



入门 Cloud Sync

NetApp
April 01, 2022

目录

- 入门 1
 - Cloud Sync 概述 1
 - Cloud Sync 快速入门 3
 - 支持的同步关系 4
 - 准备源和目标 11
 - Cloud Sync 网络概述 17
 - 安装数据代理 19

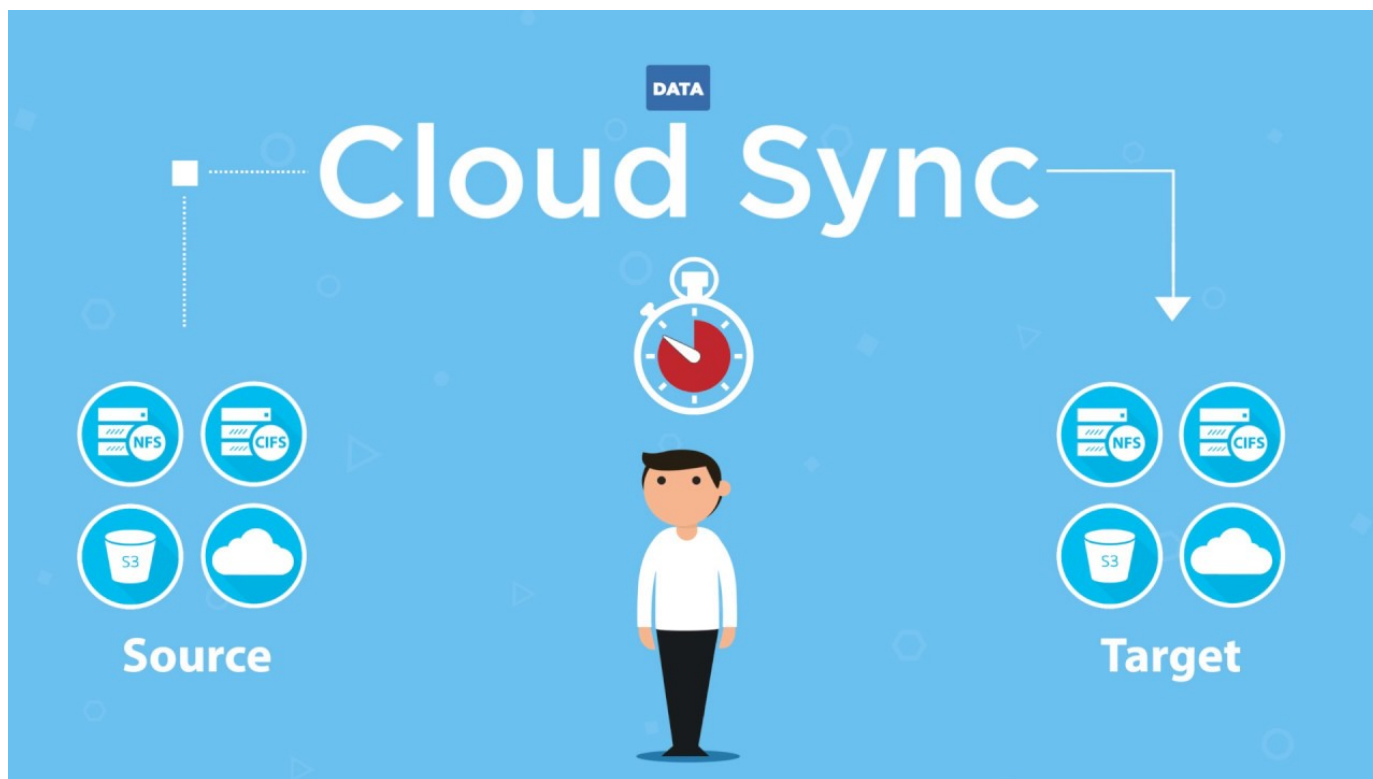
入门

Cloud Sync 概述

NetApp Cloud Sync 服务提供了一种简单、安全且自动化的方式、可以将数据迁移到云或您的企业内部的任何目标。无论是基于文件的 NAS 数据集（NFS 或 SMB）、Amazon Simple Storage Service（S3）对象格式、NetApp StorageGRID® 设备还是任何其他云提供商对象存储、Cloud Sync 都可以为您转换和移动它。

功能

观看以下视频以了解 Cloud Sync 的概述：



Cloud Sync 的工作原理

Cloud Sync 是一个软件即服务（SaaS）平台，由一个数据代理组，一个通过 Cloud Manager 提供的基于云的界面以及一个源和目标组成。

下图显示了 Cloud Sync 组件之间的关系：



NetApp 数据代理软件会将数据从源同步到目标（称为 *sync relationship*）。您可以在 AWS、Azure、Google Cloud 平台或内部运行数据代理。一个由一个或多个数据代理组成的数据代理组需要通过端口 443 建立出站 Internet 连接，以便能够与 Cloud Sync 服务进行通信并联系其他一些服务和存储库。["查看端点列表。"](#)

初始副本之后、服务将根据您设置的计划同步所有更改的数据。

支持的存储类型

Cloud Sync 支持以下存储类型：

- 任何 NFS 服务器
- 任何 SMB 服务器
- Amazon EFS
- 适用于 ONTAP 的 Amazon FSX
- Amazon S3
- Azure Blob
- Azure NetApp Files
- 框（可作为预览版使用）
- Cloud Volumes Service
- Cloud Volumes ONTAP
- Google Cloud 存储
- IBM 云对象存储

- 内部 ONTAP 集群
- ONTAP S3 存储
- SFTP （仅使用 API ）
- StorageGRID

["查看支持的同步关系"](#)。

Costs

使用 Cloud Sync 会产生两种成本：资源费用和服务费用。

资源费用

资源费用与在云中运行一个或多个数据代理的计算和存储成本有关。

服务费用

在 14 天免费试用结束后、有两种方式可以支付同步关系的费用。第一种选择是从 AWS 或 Azure 订阅，这样您就可以按小时或按年付费。第二种选择是直接从 NetApp 购买许可证。

["了解许可的工作原理"](#)。

Cloud Sync 快速入门

Cloud Sync 服务入门包括几个步骤。

验证您的源和目标是否受支持和设置。最重要的要求是验证数据代理组与源位置和目标位置之间的连接。["了解更多信息"](#)。

NetApp 数据代理软件会将数据从源同步到目标（称为 *sync relationship*）。您可以在 AWS、Azure、Google Cloud 平台或内部运行数据代理。一个由一个或多个数据代理组成的数据代理组需要通过端口 443 建立出站 Internet 连接，以便能够与 Cloud Sync 服务进行通信并联系其他一些服务和存储库。["查看端点列表"](#)。

在创建同步关系时，Cloud Sync 会指导您完成安装过程，此时您可以在云中部署数据代理或为自己的 Linux 主机下载安装脚本。

- ["查看 AWS 安装"](#)
- ["查看 Azure 安装"](#)
- ["查看 Google Cloud 安装"](#)
- ["查看 Linux 主机安装"](#)

登录到 ["云管理器"](#)，单击 * 同步 *，然后拖放您为源和目标选择的内容。按照提示完成设置。["了解更多信息"](#)。

从 AWS 或 Azure 订阅以按需购买或每年支付。或直接从 NetApp 购买许可证。只需转到 Cloud Sync 中的“许可证设置”页面即可进行设置。["了解更多信息"](#)。

支持的同步关系

使用 Cloud Sync 可以将数据从源同步到目标。这称为同步关系。您应该先了解支持的关系，然后再开始。

源位置	支持的目标位置
Amazon EFS	<ul style="list-style-type: none">• Amazon EFS• 适用于 ONTAP 的 Amazon FSX• Amazon S3• Azure Blob• Azure NetApp Files• Cloud Volumes ONTAP• Cloud Volumes Service• Google Cloud 存储• IBM 云对象存储• NFS 服务器• 内部 ONTAP 集群• SMB 服务器• StorageGRID
适用于 ONTAP 的 Amazon FSX	<ul style="list-style-type: none">• Amazon EFS• 适用于 ONTAP 的 Amazon FSX• Amazon S3• Azure Blob• Azure NetApp Files• Cloud Volumes ONTAP• Cloud Volumes Service• Google Cloud 存储• IBM 云对象存储• NFS 服务器• 内部 ONTAP 集群• SMB 服务器• StorageGRID

源位置	支持的目标位置
Amazon S3	<ul style="list-style-type: none"> • Amazon EFS • 适用于 ONTAP 的 Amazon FSX • Amazon S3 • Azure Blob • Azure NetApp Files • 盒装 ^2 , 3^ • Cloud Volumes ONTAP • Cloud Volumes Service • Google Cloud 存储 • IBM 云对象存储 • NFS 服务器 • 内部 ONTAP 集群 • SMB 服务器 • StorageGRID
Azure Blob	<ul style="list-style-type: none"> • Amazon EFS • 适用于 ONTAP 的 Amazon FSX • Amazon S3 • Azure Blob • Azure NetApp Files • Cloud Volumes ONTAP • Cloud Volumes Service • Google Cloud 存储 • IBM 云对象存储 • NFS 服务器 • 内部 ONTAP 集群 • SMB 服务器 • StorageGRID

源位置	支持的目标位置
Azure NetApp Files	<ul style="list-style-type: none"> • Amazon EFS • 适用于 ONTAP 的 Amazon FSX • Amazon S3 • Azure Blob • Azure NetApp Files • Cloud Volumes ONTAP • Cloud Volumes Service • Google Cloud 存储 • IBM 云对象存储 • NFS 服务器 • 内部 ONTAP 集群 • SMB 服务器 • StorageGRID
盒 ²	<ul style="list-style-type: none"> • Amazon S3 • IBM 云对象存储 • NFS 服务器 • SMB 服务器 • StorageGRID
Cloud Volumes ONTAP	<ul style="list-style-type: none"> • Amazon EFS • 适用于 ONTAP 的 Amazon FSX • Amazon S3 • Azure Blob • Azure NetApp Files • Cloud Volumes ONTAP • Cloud Volumes Service • Google Cloud 存储 • IBM 云对象存储 • NFS 服务器 • 内部 ONTAP 集群 • SMB 服务器 • StorageGRID

源位置	支持的目标位置
Cloud Volumes Service	<ul style="list-style-type: none"> • Amazon EFS • 适用于 ONTAP 的 Amazon FSX • Amazon S3 • Azure Blob • Azure NetApp Files • Cloud Volumes ONTAP • Cloud Volumes Service • Google Cloud 存储 • IBM 云对象存储 • NFS 服务器 • 内部 ONTAP 集群 • SMB 服务器 • StorageGRID
Google Cloud 存储	<ul style="list-style-type: none"> • Amazon EFS • 适用于 ONTAP 的 Amazon FSX • Amazon S3 • Azure Blob • Azure NetApp Files • Cloud Volumes ONTAP • Cloud Volumes Service • Google Cloud 存储 • IBM 云对象存储 • NFS 服务器 • 内部 ONTAP 集群 • ONTAP S3 存储 • SMB 服务器 • StorageGRID

源位置	支持的目标位置
IBM 云对象存储	<ul style="list-style-type: none"> • Amazon EFS • 适用于 ONTAP 的 Amazon FSX • Amazon S3 • Azure Blob • Azure NetApp Files • 盒装 ^2 , 3^ • Cloud Volumes ONTAP • Cloud Volumes Service • Google Cloud 存储 • IBM 云对象存储 • NFS 服务器 • 内部 ONTAP 集群 • SMB 服务器 • StorageGRID
NFS 服务器	<ul style="list-style-type: none"> • Amazon EFS • 适用于 ONTAP 的 Amazon FSX • Amazon S3 • Azure Blob • Azure NetApp Files • Cloud Volumes ONTAP • Cloud Volumes Service • Google Cloud 存储 • IBM 云对象存储 • NFS 服务器 • 内部 ONTAP 集群 • SMB 服务器 • StorageGRID

源位置	支持的目标位置
内部 ONTAP 集群	<ul style="list-style-type: none"> • Amazon EFS • 适用于 ONTAP 的 Amazon FSX • Amazon S3 • Azure Blob • Azure NetApp Files • Cloud Volumes ONTAP • Cloud Volumes Service • Google Cloud 存储 • IBM 云对象存储 • NFS 服务器 • 内部 ONTAP 集群 • SMB 服务器 • StorageGRID
ONTAP S3 存储	<ul style="list-style-type: none"> • Google Cloud 存储 • SMB 服务器 • StorageGRID • ONTAP S3 存储
SFTP ¹	S3
SMB 服务器	<ul style="list-style-type: none"> • Amazon EFS • 适用于 ONTAP 的 Amazon FSX • Amazon S3 • Azure Blob • Azure NetApp Files • Cloud Volumes ONTAP • Cloud Volumes Service • Google Cloud 存储 • IBM 云对象存储 • NFS 服务器 • 内部 ONTAP 集群 • ONTAP S3 存储 • SMB 服务器 • StorageGRID

源位置	支持的目标位置
StorageGRID	<ul style="list-style-type: none"> • Amazon EFS • 适用于 ONTAP 的 Amazon FSX • Amazon S3 • Azure Blob • Azure NetApp Files • 盒装 ^2 , 3^ • Cloud Volumes ONTAP • Cloud Volumes Service • Google Cloud 存储 • IBM 云对象存储 • NFS 服务器 • 内部 ONTAP 集群 • ONTAP S3 存储 • SMB 服务器 • StorageGRID

注释：

1. 仅使用 Cloud Sync API 支持与此源 / 目标的同步关系。
2. 预览版可提供盒式支持。
3. 当 BLOB 容器是目标容器时，可以选择特定的 Azure Blob 存储层：
 - 热存储
 - 冷却存储
4. **【存储类】** 当 Amazon S3 为目标时，您可以选择特定的 S3 存储类：
 - 标准（这是默认类）
 - Intelligent-Hierarchy
 - 标准—不经常访问
 - 一个 ZONE 不常访问
 - 冰河
 - Glacier 深度存档
5. 当 Google Cloud Storage 存储分段为目标时，您可以选择特定的存储类：
 - 标准
 - 近线
 - 冷线

- 归档

准备源和目标

验证源和目标是否满足以下要求。

网络

- 源和目标必须与数据代理组建立网络连接。

例如，如果数据中心中有 NFS 服务器，而 AWS 中有数据代理，则需要从网络到 VPC 的网络连接（VPN 或 Direct Connect）。

- NetApp 建议配置源，目标和数据代理以使用网络时间协议（NTP）服务。三个组件之间的时间差不应超过 5 分钟。

目标目录

创建同步关系时，您可以使用 Cloud Sync 选择现有目标目录，然后选择在该目录中创建一个新文件夹。因此，请确保您的首选目标目录已存在。

读取目录的权限

为了显示源或目标中的每个目录或文件夹，Cloud Sync 需要对目录或文件夹具有读取权限。

NFS

必须使用文件和目录上的 uid/GID 在源 / 目标上定义权限。

对象存储

- 对于 AWS 和 Google Cloud，数据代理必须具有列表对象权限（如果您按照数据代理安装步骤进行操作，则默认情况下会提供这些权限）。
- 对于 Azure，StorageGRID 和 IBM，您在设置同步关系时输入的凭据必须具有列表对象权限。

SMB

设置同步关系时输入的 SMB 凭据必须具有列表文件夹权限。



默认情况下，数据代理会忽略以下目录：.snapshot，~snapshot，.copy-of-overt

Amazon S3 存储分段要求

确保您的 Amazon S3 存储分段满足以下要求。

Amazon S3 支持的数据代理位置

包含 S3 存储的同步关系需要在 AWS 或您的内部部署数据代理。在这两种情况下、Cloud Sync 都会提示您在安装期间将数据代理与 AWS 帐户关联。

- ["了解如何部署 AWS 数据代理"](#)

- ["了解如何在 Linux 主机上安装数据代理"](#)

支持的 **AWS** 区域

除中国地区外，所有地区均受支持。

在其他 **AWS** 帐户中 **S3** 数据段所需的权限

在设置同步关系时，您可以指定一个 S3 存储分段，该存储分段驻留在与数据代理不关联的 AWS 帐户中。

"此 [JSON 文件中包含的权限](#)" 必须应用于该 S3 存储分段，以便数据代理可以访问它。这些权限使数据代理可以将数据复制到存储桶中或从中复制数据、并列出存储桶中的对象。


请注意以下有关 JSON 文件中包含的权限的信息：

1. `<BucketName>` 是位于 AWS 帐户中且与数据代理无关的存储分段名称。
2. 应将 `<RoleARN>` 替换为以下项之一：
 - 如果在 Linux 主机上手动安装了数据代理，则 `RoleARN` 应是部署数据代理时为其提供 AWS 凭据的 AWS 用户的 ARN。
 - 如果使用 CloudFormation 模板在 AWS 中部署了数据代理，则 `RoleARN` 应是此模板创建的 IAM 角色的 ARN。

您可以通过转至 EC2 控制台、选择数据代理实例并从“描述”选项卡中单击 IAM 角色来查找角色 ARN。然后，应在 IAM 控制台中查看包含角色 ARN 的“摘要”页面。

Summary

Delete role

Role ARN `arn:aws:iam::428991742922:role/tanyaBroker0304-DataBrokerIamRole-1VMHWXMW3AQ05` 

Role description [Edit](#)

Azure Blob 存储要求

确保 Azure Blob 存储满足以下要求。

Azure Blob 支持的数据代理位置

如果同步关系包括 Azure Blob 存储，则数据代理可以驻留在任何位置。

支持的 **Azure** 区域

除中国、美国政府和美国国防部地区外，所有地区均受支持。

包含 **Azure Blob** 和 **NFS/SMB** 的关系的连接字符串

在 Azure Blob 容器和 NFS 或 SMB 服务器之间创建同步关系时，需要使用存储帐户连接字符串提供 Cloud Sync：

a63cde60b553020 - Access keys

Storage account

Search (Ctrl+/)

Overview

Activity log

Access control (IAM)

Tags

Diagnose and solve problems

Storage Explorer (preview)

Settings

Access keys

CORS

Configuration

Encryption

Use access keys to authenticate your applications when making requests to this Azure storage account. Store your access keys securely - for example, using Azure Key Vault - and don't share them. We recommend regenerating your access keys regularly. You are provided two access keys so that you can maintain connections using one key while regenerating the other.

When you regenerate your access keys, you must update any Azure resources and applications that access this storage account to use the new keys. This action will not interrupt access to disks from your virtual machines. [Learn more](#)

Storage account name

a63cde60b553020

key1

Key

vScjFdvVZqIPyO/

Connection string

DefaultEndpoints

如果要在两个 Azure Blob 容器之间同步数据，则连接字符串必须包含 **"共享访问签名"**（SAS）。在 blob 容器和 NFS 或 SMB 服务器之间同步时，您还可以选择使用 SAS。

SAS 必须允许访问 Blob 服务和所有资源类型（服务、容器和对象）。SAS 还必须包括以下权限：

- 对于源 blob 容器：读取并列出
- 对于目标 BLOB 容器：读取、写入、列出、添加和创建

Search (Ctrl+/)

Overview

Activity log

Access control (IAM)

Tags

Diagnose and solve problems

Storage Explorer (preview)

Settings

Access keys

CORS

Configuration

Encryption

Shared access signature

Firewalls and virtual networks

Advanced Threat Protection (pr...

Properties

Locks

Allowed services ⓘ
☒ Blob ☐ File ☐ Queue ☐ Table

Allowed resource types ⓘ
☒ Service ☒ Container ☒ Object

Allowed permissions ⓘ
☒ Read ☒ Write ☒ Delete ☒ List ☒ Add ☒ Create ☐ Update ☐ Process

Start and expiry date/time ⓘ
Start
2018-10-23 10:07:32 AM
End
2019-10-23 6:07:32 PM
(UTC-04:00) --- Current Time Zone ---

Allowed IP addresses ⓘ
for example, 168.1.5.65 or 168.1.5.65-168.1.5.70

Allowed protocols ⓘ
☒ HTTPS only ☐ HTTPS and HTTP

Signing key ⓘ
key1

Generate SAS and connection string

Azure NetApp Files 要求

在与 Azure NetApp Files 同步数据时，请使用高级或超高级服务级别。如果磁盘服务级别为标准，则可能会出现故障和性能问题。



如果您需要帮助确定合适的服务级别，请咨询解决方案架构师。卷大小和卷层决定了您可以获得的吞吐量。

["详细了解 Azure NetApp Files 服务级别和吞吐量"。](#)

包装箱要求

- 要创建包含框的同步关系，您需要提供以下凭据：
 - 客户端 ID
 - 客户端密钥
 - 专用密钥。
 - 公有密钥 ID
 - 密码短语

- 企业 ID
- 如果要创建从 Amazon S3 到 Box 的同步关系，则必须使用具有统一配置且以下设置设置为 1 的数据代理组：
 - 扫描程序并发
 - 扫描程序进程限制
 - 传输并发性
 - 传输程序进程限制

["了解如何为数据代理组定义统一配置"](#)。

Google Cloud 存储桶要求

确保 Google Cloud 存储桶满足以下要求。

Google Cloud 存储支持的数据代理位置

包含 Google Cloud Storage 的同步关系要求在 Google Cloud 或内部部署一个数据代理。在创建同步关系时，Cloud Sync 将指导您完成数据代理安装过程。

- ["了解如何部署 Google Cloud 数据代理"](#)
- ["了解如何在 Linux 主机上安装数据代理"](#)

支持的 Google Cloud 地区

支持所有区域。

其他 Google Cloud 项目中的存储分段的权限

在设置同步关系时，如果您为数据代理的服务帐户提供了所需的权限，则可以从不同项目中的 Google Cloud 存储分段中进行选择。 ["了解如何设置服务帐户"](#)。

SnapMirror 目标的权限

如果同步关系的源是 SnapMirror 目标（只读），则 "读 / 列表" 权限足以将数据从源同步到目标。

NFS 服务器要求

- NFS 服务器可以是 NetApp 系统或非 NetApp 系统。
- 文件服务器必须允许数据代理主机访问导出。
- 支持 NFS 版本 3、4.0、4.1 和 4.2。

必须在服务器上启用所需的版本。

- 如果要从 ONTAP 系统同步 NFS 数据，请确保已启用对 SVM NFS 导出列表的访问（已启用 `vserver nfs modify -vserver svm_name -showmount`）。



从 ONTAP 9.2 开始，showmount 的默认设置为 `_enabled`。

ONTAP 要求

如果同步关系包括 Cloud Volumes ONTAP 或内部 ONTAP 集群，并且您选择了 NFSv4 或更高版本，则需要在 ONTAP 系统上启用 NFSv4 ACL。复制 ACL 时需要执行此操作。

ONTAP S3 存储要求

设置包括的同步关系时 **"ONTAP S3 存储"**，您需要提供以下内容：

- 连接到 ONTAP S3 的 LIF 的 IP 地址
- ONTAP 配置为使用的访问密钥和机密密钥

SMB 服务器要求

- SMB 服务器可以是 NetApp 系统或非 NetApp 系统。
- 您需要为 Cloud Sync 提供对 SMB 服务器具有权限的凭据。
 - 对于源 SMB 服务器，需要以下权限：list 和 read。

源 SMB 服务器支持备份操作员组的成员。

- 对于目标 SMB 服务器，需要以下权限：list，read 和 write。
- 文件服务器必须允许数据代理主机访问导出。
- 支持 SMB 版本 1.0，2.0，2.1，3.0 和 3.11。
- 向 "管理员" 组授予对源文件夹和目标文件夹的 "完全控制" 权限。

如果不授予此权限，则数据代理可能没有足够的权限来获取文件或目录上的 ACL。如果发生这种情况，您将收到以下错误："getxattr error 95"

隐藏目录和文件的 **SMB** 限制

在 SMB 服务器之间同步数据时，SMB 限制会影响隐藏的目录和文件。如果源 SMB 服务器上的任何目录或文件通过 Windows 隐藏，则隐藏属性不会复制到目标 SMB 服务器。

由于大小写不敏感限制而导致的 **SMB** 同步行为

SMB 协议不区分大小写，这意味着大小写字母将被视为相同。如果同步关系包含 SMB 服务器且目标上已存在数据，则此行为可能会导致文件被覆盖和目录复制错误。

例如，假设源上有一个名为 "A" 的文件，目标上有一个名为 "A" 的文件。当 Cloud Sync 将名为 "A" 的文件复制到目标时，文件 "A" 将被源中的文件 "A" 覆盖。

对于目录，假设源上有一个名为 "b" 的目录，目标上有一个名为 "B" 的目录。当 Cloud Sync 尝试将名为 "b" 的目录复制到目标时，Cloud Sync 会收到一条错误，指出此目录已存在。因此，Cloud Sync 始终无法复制名为 "b" 的目录。

避免此限制的最佳方法是确保将数据同步到空目录。

Cloud Sync 网络概述

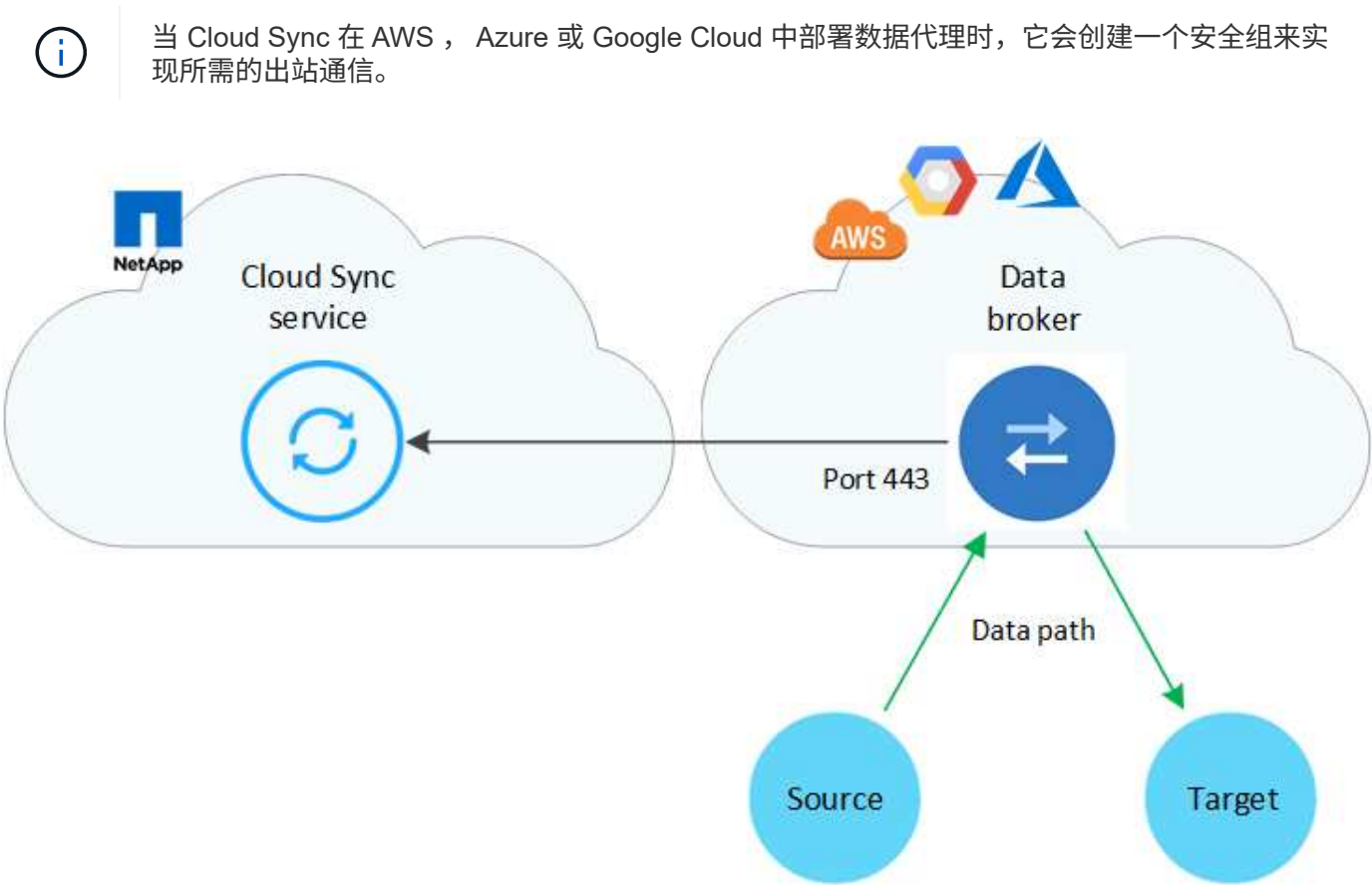
Cloud Sync 网络包括数据代理组与源位置和目标位置之间的连接，以及数据代理通过端口 443 的出站 Internet 连接。

数据代理位置

数据代理组由一个或多个安装在云或内部的数据代理组成。

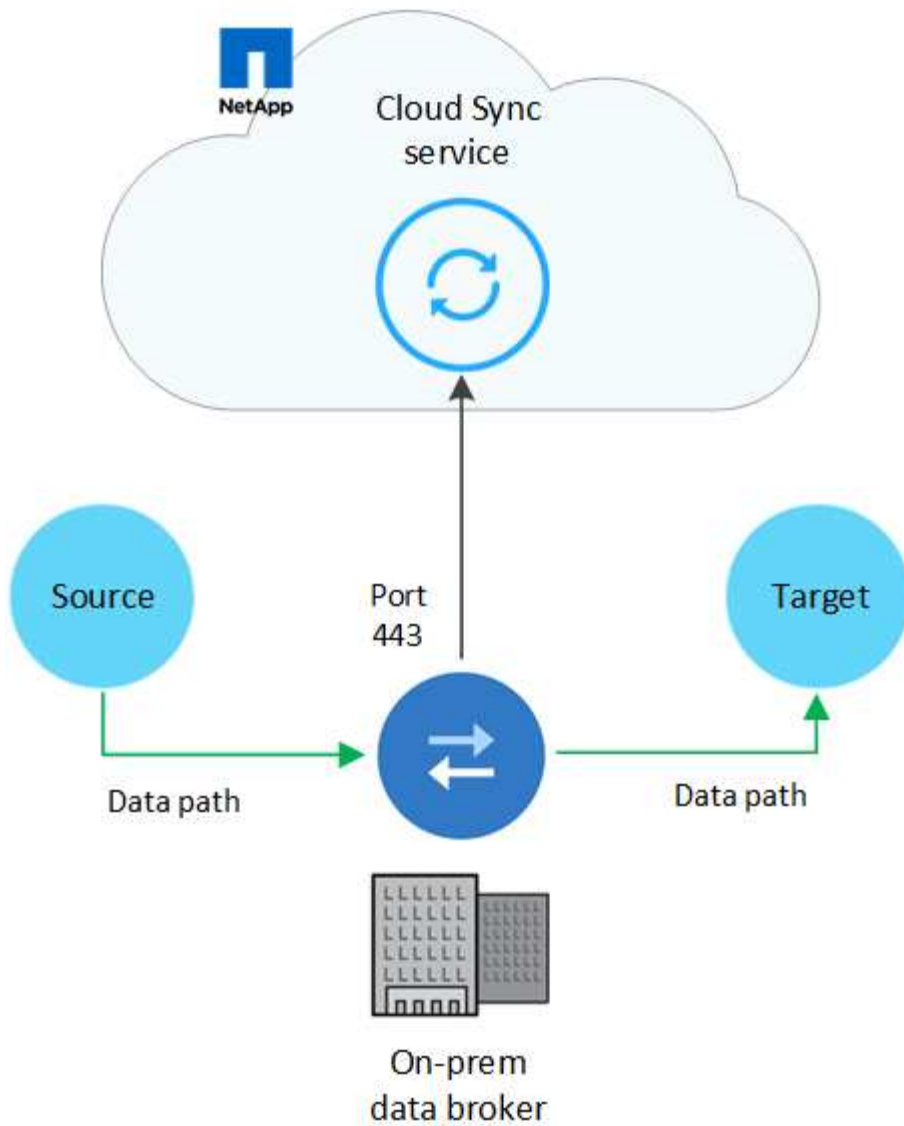
云中的数据代理

下图显示了在云中运行的数据代理，该代理可以在 AWS ， Google Cloud 或 Azure 中运行。只要与数据代理建立连接，源和目标就可以位于任何位置。例如，您可能会从数据中心连接到云提供商。



数据代理

下图显示了在数据中心内运行在 Prem 上的数据代理。同样，只要与数据代理建立连接，源和目标就可以位于任何位置。



网络要求

- 源和目标必须与数据代理组建立网络连接。

例如，如果数据中心中有 NFS 服务器，而 AWS 中有数据代理，则需要从网络到 VPC 的网络连接（VPN 或 Direct Connect）。

- 数据代理需要出站 Internet 连接，以便可以通过端口 443 轮询 Cloud Sync 服务以查找任务。
- NetApp 建议配置源，目标和数据代理以使用网络时间协议（NTP）服务。三个组件之间的时间差不应超过 5 分钟。

网络端点

NetApp 数据代理需要通过端口 443 进行出站 Internet 访问、以便与 Cloud Sync 服务进行通信、并与其他一些服务和存储库联系。您的本地 Web 浏览器还需要访问端点才能执行某些操作。如果需要限制出站连接、请在为出站通信配置防火墙时参考以下端点列表。

数据代理端点

数据代理会联系以下端点：

端点	目的
https://olcentgbl.trafficmanager.net	可与存储库联系以更新数据代理主机的 CentOS 软件包。仅当您在 CentOS 主机上手动安装数据代理时，才会联系此端点。
https://rpm.nodesource.com https://registry.npmjs.org https://nodejs.org :	与存储库联系以更新节点 .js 、 NPM 和开发中使用的其他第三方软件包。
https://tgz.pm2.io	访问用于更新 PM2 的存储库、该存储库是用于监控云同步的第三方软件包。
https://sqs.us-east-1.amazonaws.com https://kinesis.us-east-1.amazonaws.com	联系 Cloud Sync 用于操作的 AWS 服务（对文件进行排队、注册操作以及向数据代理提供更新）。
https://s3.region.amazonaws.com ， 例如： s3.us-east-2.amazonaws.com:443 https://docs.aws.amazon.com/general/latest/gr/rande.html#s3_region ["有关 S3 端点的列表，请参见 AWS 文档"^]	在同步关系包括 S3 存储区时联系 Amazon S3 。
https://s3.us-east-1.amazonaws.com	从 Cloud Sync 下载数据代理日志时，数据代理会将其日志目录置于 Z 形结构中，并将日志上传到 us-east-1 区域的预定义 S3 存储分段。
https://cf.cloudsync.netapp.com https://repo.cloudsync.netapp.com	可与 Cloud Sync 服务联系。
https://support.netapp.com	在使用 BYOL 许可证进行同步关系时联系 NetApp 支持。
https://fedoraproject.org	在安装和更新期间在数据代理虚拟机上安装 7z 。需要 7z 才能向 NetApp 技术支持发送 AutoSupport 消息。
https://sts.amazonaws.com	在 AWS 中部署数据代理或在内部部署数据代理时验证 AWS 凭据，并提供 AWS 凭据。数据代理会在部署期间，更新时以及重新启动时联系此端点。
https://cloudmanager.cloud.netapp.com https://netapp-cloud-account.auth0.com	在使用 Data sense 为新同步关系选择源文件时联系 Cloud Data sense 。

Web 浏览器端点

您的 Web 浏览器需要访问以下端点才能下载日志以进行故障排除：

logs.cloudsync.netapp.com:443

安装数据代理

在 AWS 中创建新的数据代理

创建新的数据代理组时，请选择 Amazon Web Services 在 VPC 中的新 EC2 实例上部署

数据代理软件。Cloud Sync 可指导您完成安装过程、但本页中重复了这些要求和步骤以帮助您做好安装准备。

您还可以选择在云中的现有 Linux 主机或您的企业中安装数据代理。"[了解更多信息](#)"。

支持的 **AWS** 区域

除中国地区外，所有地区均受支持。

网络要求

- 数据代理需要出站 Internet 连接、因此可以通过端口 443 轮询 Cloud Sync 服务以了解任务。

当 Cloud Sync 在 AWS 中部署数据代理时，它会创建一个安全组来启用所需的出站通信。请注意，您可以在安装过程中将数据代理配置为使用代理服务器。

如果需要限制出站连接，请参见 "[数据代理所联系的端点的列表](#)"。

- NetApp 建议将源、目标和数据代理配置为使用网络时间协议（NTP）服务。三个组件之间的时间差不应超过 5 分钟。

在 **AWS** 中部署数据代理所需的权限

用于部署数据代理的 AWS 用户帐户必须具有中包含的权限 "[此 NetApp 提供的策略](#)"。

在 **AWS** 数据代理中使用您自己的 **IAM** 角色的要求

当 Cloud Sync 部署数据代理时、它会为数据代理实例创建 IAM 角色。如果您愿意，可以使用自己的 IAM 角色部署数据代理。如果您的组织具有严格的安全策略，则可以使用此选项。

IAM 角色必须满足以下要求：

- 必须允许 EC2 服务作为受信任实体承担 IAM 角色。
- "[此 JSON 文件中定义的权限](#)" 必须附加到 IAM 角色，以便数据代理可以正常运行。

在部署数据代理时，请按照以下步骤指定 IAM 角色。

创建数据代理

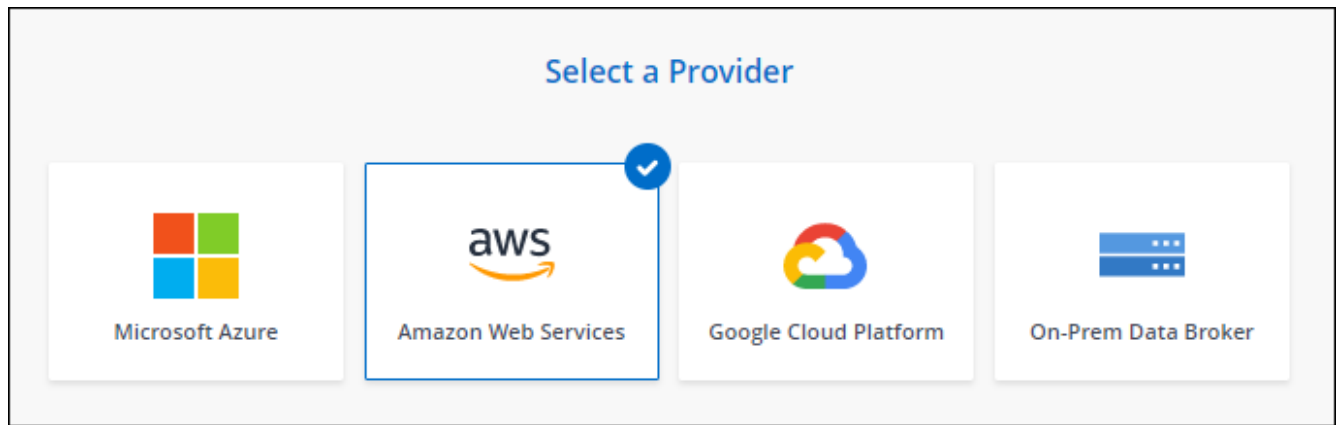
可以通过几种方法创建新的数据代理。以下步骤介绍如何在创建同步关系时在 AWS 中安装数据代理。

步骤

1. 单击 * 创建新同步 *。
2. 在 * 定义同步关系 * 页面上，选择一个源和目标，然后单击 * 继续 *。

完成这些步骤，直到显示 * 数据代理组 * 页面为止。

3. 在 * 数据代理组 * 页面上，单击 * 创建数据代理 *，然后选择 * Amazon Web Services*。



4. 输入数据代理的名称，然后单击 * 继续 *。
5. 输入 AWS 访问密钥，以便 Cloud Sync 可以代表您在 AWS 中创建数据代理。

这些密钥不会保存或用于任何其他目的。

如果您不想提供访问密钥，请单击页面底部的链接以改用 CloudFormation 模板。使用此选项时，您无需提供凭据，因为您直接登录到 AWS。

以下视频显示了如何使用 CloudFormation 模板启动数据代理实例：

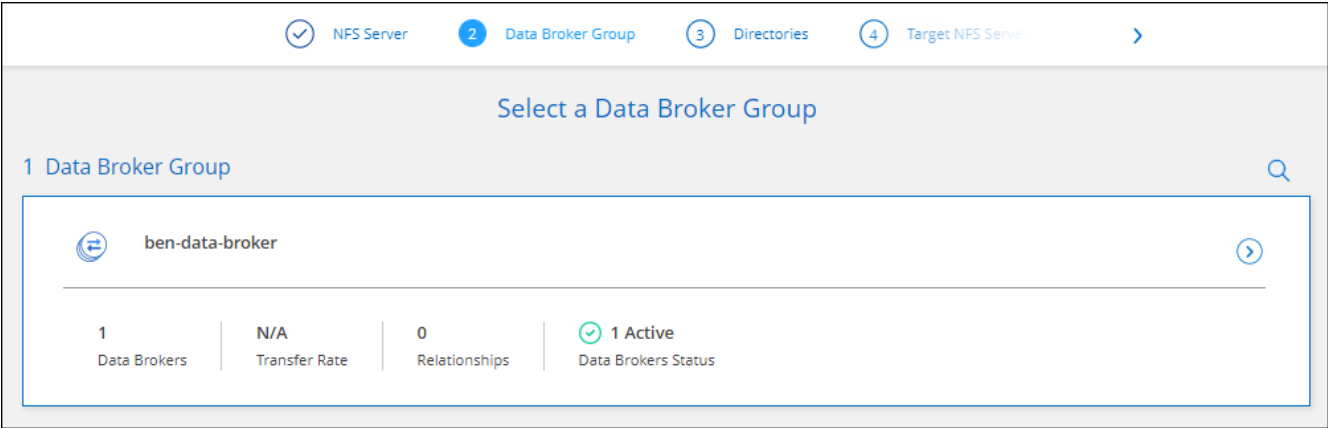
► https://docs.netapp.com/zh-cn/cloud-manager-sync//media/video_cloud_sync.mp4 (video)

6. 如果您输入了 AWS 访问密钥，请为实例选择一个位置，选择一个密钥对，选择是否启用公有 IP 地址，然后选择现有的 IAM 角色，或者将此字段留空，以便 Cloud Sync 为您创建角色。

如果您选择自己的 IAM 角色，[您需要提供所需的权限](#)。

7. 如果 VPC 中的 Internet 访问需要代理，请指定代理配置。
8. 数据代理可用后，单击 Cloud Sync 中的 * 继续 *。

下图显示了 AWS 中已成功部署的实例：



9. 完成向导中的页面以创建新的同步关系。

您已在 AWS 中部署了数据代理并创建了新的同步关系。您可以将此数据代理组与其他同步关系结合使用。

有关数据代理实例的详细信息

Cloud Sync 使用以下配置在 AWS 中创建数据代理。

Instance type

m5n.xlarge （如果在区域中可用），否则为 m5.xlarge

vCPU

4.

RAM

16 GB

操作系统

Amazon Linux 2.

磁盘大小和类型

10 GB GP2 SSD

在 **Azure** 中创建新的数据代理

创建新的数据代理组时，请选择 Microsoft Azure 以在 vNet 中的新虚拟机上部署数据代理软件。Cloud Sync 可指导您完成安装过程、但本页中重复了这些要求和步骤以帮助您做好安装准备。

您还可以选择在云中的现有 Linux 主机或您的企业中安装数据代理。"[了解更多信息。](#)"。

支持的 **Azure** 区域

除中国、美国政府和美国国防部地区外，所有地区均受支持。

网络要求

- 数据代理需要出站 Internet 连接、因此可以通过端口 443 轮询 Cloud Sync 服务以了解任务。

当 Cloud Sync 在 Azure 中部署数据代理时，它会创建一个安全组来启用所需的出站通信。

如果需要限制出站连接，请参见 ["数据代理所联系的端点的列表"](#)。

- NetApp 建议将源、目标和数据代理配置为使用网络时间协议（NTP）服务。三个组件之间的时间差不应超过 5 分钟。

身份验证方法

部署数据代理时，您需要选择一种身份验证方法：密码或 SSH 公共 - 私有密钥对。

有关创建密钥对的帮助，请参见 ["Azure 文档：在 Azure 中为 Linux VM 创建和使用 SSH 公共 - 私有密钥对"](#)。

创建数据代理

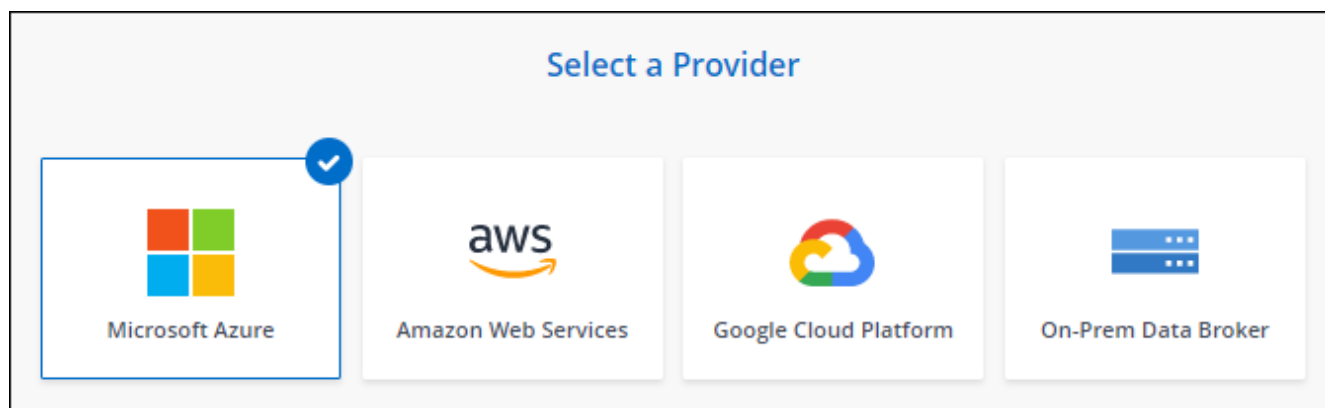
可以通过几种方法创建新的数据代理。以下步骤介绍如何在创建同步关系时在 Azure 中安装数据代理。

步骤

1. 单击 * 创建新同步 *。
2. 在 * 定义同步关系 * 页面上，选择一个源和目标，然后单击 * 继续 *。

完成这些步骤，直到显示 * 数据代理组 * 页面为止。

3. 在 * 数据代理组 * 页面上，单击 * 创建数据代理 *，然后选择 * Microsoft Azure*。



4. 输入数据代理的名称，然后单击 * 继续 *。
5. 如果出现提示，请登录到您的 Microsoft 帐户。如果未出现提示，请单击 * 登录到 Azure*。

此表由 Microsoft 拥有和托管。您的凭据不会提供给 NetApp。

6. 为数据代理选择一个位置，然后输入有关虚拟机的基本详细信息。

Location	Virtual Machine
<p>Subscription</p> <p>OCCM Dev ▼</p>	<p>VM Name</p> <p>netappdatabroker</p>
<p>Azure Region</p> <p>West US 2 ▼</p>	<p>User Name</p> <p>databroker</p>
<p>VNet</p> <p>Vnet1 ▼</p>	<p>Authentication Method:</p> <p><input checked="" type="radio"/> Password <input type="radio"/> Public Key</p>
<p>Subnet</p> <p>Subnet1 ▼</p>	<p>Enter Password</p> <p>.....</p>
	<p>Resource Group:</p> <p><input checked="" type="radio"/> Generate a new group <input type="radio"/> Use an existing group</p>

7. 如果在 vNet 中需要代理才能访问 Internet ，请指定代理配置。

8. 单击 * 继续 * 并保持此页面打开，直到部署完成。

此过程可能需要长达 7 分钟。

9. 在 Cloud Sync 中，一旦数据代理可用，请单击 * 继续 * 。

10. 完成向导中的页面以创建新的同步关系。

您已在 Azure 中部署了数据代理并创建了新的同步关系。您可以将此数据代理与其他同步关系一起使用。

是否收到有关需要管理员同意的消息？

如果 Microsoft 通知您需要管理员批准，因为 Cloud Sync 需要您的权限来代表您访问您组织中的资源，则您有两种选择：

1. 请您的 AD 管理员为您提供以下权限：

在 Azure 中，转到 * 管理中心 > Azure AD > 用户和组 > 用户设置 * 并启用 * 用户可以同意应用程序代表其访问公司数据 *。

2. 请您的 AD 管理员使用以下 URL 代表您同意使用 * CloudSync-AzureDataBrokerCreator*（这是管理员同意的端点）：

[https://login.microsoftonline.com/{FILL 此处为您的租户 ID } /v2.0/adminconsent?client_id=8ee4ca3a-BAFA-4831-97CC-5a38923cab85&redirect_uri =
https://cloudsync.netapp.com&scope=https://management.azure.com/user_impersonationhttps://graph.microsoft.com/User.Read](https://login.microsoftonline.com/{FILL 此处为您的租户 ID } /v2.0/adminconsent?client_id=8ee4ca3a-BAFA-4831-97CC-5a38923cab85&redirect_uri=https://cloudsync.netapp.com&scope=https://management.azure.com/user_impersonationhttps://graph.microsoft.com/User.Read)

如 URL 中所示，我们的应用程序 URL 为 <https://cloudsync.netapp.com>，应用程序客户端 ID 为 8ee4ca3a-BAFA-4831-97CC-5a38923cab85。

有关数据代理 VM 的详细信息

Cloud Sync 使用以下配置在 Azure 中创建数据代理。

VM 类型

标准 DS4 v2

vCPU

8.

RAM

28 GB

操作系统

CentOS 7.7

磁盘大小和类型

64 GB 高级 SSD

在 Google Cloud 中创建新的数据代理

创建新的数据代理组时，请选择 Google Cloud Platform 在 Google Cloud VPC 中的新虚拟机实例上部署数据代理软件。Cloud Sync 可指导您完成安装过程、但本页中重复了这些要求和步骤以帮助您做好安装准备。

您还可以选择在云中的现有 Linux 主机或您的企业中安装数据代理。 [了解更多信息。](#)

支持的 Google Cloud 地区

支持所有区域。

网络要求

- 数据代理需要出站 Internet 连接、因此可以通过端口 443 轮询 Cloud Sync 服务以了解任务。

当 Cloud Sync 在 Google Cloud 中部署数据代理时，它会创建一个安全组，以启用所需的出站通信。

如果需要限制出站连接，请参见 ["数据代理所联系的端点的列表"](#)。

- NetApp 建议将源、目标和数据代理配置为使用网络时间协议（NTP）服务。三个组件之间的时间差不应超过 5 分钟。

在 Google Cloud 中部署数据代理所需的权限

确保部署数据代理的 Google Cloud 用户具有以下权限：

```
- compute.networks.list
- compute.regions.list
- deploymentmanager.deployments.create
- deploymentmanager.deployments.delete
- deploymentmanager.operations.get
- iam.serviceAccounts.list
```

服务帐户所需的权限

部署数据代理时，您需要选择具有以下权限的服务帐户：

```
- logging.logEntries.create
- resourceManager.projects.get
- storage.buckets.get
- storage.buckets.list
- storage.objects.*
- iam.serviceAccounts.signJwt
```



只有在计划设置数据代理以使用外部 HashiCorp 存储时，才需要 iam.serviceAccounts.signJwt" 权限。

创建数据代理

可以通过几种方法创建新的数据代理。以下步骤介绍如何在创建同步关系时在 Google Cloud 中安装数据代理。

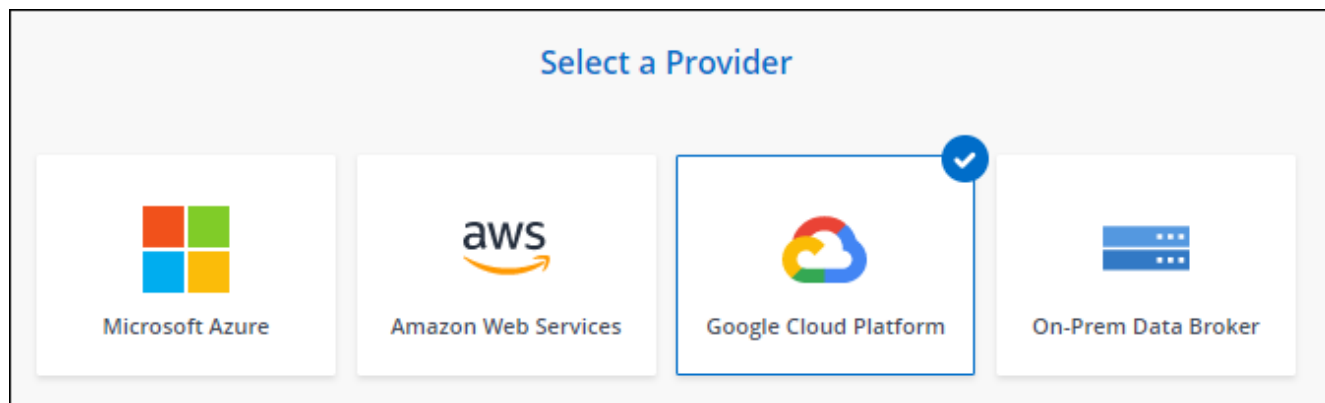
步骤

1. 单击 * 创建新同步 *。

2. 在 * 定义同步关系 * 页面上，选择一个源和目标，然后单击 * 继续 * 。

完成这些步骤，直到显示 * 数据代理组 * 页面为止。

3. 在 * 数据代理组 * 页面上，单击 * 创建数据代理 * ，然后选择 * Microsoft Azure* 。



4. 输入数据代理的名称，然后单击 * 继续 * 。

5. 如果出现提示，请使用 Google 帐户登录。

此表由 Google 拥有和托管。您的凭据不会提供给 NetApp 。

6. 选择项目和服务帐户，然后为数据代理选择一个位置，包括是否要启用或禁用公有 IP 地址。

如果不启用公有 IP 地址，则需要在下一步中定义代理服务器。

Basic Settings

Project

Project

OCCM-Dev

Location

Region

us-west1

Service Account

test

Zone

us-west1-a

Select a Service Account that includes [these permissions](#)

VPC

default

Subnet

default

Public IP

Enable

7. 如果 VPC 中的 Internet 访问需要代理，请指定代理配置。

如果需要代理才能访问 Internet，则该代理必须位于 Google Cloud 中，并使用与数据代理相同的服务帐户。

8. 数据代理可用后，单击 Cloud Sync 中的 * 继续 *。

部署该实例大约需要 5 到 10 分钟。您可以从 Cloud Sync 服务监控进度、该服务在实例可用时自动刷新。

9. 完成向导中的页面以创建新的同步关系。

您已在 Google Cloud 中部署数据代理并创建了新的同步关系。您可以将此数据代理与其他同步关系一起使用。

提供在其他 **Google Cloud** 项目中使用存储分段的权限

在创建同步关系并选择 Google Cloud Storage 作为源或目标时，您可以通过 Cloud Sync 从数据代理的服务帐户有权使用的分段中进行选择。默认情况下，此值包括与数据代理服务帐户位于 *same* 项目中的分段。但是，如果您提供了所需的权限，则可以从 *_other* 项目中选择分段。

步骤

1. 打开 Google Cloud Platform 控制台并加载云存储服务。
2. 单击要在同步关系中用作源或目标的存储分段的名称。
3. 单击 * 权限 *。

4. 单击 * 添加 *。
5. 输入数据代理的服务帐户的名称。
6. 选择提供的角色 [与上述权限相同](#)。
7. 单击 * 保存 *。

设置同步关系时，您现在可以选择该存储分段作为同步关系中的源或目标。

有关数据代理 **VM** 实例的详细信息

Cloud Sync 使用以下配置在 Google Cloud 中创建数据代理。

计算机类型

N1-standard-4

vCPU

4.

RAM

15 GB

操作系统

Red Hat Enterprise Linux 7.7

磁盘大小和类型

标准配置 20 GB HDD PD

在 **Linux** 主机上安装数据代理

创建新的数据代理组时，请选择 On-Prem Data Broker 选项，以便在内部 Linux 主机或云中的现有 Linux 主机上安装数据代理软件。Cloud Sync 可指导您完成安装过程、但本页中重复了这些要求和步骤以帮助您做好安装准备。

Linux 主机要求

- * 操作系统 * :
 - CentOS 7.0 , 7.7 和 8.0
 - 不支持 CentOS 流。
 - Red Hat Enterprise Linux 7.7 和 8.0
 - Ubuntu Server 20.04 LTS
 - SUSE Linux Enterprise Server 15 SP1

安装数据代理之前，必须在主机上运行命令 `yum update all`。

Red Hat Enterprise Linux 系统必须在 Red Hat 订购管理中注册。如果未注册、系统将无法在安装期间访问存储库以更新所需的第三方软件。

- * RAM * : 16 GB
- * CPU * : 4 个核心
- * 可用磁盘空间 * : 10 GB
- * SELinux* : 建议您禁用 "SELinux" 在主机上。

SELinux 会强制实施一种策略，该策略会阻止数据代理软件更新，并阻止数据代理与正常运行所需的端点联系。

网络要求

- Linux 主机必须与源主机和目标主机建立连接。
- 文件服务器必须允许 Linux 主机访问导出。
- 必须在 Linux 主机上打开端口 443，以便向 AWS 发送出站流量（数据代理会不断与 Amazon SQS 服务进行通信）。
- NetApp 建议将源、目标和数据代理配置为使用网络时间协议（NTP）服务。三个组件之间的时间差不应超过 5 分钟。

启用对 **AWS** 的访问

如果您计划使用包含 S3 存储分段的同步关系中的数据代理，则应准备 Linux 主机以进行 AWS 访问。安装数据代理时，您需要为具有编程访问权限和特定权限的 AWS 用户提供 AWS 密钥。

步骤

1. 使用创建 IAM 策略 ["此 NetApp 提供的策略"](#)

["查看 AWS 说明。"](#)

2. 创建具有编程访问权限的 IAM 用户。

["查看 AWS 说明。"](#)

请务必复制 AWS 密钥，因为您需要在安装数据代理软件时指定这些密钥。

启用对 **Google Cloud** 的访问

如果您计划将数据代理与包含 Google Cloud 存储分段的同步关系结合使用，则应准备 Linux 主机以进行 Google Cloud 访问。安装数据代理时，您需要为具有特定权限的服务帐户提供一个密钥。

步骤

1. 如果您还没有存储管理员权限，请创建一个 Google Cloud 服务帐户。
2. 创建以 JSON 格式保存的服务帐户密钥。

["查看 Google Cloud 说明"](#)

此文件应至少包含以下属性："project_id"，"private_key" 和 "client_email"



创建密钥时、文件将生成并下载到计算机中。

3. 将 JSON 文件保存到 Linux 主机。

启用对 **Microsoft Azure** 的访问

通过在 "同步关系" 向导中提供存储帐户和连接字符串，可以按关系定义对 Azure 的访问。

安装数据代理

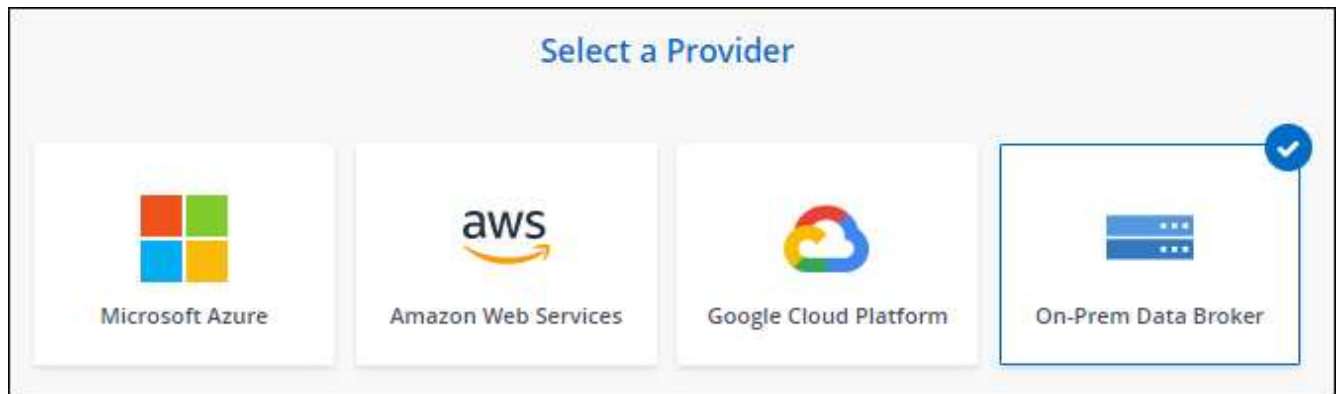
您可以在创建同步关系时在 Linux 主机上安装数据代理。

步骤

1. 单击 * 创建新同步 *。
2. 在 * 定义同步关系 * 页面上，选择一个源和目标，然后单击 * 继续 *。

完成这些步骤，直到显示 * 数据代理组 * 页面为止。

3. 在 * 数据代理组 * 页面上，单击 * 创建数据代理 *，然后选择 * 本地数据代理 *。



即使此选项标记为 * ; on-Prem_ Data Broker*，它也会对内部或云中的 Linux 主机进行适用场景。

4. 输入数据代理的名称，然后单击 * 继续 *。

此时将很快加载说明页面。您需要按照这些说明进行操作，其中包括一个用于下载安装程序的唯一链接。

5. 在说明页面上：
 - a. 选择是启用对 * AWS *，* Google Cloud* 的访问，还是同时启用这两者的访问。
 - b. 选择一个安装选项：* 无代理 *，* 使用代理服务器 * 或 * 使用代理服务器进行身份验证 *。
 - c. 使用命令下载并安装数据代理。

以下步骤提供了有关每个可能安装选项的详细信息。按照说明页面根据您的安装选项获取确切的命令。

- d. 下载安装程序：

- 无代理：

```
cURL <URI> -o data_broker_installer.sh
```

- 使用代理服务器：

```
curl <URI> -o data_broker_installer.sh -x <proxy_host> : <proxy_port>
```

- 使用具有身份验证的代理服务器：

```
cURL <URI> -o data_broker_installer.sh -x <proxy_username> :  
<proxy_password>@ <proxy_host> : <proxy_port>
```

URI

Cloud Sync 会在说明页面上显示安装文件的 URI，在按照提示部署内部数据代理时会加载此 URI。此 URI 不会在此重复，因为此链接是动态生成的，只能使用一次。[按照以下步骤从 Cloud Sync 获取 URI。](#)

- e. 切换到超级用户，使安装程序可执行并安装软件：



下面列出的每个命令都包含 AWS 访问和 Google Cloud 访问的参数。按照说明页面根据您的安装选项获取确切的命令。

- 无代理配置：

```
sudo -s chmod +x data_broker_installer.sh ./data_broker_installer.sh -a  
<AWS_access_key> -s <AWS_secret_key> -g <abasal_path_to_the_json_file>
```

- 代理配置：

```
sudo -s chmod +x data_broker_installer.sh ./data_broker_installer.sh -a  
<AWS_access_key> -s <AWS_secret_key> -g <abasal_path_to_the_json_file> -h  
<proxy_host> -p <proxy_port>
```

- 具有身份验证的代理配置：

```
sudo -s chmod +x data_broker_installer.sh ./data_broker_installer.sh -a  
<AWS_access_key> -s <AWS_secret_key> -g <abasal_path_to_the_json_file> -h  
<proxy_host> -p <proxy_port> -u <proxy_username> -w <proxy_password>
```

AWS 密钥

这些是您应准备好的用户密钥 [请按照以下步骤操作](#)。AWS 密钥存储在数据代理上，该代理在内部或云网络中运行。NetApp 不会在数据代理之外使用密钥。

JSON 文件

此 JSON 文件包含您应已准备好的服务帐户密钥 [请按照以下步骤操作](#)。

6. 数据代理可用后，单击 Cloud Sync 中的 * 继续 *。
7. 完成向导中的页面以创建新的同步关系。

Copyright Information

Copyright © 2022 NetApp, Inc. All rights reserved. Printed in the U.S. No part of this document covered by copyright may be reproduced in any form or by any means-graphic, electronic, or mechanical, including photocopying, recording, taping, or storage in an electronic retrieval system-without prior written permission of the copyright owner.

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

THIS SOFTWARE IS PROVIDED BY NETAPP "AS IS" AND WITHOUT ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE, WHICH ARE HEREBY DISCLAIMED. IN NO EVENT SHALL NETAPP BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

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

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

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).

Trademark Information

NETAPP, the NETAPP logo, and the marks listed at <http://www.netapp.com/TM> are trademarks of NetApp, Inc. Other company and product names may be trademarks of their respective owners.