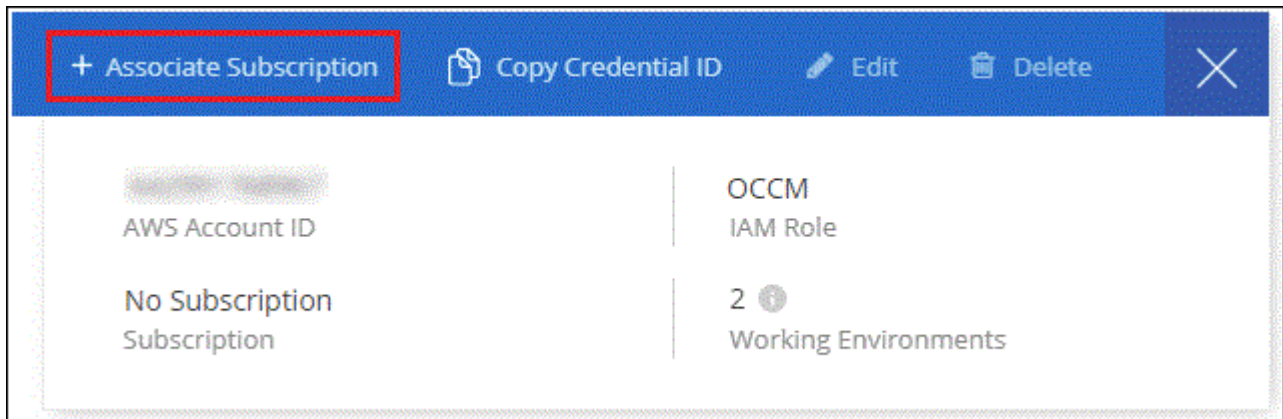
2. 查找 AWS 实例配置文件，Azure 托管服务身份或 Google Project 的凭据。

必须将订阅添加到实例配置文件，托管服务身份或 Google Project 中。否则，充电将不起作用。

如果您已订阅（如下所示为 AWS），则表示您已准备就绪—无需执行任何其他操作。



3. 如果您尚未订阅，请将鼠标悬停在凭据上，单击操作菜单，然后单击 * 关联订阅 *。



4. 选择现有订阅并单击 * 关联 *，或者单击 * 添加订阅 * 并按照步骤进行操作。

以下视频显示了如何关联 "AWS Marketplace" 订阅 AWS 订阅：

► https://docs.netapp.com/zh-cn/cloud-manager-data-sense//media/video_subscribing_aws.mp4 (video)

以下视频显示了如何关联 "Azure Marketplace" 订阅 Azure：

► https://docs.netapp.com/zh-cn/cloud-manager-data-sense//media/video_subscribing_azure.mp4 (video)

以下视频显示了如何关联 "GCP 市场" 订阅 GCP：

► https://docs.netapp.com/zh-cn/cloud-manager-data-sense//media/video_subscribing_gcp.mp4 (video)

使用 Cloud Data sense BYOL 许可证

NetApp 自带许可证的期限为 1 年，2 年或 3 年。BYOL * 云数据感知 * 许可证是一种 *float* 许可证，其中总容量由所有 * 工作环境和数据源共享，从而使初始许可和续订变得轻松。

如果您没有 Cloud Data sense 许可证，请联系我们购买一个：

- mailto : ng-contact-data-sense@netapp.com ? Subject=Licensing[发送电子邮件以购买许可证] 。
- 单击 Cloud Manager 右下角的聊天图标以请求许可证。

或者，如果您已为 Cloud Volumes ONTAP 取消分配了基于节点的许可证，而您不会使用该许可证，则可以将其转换为具有相同美元等价性和相同到期日期的云数据感知许可证。"有关详细信息，请访问此处"。

您可以使用 Cloud Manager 中的数字电子钱包页面管理 Cloud Data sense BYOL 许可证。您可以添加新许可证并更新现有许可证。

获取 Cloud Data sense 许可证文件

购买 Cloud Data sense 许可证后，您可以通过输入 Cloud Data sense 序列号和 NSS 帐户或上传 NLF 许可证文件在 Cloud Manager 中激活此许可证。以下步骤显示了如果您计划使用此方法，如何获取 NLF 许可证文件。

如果您已在无法访问 Internet 的内部站点中的主机上部署 Cloud Data sense ，则需要从已连接 Internet 的系统获取许可证文件。使用序列号和 NSS 帐户激活许可证不适用于非公开站点安装。

步骤

1. 登录到 "NetApp 支持站点" 然后单击 * 系统 > 软件许可证 * 。
2. 输入 Cloud Data sense 许可证序列号。


Software Licenses

Serial Number

481*

Serial #	Cluster SN	License Name	License Key	Host ID	Value	End Date
Serial #	Cluster SN	License Name		Host ID	Value	End Date
4810		SUBS-CLD-DAT-SENSE-TB-2Y	Get NetApp License File		100	12/31/9998

3. 在 * 许可证密钥 * 下，单击 * 获取 NetApp 许可证文件 * 。
4. 输入您的 Cloud Manager 帐户 ID （在支持站点上称为租户 ID ），然后单击 * 提交 * 下载许可证文件。



Get License

SERIAL NUMBER: 4810

LICENSE: SUBS-CLD-DAT-SENSE-TB-2Y

SALES ORDER: 3005

TENANT ID:

Example: account-xxxxxxx

[Cancel](#) [Submit](#)

您可以通过从 Cloud Manager 顶部选择 * 帐户 * 下拉列表，然后单击您帐户旁边的 * 管理帐户 * 来查找 Cloud Manager 帐户 ID。您的帐户 ID 位于概述选项卡中。

将 **Cloud Data sense BYOL** 许可证添加到您的帐户中

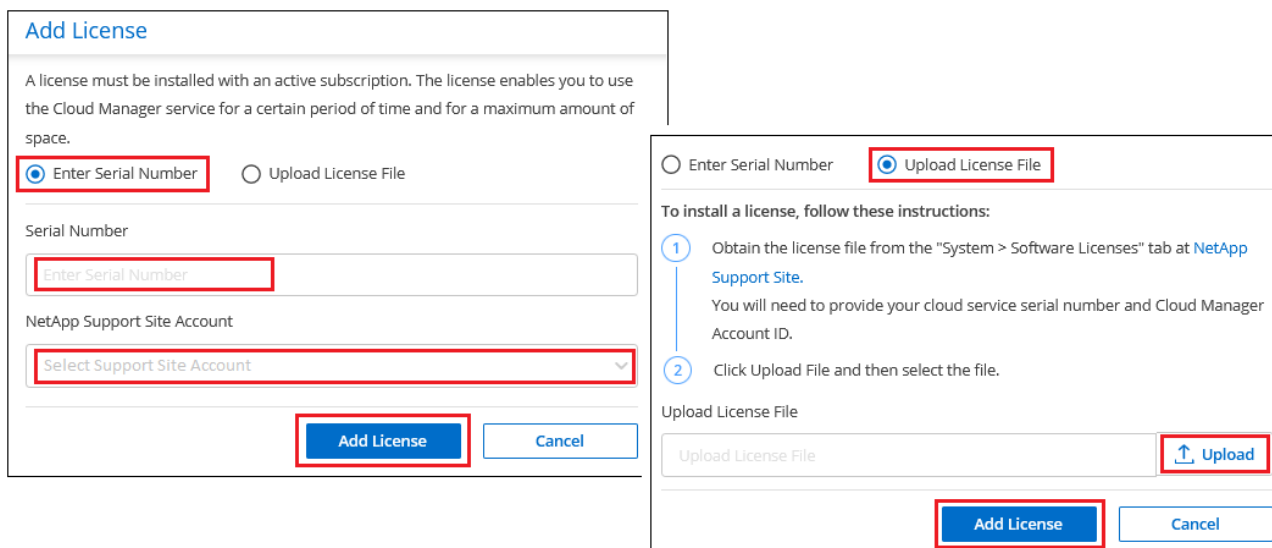
为 Cloud Manager 帐户购买 Cloud Data sense 许可证后，您需要将此许可证添加到 Cloud Manager 中才能使用 Data sense 服务。

步骤

1. 从 Cloud Manager 左侧导航菜单中、单击 * 数字电子邮件 *、然后选择 * 数据服务许可证 * 选项卡。
2. 单击 * 添加许可证 *。
3. 在 *Add License* 对话框中，输入许可证信息并单击 * 添加许可证 *：
 - 如果您拥有数据感知许可证序列号并知道您的 NSS 帐户，请选择 * 输入序列号 * 选项并输入该信息。

如果下拉列表中没有您的 NetApp 支持站点帐户，["将 NSS 帐户添加到 Cloud Manager"](#)。

- 如果您有数据感知许可证文件（安装在非公开站点时需要），请选择 * 上传许可证文件 * 选项，然后按照提示附加该文件。



Add License

A license must be installed with an active subscription. The license enables you to use the Cloud Manager service for a certain period of time and for a maximum amount of space.

☒ Enter Serial Number ☐ Upload License File

Serial Number

NetApp Support Site Account

[Add License](#) [Cancel](#)

Upload License File

☐ Enter Serial Number ☒ Upload License File

To install a license, follow these instructions:

- 1 Obtain the license file from the "System > Software Licenses" tab at [NetApp Support Site](#). You will need to provide your cloud service serial number and Cloud Manager Account ID.
- 2 Click Upload File and then select the file.

Upload License File

[Upload](#)

[Add License](#) [Cancel](#)

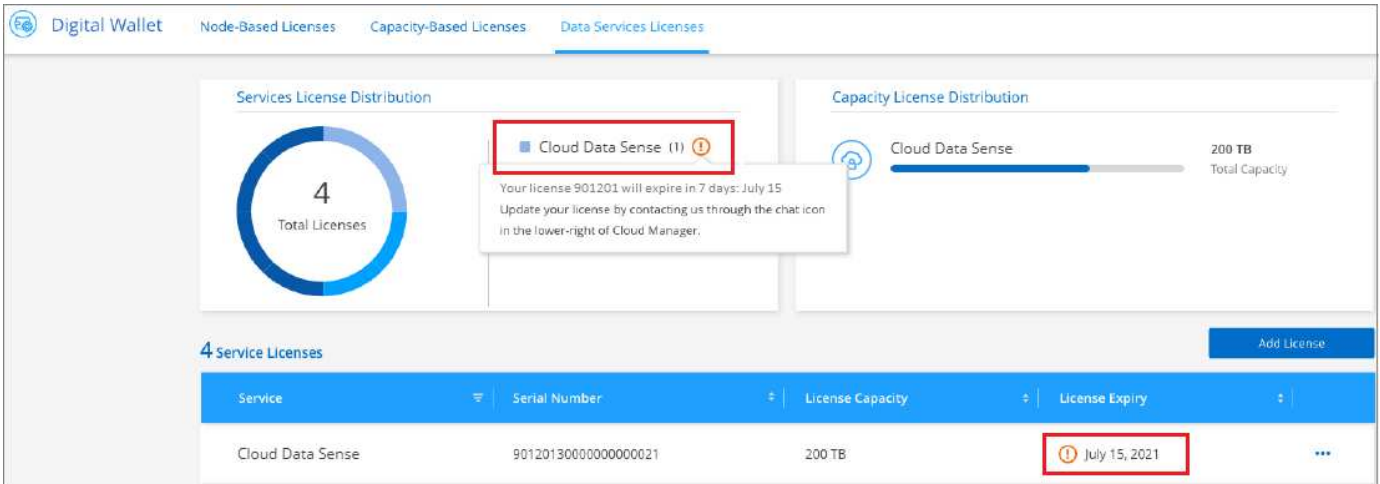
Cloud Manager 会添加许可证，以便 Cloud Data sense 服务处于活动状态。

更新 Cloud Data sense BYOL 许可证

如果您的许可期限即将到期，或者您的许可容量即将达到限制，您将在 Cloud Data sense 中收到通知。



此状态也会显示在数字电子钱包中。



您可以在 Cloud Data sense 许可证到期之前对其进行更新，以便访问扫描数据的能力不会中断。

步骤

1. 单击 Cloud Manager 右下角的聊天图标以请求延长您的期限或为特定序列号请求 Cloud Data sense 许可证的额外容量。您也可以发送电子邮件至：ng-contact-data-sense@netapp.com ? Subject=Licensing[发送电子邮件以请求更新您的许可证]。

在您为许可证付费并将其注册到 NetApp 支持站点后，Cloud Manager 会自动在数字电子邮件中更新许可证，并且数据服务许可证页面将在 5 到 10 分钟内反映此更改。

2. 如果 Cloud Manager 无法自动更新许可证（例如，安装在非公开站点时），则需要手动上传许可证文件。
 - a. 您可以从 [NetApp 支持站点](#) 获取许可证文件。
 - b. 在数字电子邮件页面的 *Data Services Licenses* 选项卡中，单击 ... 对于要更新的服务序列号，请单击 * 更新许可证 *。



c. 在 *Update License* 页面中，上传许可证文件并单击 * 更新许可证 *。

Cloud Manager 会更新许可证，以便 Cloud Data sense 服务继续保持活动状态。

BYOL 许可证注意事项

使用 Cloud Data sense BYOL 许可证时，当您正在扫描的所有数据的大小接近容量限制或接近许可证到期日期时，Cloud Manager 会在 Data sense UI 和 Digital Wallet UI 中显示警告。您会收到以下警告：

- 扫描的数据量达到许可容量的 80% 时，再次达到限制时
- 许可证到期前 30 天，许可证到期后再次

如果您看到这些警告，请使用 Cloud Manager 界面右下角的聊天图标续订许可证。

如果许可证到期，Data sense 将继续运行，但会阻止对信息板的访问，因此您无法查看有关任何已扫描数据的信息。如果您希望减少要扫描的卷数量，从而可能使容量使用量低于许可证限制，则只能使用 *Configuration* 页面。

续订 BYOL 许可证后，Cloud Manager 会自动更新 Digital Wallet 中的许可证，并提供对所有信息板的完全访问权限。如果 Cloud Manager 无法通过安全 Internet 连接访问此许可证文件（例如，安装在非公开站点时），您可以自行获取此文件并手动将其上传到 Cloud Manager。有关说明，请参见 [如何更新 Cloud Data sense 许可证](#)。



如果您正在使用的帐户同时具有 BYOL 许可证和 PAYGO 订阅，则在 BYOL 许可证到期后，Data sense *will not* 将转移到 PAYGO 订阅。您必须续订 BYOL 许可证。

有关 Cloud Data sense 的常见问题

如果您只是在寻找一个问题的快速答案，此常见问题解答将会有所帮助。

云数据感知服务

以下问题概括说明了云数据的意义。

什么是云数据的意义？

Cloud Data sense 是一款云产品，它使用人工智能(AI)驱动的技术帮助您了解数据环境并识别存储系统中的敏感数据。这些系统可以是您添加到 Cloud Manager Canvas 中的工作环境，也可以是 Data sense 可以通过网络访问的多种类型的数据源。 [请参见下面的完整列表](#)。

Cloud Data sense 提供了预定义参数（例如敏感信息类型和类别），用于满足有关数据隐私和敏感度的新数据合规性法规，例如 GDPR，CCPA，HIPAA 等。

为什么要使用 **Cloud Data sense**？

Cloud Data sense 可以为您提供数据，帮助您：

- 遵守数据合规性和隐私法规。
- 将数据从传统系统迁移到云。
- 遵守数据保留策略。
- 根据 GDPR，CCPA，HIPAA 和其他数据隐私法规的要求，根据数据主题轻松查找和报告特定数据。

Cloud Data sense 的常见用例有哪些？

- 识别个人身份信息（PII）。
- 根据 GDPR 和 CCPA 隐私法规的要求确定广泛的敏感信息。
- 遵守新的和即将出台的数据隐私法规。

["详细了解 Cloud Data sense 的用例"](#)。

云数据感知的工作原理是什么？

Cloud Data sense 可在 Cloud Manager 系统和存储系统的同时部署另一层人工智能。然后，它会扫描卷、分段、数据库和其他存储帐户上的数据，并为找到的数据洞察力编制索引。与通常围绕正则表达式和模式匹配构建的替代解决方案相比，Data sense 可以利用人工智能和自然语言处理。Cloud Data sense 利用 AI 提供对数据的上下文了解，以实现准确的检测和分类。它由 AI 驱动，因为它专为现代数据类型和规模而设计。它还了解数据环境，以便提供强大、准确的发现和分类。

["详细了解 Cloud Data sense 的工作原理"](#)。

Cloud Data sense 扫描我的数据的频率如何？

数据经常发生变化，因此 Cloud Data sense 可以持续扫描数据，而不会对数据造成影响。虽然数据的初始扫描可能需要较长时间，但后续扫描只会扫描增量更改，从而缩短系统扫描时间。

["了解扫描的工作原理"](#)。

数据扫描对存储系统和数据的影响可以忽略不计。但是，如果您担心即使影响很小，也可以将 Data sense 配置为执行 "缓慢" 扫描。 ["请参见如何降低扫描速度"](#)。

Cloud Data sense 支持哪些部署模式？

Cloud Data sense 通常使用 SaaS 模式进行部署，在这种模式中，服务通过 Cloud Manager 界面启用。Cloud Data sense 可以部署在几乎任何位置的系统上进行扫描和报告，包括内部环境、云和混合环境。对于安全安装，Cloud Manager 和 Cloud Data sense 可以部署在 "非公开站点" 模式中，该模式作为软件包安装在内部，不需要外部网络连接。

是否有人可以访问在我的浏览器和**Data sense**之间发送的私有数据？

否在浏览器和Data sense实例之间发送的私有数据通过端到端加密得到保护、这意味着NetApp和第三方无法读取。除非您请求并批准访问、否则Data sense不会与NetApp共享任何数据或结果。

Cloud Data sense 是否提供报告？

是的。Cloud Data sense 提供的信息可能与您组织中的其他利益相关方相关，因此我们可以帮助您生成报告以分享这些见解。以下报告可用于 Data sense：

隐私风险评估报告

根据您的数据提供隐私洞察力并获得隐私风险得分。 ["了解更多信息。"](#)。

数据主体访问请求报告

用于提取包含数据主体的特定名称或个人标识符相关信息的所有文件的报告。 ["了解更多信息。"](#)。

PCI DSS 报告

帮助您确定信用卡信息在整个文件中的分布情况。 ["了解更多信息。"](#)。

HIPAA 报告

帮助您确定运行状况信息在文件中的分布情况。 ["了解更多信息。"](#)。

数据映射报告

提供有关工作环境中文件大小和数量的信息。其中包括使用容量，数据期限，数据大小和文件类型。 ["了解更多信息。"](#)。

报告特定信息类型

我们提供的报告包含有关包含个人数据和敏感个人数据的已识别文件的详细信息。您还可以查看按类别和文件类型细分的文件。 ["了解更多信息。"](#)。

扫描性能是否有所不同？

扫描性能可能因网络带宽和环境中的平均文件大小而异。它还可能取决于主机系统（在云端或内部）的大小特征。请参见 ["云数据感知实例"](#) 和 ["部署 Cloud Data sense"](#) 有关详细信息 ...

在首次添加新数据源时，您还可以选择仅执行 "映射" 扫描，而不是执行完整的 "分类" 扫描。由于无法访问文件以查看数据源中的数据，因此可以非常快速地对数据源进行映射。 ["查看映射扫描与分类扫描之间的区别。"](#)

如何启用 **Cloud Data sense** ？

首先，您需要在 Cloud Manager 中部署 Cloud Data sense 实例。实例运行后、您可以从*数据感知*选项卡或通过选择特定的工作环境在现有工作环境、数据库和其他数据源上启用此服务。

["了解如何开始使用"](#)。



在数据源上激活Cloud Data sense会立即执行初始扫描。扫描结果会在之后不久显示。

如何禁用 **Cloud Data sense** ？

您可以从 "数据感知配置" 页面禁用 Cloud Data sense 扫描单个工作环境，数据库，文件共享组，OneDrive 帐

户或 SharePoint 帐户。

["了解更多信息。"](#)



要完全删除 Cloud Data sense 实例，您可以从云提供商的门户或内部位置手动删除 Data sense 实例。

如果在 **ONTAP** 卷上启用了数据分层，会发生什么情况？

您可能希望在将冷数据分层到对象存储的 ONTAP 系统上启用云数据感知。如果启用了数据分层，则 Data sense 会扫描所有数据—磁盘上的数据以及分层到对象存储的冷数据。

合规性扫描不会加热冷数据，它会保持冷数据并分层到对象存储。

Cloud Data sense 能否向我的组织发送通知？

是的。通过与策略功能结合使用，您可以在策略返回结果时向 Cloud Manager 用户发送电子邮件警报（每日，每周或每月），以便您可以收到保护数据的通知。了解更多信息 ["策略"](#)。

您还可以从 "监管" 页面和 "调查" 页面下载状态报告，并在组织内部共享这些报告。

我是否可以根据组织的需求自定义服务？

Cloud Data sense 提供对数据的即装即用洞察力。您可以根据组织的需求提取和利用这些洞察信息。

此外，您还可以使用 * 数据检测 * 功能让 Fusion 根据您正在扫描的数据库中特定列中的标准扫描所有数据，这实际上使您可以创建自己的自定义个人数据类型。

["了解更多信息。"](#)

Cloud Data sense 是否可以与我的文件中嵌入的 **AIP** 标签配合使用？

是的。如果您已订阅，则可以管理 Cloud Data sense 正在扫描的文件中的 AIP 标签 ["Azure 信息保护（AIP）"](#)。您可以查看已分配给文件的标签，向文件添加标签以及更改现有标签。

["了解更多信息。"](#)

是否可以将云数据感知信息限制为特定用户？

是的，Cloud Data sense 已与 Cloud Manager 完全集成。Cloud Manager 用户只能根据其工作空间权限查看其有资格查看的工作环境的信息。

此外，如果您希望允许某些用户只查看数据感知扫描结果而不能管理数据感知设置，则可以为这些用户分配 *Cloud Compliance Viewer* 角色。

["了解更多信息。"](#)

支持哪些云提供商？

Cloud Data sense 作为 Cloud Manager 的一部分运行，并支持 AWS，Azure 和 GCP。这样，您的组织就可以在不同的云提供商之间实现统一的隐私可见性。

源系统的类型和数据类型

以下问题与可扫描的存储类型以及所扫描的数据类型有关。

可以使用**Data sense**扫描哪些数据源？

Cloud Data sense可以扫描您添加到Cloud Manager Canvas的工作环境中的数据、以及Data sense可以通过网络访问的多种类型的数据源中的数据。

- 工作环境： *
- Cloud Volumes ONTAP （部署在 AWS ， Azure 或 GCP 中）
- 内部 ONTAP 集群
- Azure NetApp Files
- 适用于 ONTAP 的 Amazon FSX
- Amazon S3
- 数据源： *
- 非 NetApp 文件共享
- 对象存储（使用 S3 协议）
- 数据库(Amazon RDS、MongoDB、MySQL、Oracle、PostgreSQL、 SAP HANA、SQL Server)
- OneDrive 帐户
- SharePoint 帐户
- Google Drive帐户

Data sense 支持 NFS 3.x ， 4.0 和 4.1 以及 CIFS 1.x ， 2.0 ， 2.1 和 3.0 版。

在政府区域部署时是否存在任何限制？

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。

如果在无法访问**Internet**的站点上安装**Data sense**、可以扫描哪些数据源？

数据感知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"非公开"站点中的以下本地数据源：

- 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

支持哪些文件类型？

Cloud Data sense会扫描所有文件以获取类别和元数据洞察力、并在信息板的文件类型部分显示所有文件类型。

当Data sense检测到个人可识别信息(PiD)或执行DSAL搜索时、仅支持以下文件格式：

.CSV、.dcm、.Dicom、.DOC、.docx、.json、.PDF、.PPTX、.RTV、.TXT、.XLS、.XLSX、文档、工作表和幻灯片

Cloud Data可以捕获哪些类型的数据和元数据？

您可以通过Cloud Data sense对数据源运行常规"映射"扫描或完整的"分类"扫描。映射仅提供数据的概览，而 "分类" 则提供数据的深度扫描。由于无法访问文件以查看数据源中的数据，因此可以非常快速地对数据源进行映射。

- 数据映射扫描。

Data sense仅扫描元数据。这对于整体数据管理和监管、快速的项目范围界定、非常大的资产和优先级排序非常有用。数据映射基于元数据、被视为*快速*扫描。

快速扫描后、您可以生成数据映射报告。本报告概述了存储在企业数据源中的数据、可帮助您确定资源利用率、迁移、备份、安全性和合规性流程。

- 数据分类(深度)扫描。

在整个客户环境中使用标准协议和只读权限进行数据感知扫描。系统会打开并扫描选定文件、以查看与业务相关的敏感数据、私有信息以及与勒索软件相关的问题。

完整扫描后、您可以对数据应用许多其他数据感知功能、例如在"数据调查"页面中查看和细化数据、搜索文件中的名称、复制、移动和删除源文件等。

许可证和成本

以下问题与使用Cloud Data sense的许可和成本有关。

云数据的成本有多高？

使用 Cloud Data sense 的成本取决于您要扫描的数据量。Data sense 在 Cloud Manager 工作空间中扫描的前 1 TB 数据是免费的。达到此限制后、您需要执行以下操作之一才能继续扫描超过1 TB的数据：

- 您的云提供商或订阅Cloud Manager Marketplace列表
- NetApp自带许可证(BYOL)

请参见 ["定价"](#) 了解详细信息。

如果我已达到**BYOL**容量限制、会发生什么情况？

如果达到BYOL容量限制、则Data sense会继续运行、但会阻止对信息板的访问、因此您无法查看有关任何已扫描数据的信息。如果您希望减少要扫描的卷数量、从而可能使容量使用量低于许可证限制、则只能使用配置页面。您必须续订BYOL许可证才能重新获得对Data sense的完全访问权限。

连接器部署

以下问题与Cloud Manager Connector相关。

什么是连接器？

Connector是在您的云帐户或内部环境中的计算实例上运行的软件、可使Cloud Manager安全地管理云资源。您必须部署Connector才能使用Cloud Data sense。

连接器需要安装在何处？

- 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。
- 在 GCP 的 Cloud Volumes ONTAP 中扫描数据时， 您可以在 GCP 中使用连接器。
- 在扫描内部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帐户中的数据时、您可以在任何这些云位置使用连接器。

因此、如果您在其中许多位置都有数据、则可能需要使用 ["多个连接器"](#)。

是否可以在自己的主机上部署此连接器？

是的。您可以 ["在内部部署 Connector"](#) 在网络或云中的 Linux 主机上。如果您计划在内部部署Data sense、则可能还需要在内部安装Connector；但这并不是必需的。

没有Internet访问的安全站点如何？

是的、也支持这种做法。您可以 ["在无法访问Internet的内部Linux主机上部署Connector"](#)。然后、您可以发现内部ONTAP 集群和其他本地数据源、并使用Data sense扫描数据。

数据感知部署

以下问题与单独的数据感知实例相关。

云数据感知需要哪种类型的实例或虚拟机？

时间 ["部署在云中"](#)：

- 在 AWS 中， Cloud Data sense 在具有 500 GB GP2 磁盘的 m5.4xlarge 实例上运行。
- 在 Azure 中， Cloud Data sense 在具有 512 GB 磁盘的 Standard_d16s_v3 VM 上运行。
- 在 GCP 中， Cloud Data sense 在具有 512 GB 标准永久性磁盘的 n2-standard-16 VM 上运行。

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。

["详细了解 Cloud Data sense 的工作原理"](#)。

是否可以在自己的主机上部署**Data sense**？

是的。您可以在可通过网络或云访问 Internet 的 Linux 主机上安装 Data sense 软件。所有功能均相同，您可以继续通过 Cloud Manager 管理扫描配置和结果。请参见 ["在内部部署 Cloud Data sense"](#) 了解系统要求和安装详细信息。

没有**Internet**访问的安全站点如何？

是的、也支持这种做法。您可以 ["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。

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。