



用户指南

# AWS DataSync



# AWS DataSync: 用户指南

Copyright © 2023 Amazon Web Services, Inc. and/or its affiliates. All rights reserved.

Amazon 的商标和商业外观不得用于任何非 Amazon 的商品或服务，也不得以任何可能引起客户混淆、贬低或诋毁 Amazon 的方式使用。所有非 Amazon 拥有的其它商标均为各自所有者的财产，这些所有者可能附属于 Amazon、与 Amazon 有关联或由 Amazon 赞助，也可能不是如此。

# Table of Contents

什么是 AWS DataSync ? .....	1
使用案例 .....	1
优点 .....	2
其他资源 .....	2
工作原理 .....	3
发现 .....	3
DataSync探索架构 .....	3
概念和术语 .....	4
限制 .....	5
数据传输 .....	5
DataSync 架构 .....	5
概念和术语 .....	8
如何DataSync传输文件和对象 .....	9
要求 .....	11
代理要求 .....	11
找一个代理 .....	11
虚拟机管理程序要求 .....	11
代理人对DataSync转账的要求 .....	12
DataSyncDiscovery 的代理要求 .....	13
代理管理要求 .....	13
网络要求 .....	13
自管理存储系统的网络要求 .....	13
AWS存储服务的网络要求 .....	16
VPC 终端节点的网络要求 .....	16
公共或 FIPS 端点的网络要求 .....	19
网络接口要求 .....	25
所需的权限 .....	27
AWS 托管策略 .....	28
客户管理型策略 .....	28
开始使用 .....	29
设置 .....	29
注册一个 AWS 账户 .....	29
创建管理用户 .....	30
我可以在哪里使用DataSync ? .....	30

我如何如何如何如何如何如何如何如何DataSync如何如何 .....	30
DataSync费用是多少? .....	31
登录到 控制台 .....	31
创建一个代理 .....	31
部署您的代理 .....	32
选择服务端点 .....	37
激活您的代理 .....	39
探索您的存储空间 .....	41
添加您的本地存储系统 .....	41
开始你的探索工作 .....	42
传输您的数据 .....	42
创建源位置 .....	43
创建目标位置 .....	44
创建并开始任务 .....	45
清理资源 .....	46
与代理合作 .....	47
创建一个代理 .....	47
将代理VPC 于 .....	47
DataSync代理如何使用 VPC 终端节点 .....	48
DataSyncVPC 的局限性 .....	48
将您的DataSync代理配置为使用 VPC 终端节点 .....	48
将您的代理部署在 AWS 区域 .....	52
将数据从云文件系统传输到另一个云文件系统或 Amazon S3 .....	52
将数据从 Amazon S3 AWS .....	53
使用多个代理进行转账 .....	54
为多个 NIC 配置代理 .....	54
管理您的代理 .....	55
代理软件更新 .....	55
代理故障排除 .....	55
编辑您的代理的属性 .....	55
代理状态 .....	56
使用代理的本地控制台 .....	56
访问代理的本地控制台 .....	57
获取代理激活密钥 .....	57
配置您的代理网络设置 .....	58
测试您的代理与的连接AWS .....	61

测试代理与存储系统的连接 .....	62
查看代理的系统资源状态 .....	62
同步您的 VMware 代理上的时间 .....	63
配置其他代理设置 .....	64
通过以下方式向您的代理人寻求帮助AWS Support .....	65
删除该代理 .....	67
发现您的存储 .....	68
添加您的本地存储系统 .....	68
访问您的本地存储系统 .....	68
添加您的本地存储系统 .....	69
移除您的本地存储系统 .....	71
将DataSync发现活动记录到亚马逊 CloudWatch .....	72
从事探索工作 .....	72
开始探索任务 .....	73
停止发现作业 .....	74
查看存储资源信息 .....	74
查看收集的有关您的存储系统的信息 .....	75
获取建议 .....	78
建议中包含什么？ .....	78
建议中未包含什么？ .....	78
获取建议 .....	79
DataSync发现状态 .....	81
Discover 状态 .....	81
推荐状态 .....	82
传输您的数据 .....	84
我可以在哪里传输我的数据？ .....	84
支持在同一时间传输AWS 账户 .....	85
支持跨境传输AWS 账户 .....	86
支持在同一时间传输AWS 区域 .....	87
支持跨境传输AWS 区域 .....	87
从本地存储转移 .....	88
配置来自 NFS 文件服务器的传输 .....	88
配置来自 SMB 文件服务器的传输 .....	89
配置来自 HDFS 的传输 .....	92
配置来自对象存储系统的传输 .....	94
传输到存储器或从AWS存储器传出 .....	96

使用 Amazon S3 配置传输 .....	97
使用亚马逊 EFS 配置传输 .....	110
使用 FSx for Windows 文件服务器配置传输 .....	114
使用 FSx for Lustre 配置传输 .....	117
使用 FSx for OpenZFS 配置传输 .....	118
使用 FsX 为 ONTAP 配置传输 .....	120
向其他云存储传输或从其他云存储传输 .....	125
使用 Google Cloud Storage .....	125
使用 Azure Blob 存储配置传输 ( 预览 ) .....	133
使用微软 Azure 文件配置传输 .....	143
传输到边缘或离线存储器或从边缘存储器传输数据 .....	146
在 Snowball Edge 上使用兼容 S3 的存储器配置 .....	146
使用 Snowcone 配置传输 .....	151
如何DataSync处理元数据和特殊文件 .....	152
元数据复制者 DataSync .....	152
复制的链接和目录 DataSync .....	160
移除传送地点 .....	161
处理转移任务 .....	161
验证数据完整性 .....	162
配置哪些DataSync传输 .....	163
设置带宽限制 .....	172
安排任务 .....	173
记录任务活动 .....	175
为任务添加标签 .....	176
开始任务 .....	179
取消任务 .....	180
删除你的任务 .....	181
任务状态 .....	182
任务示例 .....	183
监控 .....	190
亚马逊的CloudWatch指标DataSync .....	190
的亚马逊EventBridge活动DataSync .....	192
DataSync转移事件 .....	192
DataSync探索活动 .....	194
允许DataSync将日志上传到CloudWatch日志组 .....	195
从命令行监控您的DataSync任务 .....	196

使用以下方法监控您的任务AWS CLI .....	197
使用该watch实用程序监控您的任务 .....	198
安全性 .....	199
数据保护 .....	199
传输中加密 .....	200
静态加密 .....	202
互连网络流量隐私 .....	203
身份和访问管理 .....	203
访问管理 .....	204
AWS 托管策略 .....	208
客户管理型策略 .....	211
使用服务相关角色 .....	214
在创建过程中标记资源 .....	217
防止跨服务混淆代理 .....	218
API 权限参考 .....	220
日志记录 .....	228
处理中的DataSync信息CloudTrail .....	229
了解 DataSync 日志文件条目 .....	230
合规性验证 .....	231
故障恢复能力 .....	232
基础设施安全性 .....	232
配额 .....	233
存储系统、文件和对象限制 .....	233
DataSync任务配额 .....	233
DataSync发现配额 .....	236
请求提高限额 .....	237
问题排查 .....	238
代理 .....	238
如何连接到 Amazon EC2 代理的本地控制台？ .....	238
无法检索代理激活密钥错误是什么意思？ .....	238
我无法使用 VPC 终端节点激活代理 .....	239
我不知道我的经纪人怎么了。有人能帮我吗？ .....	239
数据传输 .....	239
DataSync如何配置为使用特定的 NFS 或 SMB 版本来挂载我的文件共享？ .....	239
错误：SyncOption值无效。选项：TransferMode，PreserveDeletedFiles，值：全部，删除。 .....	240

我的任务一直因EniNotFound错误而失败 .....	240
我的任务因DataSync当前不支持服务器端 NFSv4 ID 映射错误而失败 .....	240
我的任务状态为不可用，表示存在装载错误 .....	241
我的任务失败，出现“无法分配内存”错误 .....	241
我的任务因输入/输出错误而失败 .....	242
我的任务执行处于启动状态，但似乎什么也没发生 .....	242
我的任务执行处于准备状态很长时间了 .....	243
我的任务因权限被拒绝错误而失败 .....	243
验证任务需要多长时间？ .....	243
转移到另一个 S3 存储桶时，我的任务失败AWS 账户 .....	244
从 Google 云存储分区传输时我的任务失败 .....	244
我的任务开始和结束时间与日志不符 .....	244
Error: SyncTaskDeletedByUser .....	245
Error: NoMem .....	245
Error: FsS3UnableToConnectToEndpoint .....	245
Error: FsS3HeadBucketFailed .....	245
成本 .....	245
教程 .....	247
从本地转移到另一个账户的 S3 .....	247
概览 .....	247
先决条件 .....	248
步骤 1：为账户 ADataSync 中创建 IAM 角色 .....	248
步骤 2：在账户 B 中禁用 S3 存储桶的 ACL .....	250
步骤 3：更新账户 B 中的 S3 存储桶策略 .....	250
步骤 4：为 S3 存储桶创建DataSync目标位置 .....	252
步骤 5：创建并启动DataSync任务 .....	253
相关资源 .....	253
在另一个账户中从 S3 转移到 S3 .....	254
概览 .....	254
先决条件 .....	256
步骤 1：DataSync在账户 A 中创建 IAM 角色 .....	256
步骤 2：在账户 B 中禁用 S3 存储桶的 ACL .....	258
步骤 3：更新账户 B 中的 S3 存储桶策略 .....	259
步骤 4：为 S3 存储桶创建DataSync目标位置 .....	260
步骤 5：创建并启动DataSync任务 .....	261
相关资源 .....	261



其他资源 .....	263
从自我管理的存储阵列传输数据 .....	263
其他使用案例 .....	263
朝相反的方向传输文件 .....	264
使用多个任务写入同一 Amazon S3 存储桶 .....	264
允许 DataSync 访问受限的 Amazon S3 存储桶 .....	265
开源组件 .....	266
使用 AWS CLI .....	267
创建代理 .....	268
创建地点 .....	271
创建 NFS 位置 .....	272
创建 SMB 位置 .....	273
创建 HDFS 位置 .....	274
创建对象存储位置 .....	275
创建 Amazon EFS 位置 .....	276
创建 Amazon FSx for Windows File Server 位置 .....	278
创建 Amazon FSx for Lustre 位置 .....	279
创建 Amazon FSx for OpenZFS 位置 .....	280
创建 Amazon FSx for NetApp ONTAP 位置 .....	280
创建 Amazon S3 位置 .....	282
创建任务 .....	287
启动任务 .....	289
筛选资源 .....	289
用于筛选的参数 .....	290
按位置筛选 .....	291
按任务筛选 .....	292
DataSync API .....	293
操作 .....	293
AddStorageSystem .....	296
CancelTaskExecution .....	302
CreateAgent .....	304
CreateLocationEfs .....	309
CreateLocationFsxLustre .....	315
CreateLocationFsxOntap .....	319
CreateLocationFsxOpenZfs .....	323
CreateLocationFsxWindows .....	327

CreateLocationHdfs .....	332
CreateLocationNfs .....	339
CreateLocationObjectStorage .....	344
CreateLocationS3 .....	349
CreateLocationSmb .....	356
CreateTask .....	362
DeleteAgent .....	369
DeleteLocation .....	371
DeleteTask .....	373
DescribeAgent .....	375
DescribeDiscoveryJob .....	379
DescribeLocationEfs .....	383
DescribeLocationFsxLustre .....	388
DescribeLocationFsxOntap .....	391
DescribeLocationFsxOpenZfs .....	395
DescribeLocationFsxWindows .....	399
DescribeLocationHdfs .....	403
DescribeLocationNfs .....	408
DescribeLocationObjectStorage .....	412
DescribeLocationS3 .....	416
DescribeLocationSmb .....	420
DescribeStorageSystem .....	425
DescribeStorageSystemResourceMetrics .....	430
DescribeStorageSystemResources .....	436
DescribeTask .....	446
DescribeTaskExecution .....	453
GenerateRecommendations .....	460
ListAgents .....	463
ListDiscoveryJobs .....	466
ListLocations .....	469
ListStorageSystems .....	472
ListTagsForResource .....	475
ListTaskExecutions .....	478
ListTasks .....	481
RemoveStorageSystem .....	484
StartDiscoveryJob .....	486

StartTaskExecution .....	490
StopDiscoveryJob .....	495
TagResource .....	497
UntagResource .....	500
UpdateAgent .....	502
UpdateDiscoveryJob .....	504
UpdateLocationHdfs .....	506
UpdateLocationNfs .....	511
UpdateLocationObjectStorage .....	514
UpdateLocationSmb .....	518
UpdateStorageSystem .....	522
UpdateTask .....	526
UpdateTaskExecution .....	530
数据类型 .....	532
AgentListEntry .....	534
Capacity .....	536
Credentials .....	538
DiscoveryJobListEntry .....	540
DiscoveryServerConfiguration .....	541
Ec2Config .....	542
FilterRule .....	544
FsxProtocol .....	546
FsxProtocolNfs .....	547
FsxProtocolSmb .....	548
HdfsNameNode .....	550
IOPS .....	551
Latency .....	553
LocationFilter .....	555
LocationListEntry .....	557
MaxP95Performance .....	559
NetAppONTAPCluster .....	562
NetAppONTAPSVM .....	565
NetAppONTAPVolume .....	569
NfsMountOptions .....	573
OnPremConfig .....	574
Options .....	575

---

P95Metrics .....	583
PrivateLinkConfig .....	584
QopConfiguration .....	586
Recommendation .....	587
ResourceDetails .....	589
ResourceMetrics .....	590
S3Config .....	592
SmbMountOptions .....	593
StorageSystemListEntry .....	595
TagListEntry .....	596
TaskExecutionListEntry .....	597
TaskExecutionResultDetail .....	598
TaskFilter .....	601
TaskListEntry .....	603
TaskSchedule .....	605
Throughput .....	606
常见错误 .....	607
常见参数 .....	609
文档历史记录 .....	612
AWS词汇表 .....	619
.....	dcxx

# 什么是 AWS DataSync ?

AWS DataSync是一项在线数据移动和发现服务，可简化数据迁移，并帮助您在AWS存储服务之间快速、轻松、安全地传输文件或对象数据。

DataSync使用以下存储系统和服务：

- [网络文件系统 \( NFS \)](#) 文件服务器
- [服务器消息块 \( SMB \)](#) 文件服务器
- [Hadoop Distributed File](#)
- [对象存储系统](#)
- [Amazon S3](#) 存储桶
- [亚马逊 EFS](#) 文件系统
- [Amazon FSx for Windows File Server](#) 文件系统
- [Amazon FSx for Lustre](#)
- [Amazon FSx for OpenZ](#) FS 文件系统
- [适用于 NetApp ONTAP 文件系统的亚马逊 FSx](#)
- [Good Storage](#)
- [微软 Azure Cloud Storage](#)
- [微软 Azure 文件](#)
- AWS Snowball Edge设备@@ [上兼容Amazon S3 的存储](#)
- [AWS Snowcone](#) 设备

## 使用案例

以下是 DataSync 的主要使用案例：

- 发现数据-了解您的本地存储性能和利用率。AWS DataSync 数据查找服务还可以提供有关将数据迁移到AWS存储服务的建议。
- 迁移数据-通过网络将活动数据集快速移动到AWS存储服务中。DataSync包括自动加密和数据完整性验证，以帮助确保您的数据安全、完好无损地到达并随时可用。
- 存档冷数据 — 将存储在本地存储中的冷数据直接移动到耐用且安全的长期存储类别，例如 S3 Glacier 灵活检索或 S3 Glacier Deep Archive。这样做可以释放本地存储容量并关闭旧系统。

- **复制数据**-将数据复制到任何 Amazon S3 存储类中，根据您的需求选择最具成本效益的存储类别。您还可以将数据发送到亚马逊 EFS、FSx for Windows File Server、适用于 Lustre 的 FsX 或备用文件系统的适用于 OpenZFS 的 FsX。
- **移动数据**以便及时在云端进行处理-将数据移入或移出AWS以进行处理。这种方法可加快许多行业内的关键混合云工作流。其中包括生命科学行业的机器学习、媒体和娱乐领域的视频制作、金融服务中的大数据分析以及石油和天然气行业的地震研究。

## 优点

通过使用 DataSync，您可以获得以下好处：

- **简化迁移规划**-通过自动数据收集和建议，DataSyncDiscovery 可以最大限度地减少与规划数据迁移到AWS相关的时间、工作量和成本。您可以使用建议为预算计划提供信息，并在迁移临近时重新运行发现任务以验证您的假设。
- **自动移动数据**-DataSync 使通过网络在存储系统和服务之间移动数据变得更加容易。DataSync自动管理数据传输过程和高性能和安全数据传输所需的基础架构。
- **安全传输数据**-DataSync 提供端到端安全性，包括加密和完整性验证，以帮助确保您的数据安全、完好无损地到达并随时可用。DataSync通过内置AWS安全机制（例如 AWS Identity and Access Management (IAM) 角色）访问您的AWS存储。它还支持虚拟私有云 (VPC) 终端节点，使您可以选择不穿越公共互联网的情况下传输数据，并进一步提高了在线复制的数据的安全性。
- **更快地移动数据**-DataSync 使用专门构建的网络协议和parallel 多线程架构来加快传输速度。这种方法可以加快迁移、用于分析和机器学习的重复数据处理工作流程以及数据保护流程。
- **降低运营成本** — 采用每千兆字节的固定定价，经济高效地移动数据。DataSync避免编写和维护自定义脚本或使用昂贵的商业传输工具。

## 其他资源

我们建议您阅读以下部分：

- [DataSync资源](#)-包括博客、视频和其他培训材料
- [AWSre: post](#) — 查看有关的最新讨论 DataSync
- [AWS DataSync 定价](#)

# AWS DataSync工作原理

直观地概述AWS DataSync工作原理并学习关键概念，以帮助您快速识别和移动数据。

主题

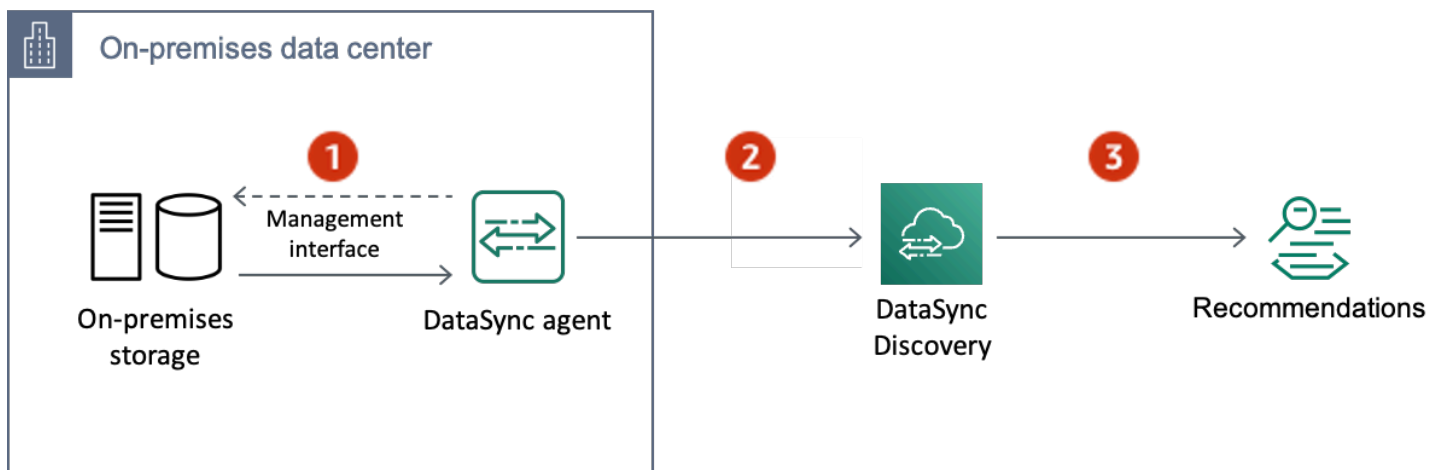
- [AWS DataSync 数据查找服务 的工作原理](#)
- [AWS DataSync转账是如何运作的](#)

## AWS DataSync 数据查找服务 的工作原理

学习与之相关的关键概念和术语AWS DataSync 数据查找服务。

### DataSync探索架构

下图说明了 DataSync Discovery 如何收集信息，并提供了将数据从本地存储系统迁移到的建议AWS。



参考	描述
1	DataSync代理连接到您的本地存储系统的管理接口（例如，使用端口 443）。然后，您运行一个发现任务来收集有关系统信息。
2	代理通过 <a href="#">公共服务端点</a> 将其收集的信息发送到 DataSync Discovery。





## 存储系统资源信息

DataSyncDiscovery 可以为您提供有关本地存储系统资源的性能和利用率信息。例如，了解特定存储卷中正在使用多少存储容量，而您最初预置的容量是多少。

您可以使用以下方法在发现任务收集此类信息时查看此类信息：

- DataSync 控制台
- 该[DescribeStorageSystemResources](#)手术
- 该[DescribeStorageSystemResourceMetrics](#)手术

有关更多信息，请参阅[查看收集的存储资源信息 AWS DataSync 数据查找服务](#)：

## AWS存储建议

DataSyncDiscovery 使用其收集的有关本地存储系统资源的信息，推荐AWS存储服务来帮助您规划迁移到AWS。

您可以使用以下方法查看推荐：

- DataSync 控制台
- 该[DescribeStorageSystemResources](#)手术

有关更多信息，请参阅[从中获得推荐 AWS DataSync 数据查找服务](#)：

## 限制

- 目前，您只能激活具有[公共服务端点](#)的 DataSync Discovery 代理。

## AWS DataSync转账是如何运作的

学习与AWS DataSync转账相关的关键概念和术语。

## DataSync 架构

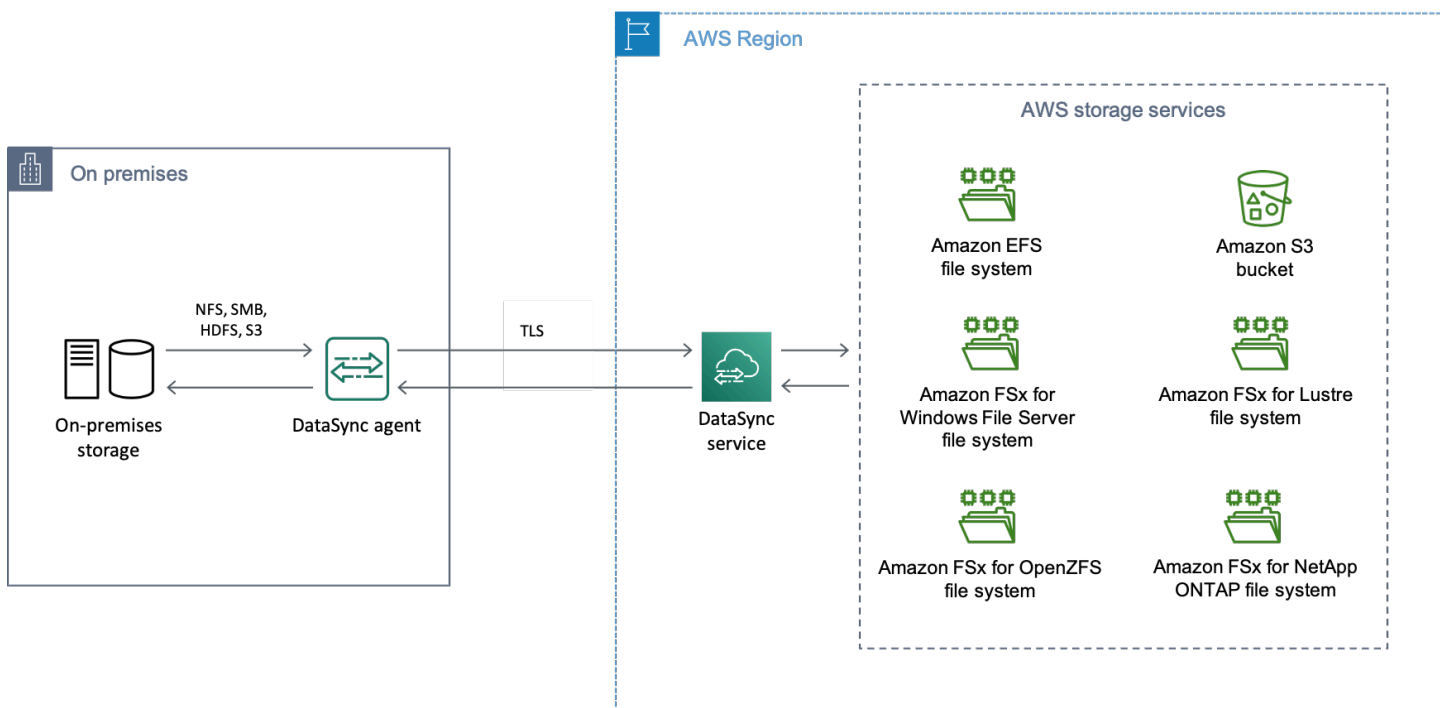
下图显示了DataSync通常如何以及在何处传输存储数据。有关DataSync支持的存储系统和服务的完整列表，请参见[我可以在哪里传输我的数据AWS DataSync？](#)。

## 主题

- [在本地存储和之间传输 AWS](#)
- [在AWS存储服务之间传输](#)
- [在云存储系统和AWS存储服务之间传输](#)

## 在本地存储和之间传输 AWS

下图高度概括在自管本地存储系统和AWS 服务之间DataSync传输文件的情况。

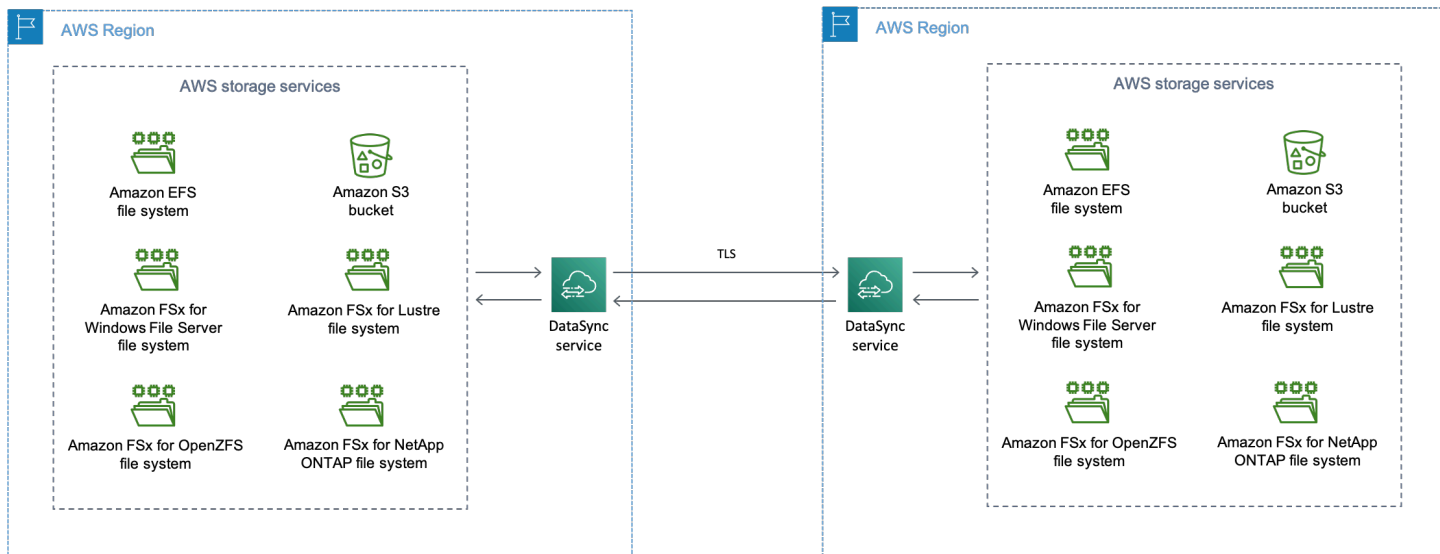


该图说明了一个常见的DataSync用例：

- DataSync代理从本地存储系统复制数据。
- 数据AWS通过传输层安全性 ( TLS ) 传入。
- DataSync将数据复制到支持的AWS存储服务。

## 在AWS存储服务之间传输

下图高度概括此处之间AWS 服务文件DataSync传输的高级概述AWS 账户。



该图说明了一个常见的DataSync用例：

- DataSync从支持的AWS存储服务复制数据。
- AWS 区域通过 TLS 传输数据。
- DataSync将数据复制到支持的AWS存储服务。

在AWS存储服务之间传输时（无论是同一服务AWS 区域还是跨存储服务AWS 区域），您的数据将保留在AWS网络中，不会通过公共互联网。

#### **⚠ Important**

您需要为在两者之间传输的数据付费AWS 区域。这按从您的来源区域向目标区域传输数据计费。有关更多信息，请参阅[数据传输定价](#)。

## 在云存储系统和AWS存储服务之间传输

使用DataSync，您可以在其他云存储系统和之间传输数据AWS 服务。在这种情况下，云存储系统可以包括：

- 由托管的自我管理存储系统AWS（例如，内部AWS虚拟私有云中的 NFS 共享）。
- 其他云提供商托管的存储系统或服务。

有关更多信息，请参阅：



## 任务执行

任务执行是DataSync传输任务的单独运行。任务执行涉及多个阶段。有关更多信息，请参阅[任务执行状态](#)：

## 如何DataSync传输文件和对象

开始传输时，DataSync检查您的源存储系统和目标存储系统以确定要同步的内容。它通过递归扫描两个系统的内容和元数据来识别两者之间的差异来做到这一点。这可能只需要几分钟或几个小时，具体取决于所涉及的文件或对象的数量（包括存储系统的性能）。

DataSync然后开始根据[您设置传输的方式](#)将数据（包括元数据）从源移动到目标。例如，在传输过程中DataSync始终执行数据完整性检查。传输完成后，DataSync还可以验证各个位置之间的整个数据集或仅验证您复制的数据。（在大多数情况下，我们建议仅验证传输的内容。）也有筛选要传输的内容的选项。

### 主题

- [如何DataSync验证数据完整性](#)
- [如何DataSync处理打开和锁定的文件](#)

## 如何DataSync验证数据完整性

DataSync在本地计算源和目标存储系统中每个文件或对象的校验和并进行比较。此外，DataSync比较源和目标中每个文件或对象的元数据。

如果两者之间存在差异，验证将失败，显示精确指定失败内容的错误代码。例如，你可能会看到Checksum failure、Metadata failureFiles were addedFiles were removed、等错误代码。

有关更多信息，请参阅[配置如何AWS DataSync验证数据完整性](#)：

## 如何DataSync处理打开和锁定的文件

尝试传输正在使用或锁定的文件时，请记住以下几点：

- 一般而言，DataSync可以传输打开的文件而没有任何限制。
- 如果文件在传输过程中处于打开状态并正在写入，则DataSync可以在传输任务的验证阶段检测到这种不一致性。要获取最新版本的文件必须再次运行任务。

- 如果文件被锁定并且服务器DataSync阻止其打开，则在传输过程中DataSync跳过该文件并记录错误。
- DataSync无法锁定或解锁文件。

# AWS DataSync 的要求

AWS DataSync代理和网络要求因您计划传输数据的位置和方式而异。

主题

- [AWS DataSync代理要求](#)
- [AWS DataSync网络要求](#)
- [使用所需的 IAM 权限AWS DataSync](#)

## AWS DataSync代理要求

代理是AWS DataSync用于存储发现和数据传输的虚拟机 (VM) 设备。例如，您需要DataSync代理将文件从本地存储系统传输到AWS。

使用这些信息来了解在存储环境中部署代理需要什么。

### 找一个代理

您可以从[DataSync控制台](#)下载代理。

### 虚拟机管理程序要求

可以在以下虚拟机管理DataSync程序上运行代理：

- VMware ESXi Hypervisor ( 版本 6.5、6.7 或 7.0 ) : VMware [网站上有免费版本的 VMware](#)。您还需要使用 VMware vSphere 客户端以连接到主机。

#### Note

当 VMware 结束对 ESXi 管理程序版本的一般支持时，DataSync 也将结束对该版本的支持。有关 VMware 支持的虚拟机管理程序版本的信息，请参阅 [VMware 网站上的 VMware 生命周期政策](#)。

- 微软 Hyper-V Hypervisor ( 2012 R2、2016 或 2019 年版本 ) : Hyper-V 的免费独立版本可在[微软下载中心](#)获得。对于此设置，您需要 Microsoft Windows 客户端计算机上的 Microsoft Hyper-V Manager 才能连接到主机。

**Note**

DataSync代理虚拟机 (VM) 是第 1 代虚拟机。有关第 1 代和第 2 代虚拟机之间差异的更多信息，请参阅[我应该如何在 Hyper-V 中创建第 1 代还是第 2 代虚拟机？](#)

- 基于 Linux 内核的虚拟机 (KVM)：一种免费的开源虚拟化技术。KVM 包含在 Linux 版本 2.6.20 及更高版本中。DataSync适用于 Centos/RHEL 7 和 8、Ubuntu、Ubuntu 16.04 LTS 和 Ubuntu 16.04 LTS 发行版、Ubuntu 16.04 LTS 和 Ubuntu 其他现代 Linux 发行版可能有效，但不能保证功能或性能。如果您已经启动并运行 KVM 环境并且已经熟悉 KVM 的工作原理，我们建议您使用此选项。

**Note**

不支持在 Amazon EC2 上运行 KVM，也不能用于DataSync代理。要在 Amazon EC2 上运行该代理，请部署一个代理Amazon EC2、Amazon EC2、AMI mazon、有关在 Amazon EC2 上部署代理 AMI 的更多信息，请参阅[在 Amazon EC2 上部署您的代理](#)。

- Amazon EC2：DataSync提供包含虚拟机映像 (AMI M)，其中包含DataSync虚拟机映像。有关建议的实例类型，请参阅[Amazon EC2 实例要求](#)。

## 代理人对DataSync转账的要求

对于DataSync转账，您的代理必须满足以下资源要求。

### 主题

- [虚拟机要求](#)
- [Amazon EC2 实例要求](#)

## 虚拟机要求

部署不是 Amazon EC2 实例的DataSync代理时（例如，在本地存储环境中），代理虚拟机需要以下资源：

- 虚拟处理器：分配给 VM 的四个虚拟处理器。
- 磁盘空间：用于安装虚拟机映像和系统数据的 80 GB 磁盘空间。
- RAM：根据您的传输方案，您需要以下内存量：
  - 为虚拟机分配 32 GB 的 RAM，用于传输多达 2000 万个文件的任务。



- 为虚拟机分配 64 GB 的 RAM，用于传输超过 2000 万个文件的任务。

## Amazon EC2 实例要求

将DataSync代理部署为 Amazon EC2 实例时，实例大小必须至少为 2 倍大。

我们建议使用以下实例大小之一：

- m5.2xlarge：用于传输多达 2000 万个文件的任务。
- m5.4xlarge：用于传输超过 2000 万个文件的任务。

### Note

这些建议的一个例外情况是，如果您在AWS Snowcone设备上运行DataSync代理。使用默认实例 snc1.medium，它提供 2 个 CPU 内核和 4 GiB 内存。

## DataSyncDiscovery 的代理要求

无论是虚拟机还是 Amazon EC2 实例，您在DataSync Discovery 中使用的代理都必须有 80 GB 的磁盘空间和 16 GB 的内存。

## 代理管理要求

[激活](#)DataSync代理后，将为您AWS管理代理。有关更多信息，请参阅[管理您的AWS DataSync代理](#)：

## AWS DataSync网络要求

配置网络是设置过程中的重要步骤AWS DataSync。您的网络配置取决于多个因素，例如您是否需要有关存储的信息，还是准备好传输数据。它还取决于您计划使用哪种服务端点向其发送数据AWS。

## 自管理存储系统的网络要求

自管理存储可以指您管理的本地或云内存存储系统。使用此类系统进行传输需要使用代DataSync理。

**Note**

根据您的网络，您可能需要允许此处列出的端口以外的端口上的流量DataSync才能连接到您的自管理存储。

From	目的	协议	端口	它是如何使用的DataSync
DataSync 代理	NFS 文件服务器	TCP/UDP	2049	装载 NFS 文件服务器。  DataSync支持 NFS 版本 3.x、4.0 和 4.1。
DataSync 代理	SMB 文件服务器	TCP/UDP	139 或 445	装载 SMB 文件服务器。  DataSync支持 SMB 版本 1.0 及更高版本 SMB 版本。
DataSync 代理	对象存储	TCP	443 (HTTPS) 或 80 (HTTP)	访问您的对象存储。
DataSync 代理	Hadoop 集群	TCP	NameNode 端口 ( 默认 为 8020 )  在大多数 集群中， 您可以 在fs.defau t 或fs.def t.name 属 性下 的core- site .xml 文件	访问您的NameNodes Hadoop 集群中的。指定 创建 HDFS 位置时使用的 端口。

From	目的	协议	端口	它是如何使用的DataSync
			中找到此端口号 ( 取决于 Hadoop 发行版 ) 。	
DataSync 代理	Hadoop 集群	TCP	DataNode 端口 ( 默认值为 50010 )  在大多数集群中，可以在 <code>dfs.dataode.address</code> 属性下的 <code>hdfs-site.xml</code> 文件中找到此端口号。	访问您的 DataNodes Hadoop 集群中的。Data Sync 代理自动确定要使用的端口。
DataSync 代理	Hadoop 密钥管理服务器 ( KMS )	TCP	KMS 端口 ( 默认值为 9600 )	访问您的 Hadoop 集群的 KMS。
DataSync 代理	Kerberos 密钥分发中心 (KDC) 服务器	TCP	KDC 端口 ( 默认值为 88 )	使用 Kerberos 领域进行身份验证。此端口仅用于 HDFS。
DataSync 代理	存储系统的管理接口	TCP/UDP	取决于您的网络	连接到您的存储系统。DataSyncDiscovery 使用此连接来收集有关您的系统的信息。

## AWS存储服务的网络要求

传输期间DataSync连接到AWS存储服务所需的网络端口各不相同。

From	目的	协议	端口
DataSync 服务	Amazon EFS	TCP	2049
DataSync 服务	FSx for Windows File Server	请参阅 <a href="#">Windows FSx for Windows File Server</a> <a href="#">FSx for Windows File Server</a> F	
DataSync 服务	FSx for Lustre	参见 <a href="#">FsX for Lustre 的文件系统访问控制</a> 。	
DataSync 服务	FSx for OpenZFS	参见 <a href="#">OpenZFS 的 FSx 文件系统访问控制</a> 。	
DataSync 服务	Amazon FSx	TCP	111、635 和 2049 ( NFS ) 445 (SMB)
DataSync 服务	Amazon S3	TCP	443 (HTTPS)

## VPC 终端节点的网络要求

虚拟PVPC 终端节点可在您的代理之间提供专用连接AWS，且不跨越互联网或使用公有 IP 地址。这还有助于防止数据包进入或退出网络。有关更多信息，请参阅[将AWS DataSync代理与 VPC 终端节点一起使用](#)：

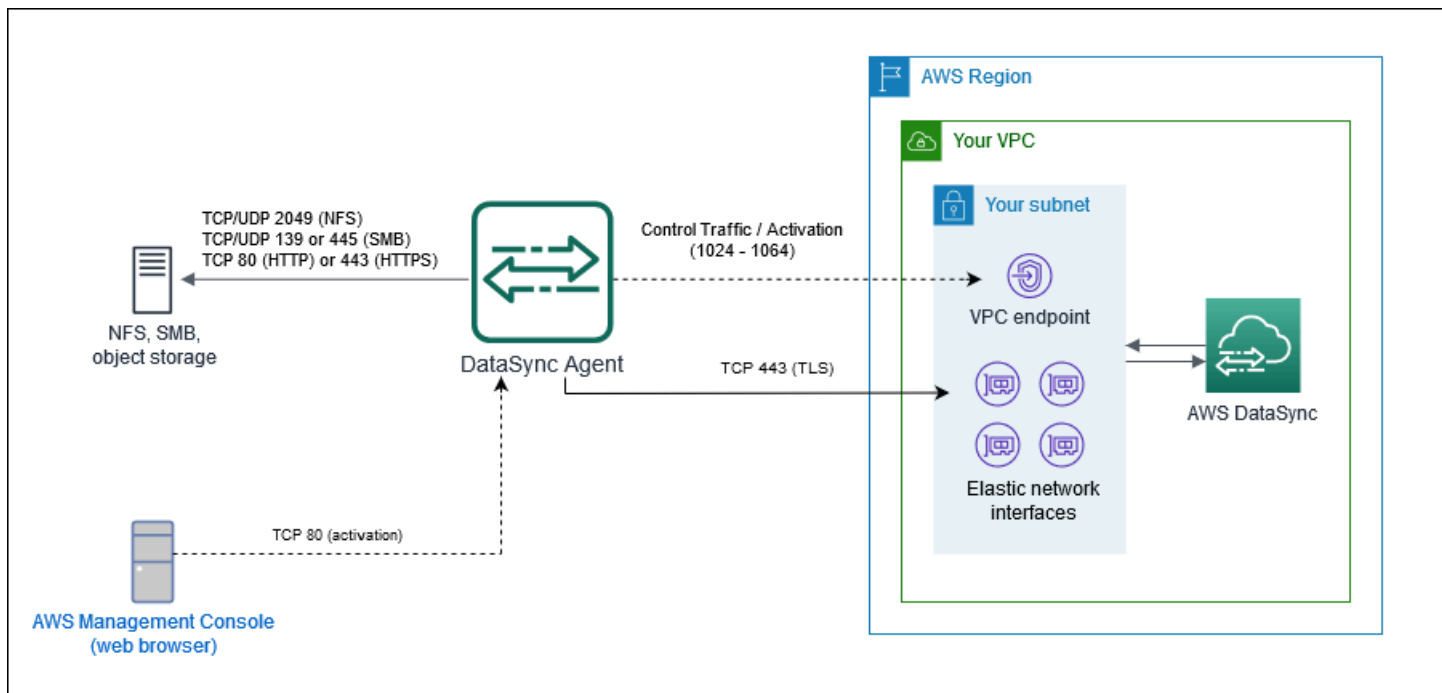
DataSync您的代理需要以下端口才能使用 VPC 终端节点。

From	目的	协议	端口	它是如何使用的
您的 Web 浏览器	你的DataSync经纪人	TCP	80 (HTTP)	由计算机用于获取代理激活密钥。成功激活后，DataSync 将关闭代理的端口 80。

From	目的	协议	端口	它是如何使用的
				<p>DataSync 代理不要求端口 80 可供公开访问。所需的端口 80 访问级别取决于网络配置。</p> <div style="border: 1px solid #00a0e3; border-radius: 10px; padding: 10px; margin-top: 10px;"> <p> <b>Note</b></p> <p>或者，您可以从代理的本地控制台获取激活密钥。此方法不需要在浏览器和代理之间建立连接。有关使用本地控制台获取激活密钥的更多信息，请参阅<a href="#">获取代理激活密钥</a>。</p> </div>
DataSync 代理	<p>您的 DataSync VPC 终端节点</p> <p>要找到正确的 IP 地址，请打开 <a href="#">Amazon VPC 控制台</a>，然后从左侧导航窗格中选择终端节点。选择 DataSync 终端节点，然后查看 Subnets (子网) 列表以查找与您为 VPC 终端节点设置选择的子网相对应的私有 IP 地址。</p> <p>有关更多信息，请参阅<a href="#">将您的 DataSync 代理配置为使用 VPC 终端节点中的步骤 5</a>。</p>	TCP	1024—1064	用于控制 DataSync 代理和 AWS 服务之间的流量。

From	目的	协议	端口	它是如何使用的
DataSync 代理	<p>你的任务的网络接口</p> <p>要查找相关的 IP 地址，请打开 Amazon EC2 控制台并从左侧导航窗格中选择网络接口。要查看任务的四个网络接口，请在搜索过滤器中输入您的任务 ID。</p> <p>有关更多信息，请参阅<a href="#">将您的DataSync代理配置为使用 VPC 终端节点</a>中的步骤 9。</p>	TCP	443 (HTTPS)	用于将数据从DataSync VM 传输到AWS 服务。
DataSync 代理	您的 DataSync VPC 终端节点	TCP	22 (支持渠道)	<p>允许AWS Support访问您的DataSync代理进行故障排除。</p> <p>您无需打开此端口即可正常运行。</p>

下图显示了使用 VPC 终端节点DataSync时所需的端口。



## 公共或 FIPS 端点的网络要求

使用公共或 FIPS 服务端点时，您的DataSync代理需要以下网络访问权限。如果您使用防火墙或路由器来过滤或限制网络流量，请将您的防火墙或路由器配置为允许这些端点。

From	目的	协议	端口	它是如何使用的	代理访问的终端节点
您的 Web 浏览器	DataSync 代理	TCP	80 (HTTP)	<p>允许您的计算机获取DataSync代理的激活密钥。成功激活后，DataSync将关闭代理的端口 80。</p> <p>代理不要求端口 80 可供公开访问。所需的端口 80 访问级别取决于网络配置。</p>	不适用

From	目的	协议	端口	它是如何使用的	代理访问的终端节点
				<p> <b>Note</b></p> <p>或者，您可以从代理的本地控制台获取激活密钥。此方法不需要在浏览器和代理之间建立连接。有关更多信息，请参阅<a href="#">获取代理激活密钥</a>：</p>	



From	目的	协议	端口	它是如何使用的	代理访问的终端节点
DataSync 代理	Amazon CloudFront	TCP	443 (HTTPS)	帮助在激活之前引导您的DataSync代理。	<p>AWS 区域:</p> <ul style="list-style-type: none"> <li>d3dvvaliwoko8h.cloudfront.net</li> </ul> <p>AWS GovCloud (US)地区 :</p> <ul style="list-style-type: none"> <li>s3.us-gov-west-1.amazonaws.com/fmrse-ndpoints-endpoints-bucket-go4p5gpna6sk</li> </ul>
DataSync 代理	AWS	TCP	443 (HTTPS)	激活您的DataSync代理并将其与您的代理关联AWS 账户。您可以在激活后封锁公共终端节点。	<p><i>activation-region</i> 是激活DataSync代理AWS 区域的地方。</p> <p>公共端点激活 :</p> <ul style="list-style-type: none"> <li>activation.datasync.c.<i>activation-region</i>.amazonaws.com</li> </ul> <p>FIPS 端点激活 :</p> <ul style="list-style-type: none"> <li>activation.datasync-fips.<i>activation-region</i>.amazonaws.com</li> </ul>

From	目的	协议	端口	它是如何使用的	代理访问的终端节点
DataSync代理	AWS	TCP	443 (HTTPS)	<p>允许DataSync代理和AWS服务端点之间的通信。</p> <p>有关信息，请参阅 <a href="#">为您的AWS DataSync代理选择服务终端节点</a>。</p>	<p><b>代理访问的终端节点</b></p> <p><i>activation-region</i> 这是您激活DataSync代理AWS区域的地方。根据您的用DataSync途，您可能不需要允许访问此处列出的每个端点。</p> <p>DataSyncAPI 终端节点：</p> <ul style="list-style-type: none"> <li>• <code>datasync. <i>activation-region</i>.amazonaws.com</code></li> </ul> <p>DataSync发现终端节点（仅适用于发现任务）：</p> <ul style="list-style-type: none"> <li>• <code>discovery-datasync. <i>activation-region</i>.amazonaws.com</code></li> </ul> <p>DataSync控制平面端点：</p> <ul style="list-style-type: none"> <li>• 公共终端节点：<code>cp.datasync. <i>activation-region</i>.amazonaws.com</code></li> <li>• FIPS 终端节点：<code>cp.datasync-fips. <i>activation-region</i>.amazonaws.com</code></li> </ul> <p>DataSync数据平面端点（仅用于传输任务）：</p> <ul style="list-style-type: none"> <li>• <code><i>your-task-id</i>.datasync-dp. <i>activation-region</i>.amazonaws.com</code></li> </ul>

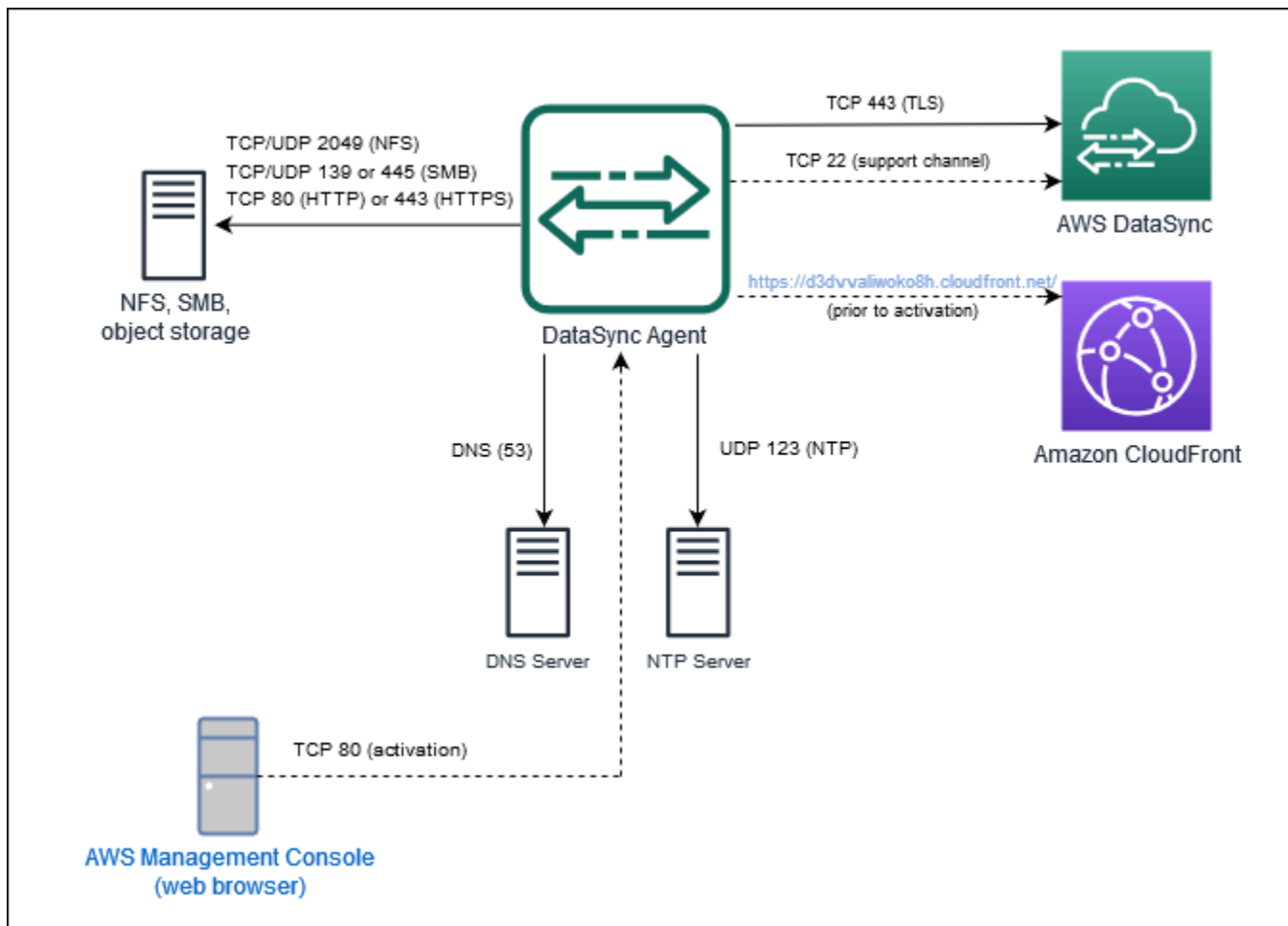
From	目的	协议	端口	它是如何使用的	代理访问的终端节点
DataSync 代理	AWS	TCP	80 (HTTP)	允许DataSync代理从中获取更新AWS。有关更多信息，请参阅 <a href="#">管理您的AWS DataSync代理</a> ：	<p><i>activation-region</i> 这是您激活DataSync代理AWS区域的地方。</p> <ul style="list-style-type: none"> <li>• <code>repo.default.amazonaws.com</code></li> <li>• <code>packages.us-west-1.amazonaws.com</code></li> <li>• <code>packages.sa-east-1.amazonaws.com</code></li> <li>• <code>repo.<i>activation-region</i>.amazonaws.com</code></li> <li>• <code>packages.<i>activation-region</i>.amazonaws.com</code></li> </ul>
DataSync 代理	AWS	TCP	443 (HTTPS)	允许DataSync代理从中获取更新AWS。有关更多信息，请参阅 <a href="#">管理您的AWS DataSync代理</a> ：	<p><i>activation-region</i> 这是您激活DataSync代理AWS区域的地方。</p> <ul style="list-style-type: none"> <li>• <code>amazonlinux.default.amazonaws.com</code></li> <li>• <code>cdn.amazonlinux.com</code></li> <li>• <code>amazonlinux-2-repos-<i>activation-region</i>.s3.dualstack.<i>activation-region</i>.amazonaws.com</code></li> <li>• <code>amazonlinux-2-repos-<i>activation-region</i>.s3.<i>activation-region</i>.amazonaws.com</code></li> <li>• <code>*.s3.<i>activation-region</i>.amazonaws.com</code></li> </ul>

From	目的	协议	端口	它是如何使用的	代理访问的终端节点
DataSync 代理	域名服务 (DNS) 服务器	TCP/ UDP	53 (DNS)	允许DataSync代理和DNS服务器之间的通信。	不适用
DataSync 代理	AWS	TCP	22 (支持渠道)	AWS Support 允许访问您的DataSync代理以帮助您解决问题。您无需打开此端口即可正常运行。	AWS Support频道 :  • 54.201.223.107
DataSync 代理	网络时间协议 (NTP) 服务器	UDP	123 (NTP)	允许本地系统将VM时间与主机时间同步。	NTP :  • 0.amazon.pool.ntp.org • 1.amazon.pool.ntp.org • 2.amazon.pool.ntp.org • 3.amazon.pool.ntp.org

 **Note**

要使用本地控制台将虚拟机代理的默认 NTP 配置更改为使用其他 NTP 服务器，请参阅[同步您的 VMware 代理上的时间](#)。

下图显示了使用公共或 FIPS 端点DataSync时所需的端口。



## 网络接口要求

对于您创建的每项任务，都会DataSync自动生成和管理用于数据传输流量的[网络接口](#)。创建了多少网络接口及其DataSync创建位置取决于有关您的任务的以下详细信息：

- 您的任务是否需要代DataSync理。
- 您的源位置和目标位置（您从中复制数据的位置和向其中复制数据）。
- 用于激活代理的端点的类型。

每个网络接口在您的子网中使用一个 IP 地址（网络接口越多，需要的 IP 地址就越多）。使用下表确保您的子网有足够的 IP 地址来完成任务。

### 主题

- [用于与代理进行传输的网络接口](#)
- [无需代理即可进行传输的网络接口](#)

- [查看您的网络接口](#)

## 用于与代理进行传输的网络接口

在AWS存储服务和非存储系统之间复制数据时，需要DataSync代理AWS。

位置	默认情况下创建的网络接口	使用公共或 FIPS 端点时在哪里创建网络接口	使用私有 (VPC) 终端节点时在哪里创建网络接口
Amazon S3	4	不适用 ( 不需要网络接口，因为直接DataSync与 S3 存储桶通信 )	您在激活DataSync代理时指定的子网。
Amazon EFS	4	您在创建 Amazon EFS 位置时指定的子网。	
Amazon FSx for Windows File Server	4	与文件系统的首选文件服务器相同的子网。	
Amazon FSx for Lustre	4	与文件系统相同的子网。	
Amazon FSx for OpenZFS	4	与文件系统相同的子网。	
Amazon FSx for NetApp ONTAP	4	与文件系统相同的子网。	

## 无需代理即可进行传输的网络接口

在两者之间复制数据时不需要DataSync代理AWS 服务。

### Note

网络接口总数取决于您的DataSync任务位置。例如，从 Amazon EFS 站点传输到 FsX for Lustre 需要四个网络接口。同时，从FSx for Windows File Server 传输到 Amazon S3 存储桶需要两个网络接口。

位置	默认情况下创建的网络接口	在哪里创建网络接口
Amazon S3	不适用 ( 不需要网络接口 , 因为直接DataSync与 S3 存储桶通信 )	
Amazon EFS	2	您在创建 Amazon EFS 位置时指定的子网。
FSx for Windows File Server	2	与文件系统的首选文件服务器相同的子网。
FSx for Lustre	2	与文件系统相同的子网。
FSx for OpenZFS	2	与文件系统相同的子网。
Amazon FSx	2	与文件系统相同的子网。

## 查看您的网络接口

要查看分配给您的DataSync任务的网络接口，请执行以下操作之一：

- 使用该[DescribeTask](#)操作。该操作返回SourceNetworkInterfaceArnsDestinationNetworkInterfaceArns，其响应如下所示：

```
arn:aws:ec2:your-region:your-account-id:network-interface/eni-f012345678abcdef0
```

在此示例中，网络接口 ID 为eni-f012345678abcdef0。

- 在 Amazon EC2 控制台中，搜索您的任务 ID ( 例如task-f012345678abcdef0 ) 以查找其网络接口。

## 使用所需的 IAM 权限AWS DataSync

AWS DataSync可以将您的数据移至 Amazon S3 存储桶、Amazon EFS 文件系统或[许多其他AWS 存储服务](#)。要将数据传输到您想要的地方，您需要向您的身份授予正确AWS Identity and Access Management (IAM) 权限。例如，您使用的 IAM 角色DataSync需要权限才能使用将数据传输到 S3 存储桶所需的 Amazon S3 操作。

您可以使用自己提供的 IAM 策略授予这些权限，AWS 也可以通过创建自己的策略来授予这些权限。

## AWS 托管策略

AWS 为常见 DataSync 用例提供以下托管策略：

- `AWSDataSyncReadOnlyAccess` – 提供对 DataSync 的只读访问权限。
- `AWSDataSyncFullAccess`— 提供对其依赖项的完全访问权限 DataSync 和最低限度的访问权限。

有关更多信息，请参阅 [AWS 适用于 AWS DataSync 的托管策略](#)：

## 客户管理型策略

您可以创建自定义 IAM policy 来使用 DataSync。有关更多信息，请参阅 [IAM 客户托管式策略 AWS DataSync](#)。



# AWS DataSync 入门

本节主要介绍如何AWS DataSync通过使用开始使用AWS Management Console。

在开始之前，我们建议您阅读[AWSDataSync工作原理](#)。

## 主题

- [设置为AWS DataSync](#)
- [登录到 AWS DataSync 控制台](#)
- [创建代AWS DataSync理](#)
- [通过以下方式发现您的存储空间 AWS DataSync 数据查找服务](#)
- [通过以下方式传输您的数据 AWS DataSync](#)
- [清除 AWS 资源](#)

## 设置为AWS DataSync

在你开始使用之前AWS DataSync，AWS 账户如果你没有，你需要注册一个。我们还建议您了解DataSync可以在哪里使用以及移动数据可能需要多少费用。

## 注册一个 AWS 账户

如果您还没有 AWS 账户，请完成以下步骤来创建一个。

### 注册 AWS 账户

1. 打开 <https://portal.aws.amazon.com/billing/signup>。
2. 按照屏幕上的说明进行操作。

在注册时，您将接到一通电话，要求您使用电话键盘输入一个验证码。

当您注册 AWS 账户时，系统将会创建一个 AWS 账户根用户。根用户有权访问该账户中的所有 AWS 服务和资源。作为安全最佳实践，请 [为管理用户分配管理访问权限](#)，并且只使用根用户执行 [需要根用户访问权限的任务](#)。

AWS注册过程完成后，会向您发送一封确认电子邮件。在任何时候，您都可以通过转至 <https://aws.amazon.com/> 并选择 My Account (我的账户) 来查看当前的账户活动并管理您的账户。

## 创建管理用户

注册 AWS 账户后，创建一个管理用户，以避免使用根用户执行日常任务。

保护您的 AWS 账户根用户

1. 选择 Root user (根用户) 并输入您的 AWS 账户电子邮件地址，以账户所有者身份登录 [AWS Management Console](#)。在下一页上，输入您的密码。

要获取使用根用户登录方面的帮助，请参阅《AWS 登录 用户指南》中的[以根用户身份登录](#)。

2. 对您的根用户启用多重身份验证 (MFA)。

有关说明，请参阅《IAM 用户指南》中的[为 AWS 账户根用户启用虚拟 MFA 设备 \(控制台\)](#)。

### 创建管理用户

- 对于您的日常管理任务，请在 AWS IAM Identity Center 中为管理用户授予管理访问权限。

有关说明，请参阅《AWS IAM Identity Center 用户指南》中的[入门](#)。

### 作为管理用户登录

- 要使用您的 IAM Identity Center 用户身份登录，请使用您在创建 IAM Identity Center 用户时发送到您的电子邮件地址的登录 URL。

要获取使用 IAM Identity Center 用户登录方面的帮助，请参阅《AWS 登录 用户指南》中的[登录 AWS 访问门户](#)。

## 我可以在哪里使用DataSync？

有关DataSync支持的AWS 区域终端节点列表，请参阅中的[AWS DataSync终端节点和配额AWS 一般参考](#)。

## 我如何如何如何如何如何如何如何如何DataSync如何如何

有多种方式DataSync。

- [DataSync控制台](#)，这是的一部分AWS Management Console。
- [DataSyncAPI](#) 或[AWS CLI](#)以编程方式配置和管理DataSync。

- [AWS CloudFormation](#) 或者 [Terraform](#) 来配置你的DataSync资源。
- AWS用于构建使用的应用程序的 [SDK](#)DataSync。

## DataSync费用是多少？

在[DataSync定价](#)页面上，使用您计划复制的数据量创建自定义估算值。

## 登录到 AWS DataSync 控制台

AWS DataSync立即通过控制台开始使用。

要AWS DataSync通过控制台开始使用

1. 通过 <https://console.aws.amazon.com/datasync/> 打开AWS DataSync主机。
2. 在右上角，选择要使用AWS 区域的地方。DataSync

我们建议选择传输过程中AWS存储资源AWS 区域使用的相同存储资源。

3. 在DataSync主页上，请选择以下选项之一：
  - 如果您DataSync想帮助您了解本地@@ 存储，请探索存储。
  - 传输数据，开始在存储服务之间、从AWS存储服务之间移动数据。

下一步：[创建代AWS DataSync理](#)

## 创建代AWS DataSync理

无论您是计划数据迁移还是准备移动数据，都可能需要代AWS DataSync理，原因如下：

- 了解您的存储 — DataSync Discovery 使用代理收集有关如何使用和配置本地存储系统的信息。
- 传输数据-DataSync 使用代理读取和写入位于本地或其他云中的存储系统。

### Tip

在同一AWS存储服务之间传输时，无需代理AWS 账户。如果这是你想要做的，请直接跳到[为其创建源位置 AWS DataSync](#)。

我们建议使用单独的代理进行DataSync发现和DataSync传输。对于DataSync传输，如果代理可以访问您的存储系统并且已在同一代理中激活，则可以重复使用AWS 区域。

## 主题

- [部署您的AWS DataSync代理](#)
- [为您的AWS DataSync代理选择服务终端节点](#)
- [激活您的AWS DataSync代理](#)

## 部署您的AWS DataSync代理

AWS DataSync为各种存储环境提供代理。你可以在 VMware ESXi、基于 Linux 内核的虚拟机 (KVM) 或微软 Hyper-V 虚拟机管理程序上部署代理。对于在中的虚拟私有云 (VPC) 中的存储AWS，您可以将代理作为 Amazon EC2 实例。

## 主题

- [在 VMware 上部署您的代理](#)
- [在 KVM 上部署您的代理](#)
- [在微软 Hyper-V 上部署您的代理](#)
- [在 Amazon EC2 上部署您的代理](#)
- [在上部署您的代理 AWS Snowcone](#)
- [在上部署您的代理 AWS Outposts](#)

## 在 VMware 上部署您的代理

您可以从DataSync控制台下载代理并将其部署到您的 VMware 环境中。

开始之前：确保您的存储环境可以支持代DataSync理。有关更多信息，请参阅[虚拟机要求](#)：

### 在 VMware 上部署代理

1. [通过 https://console.aws.amazon.com/datasync/](https://console.aws.amazon.com/datasync/) 打开AWS DataSync主机。
2. 在左侧导航窗格中，选择 Agents。
3. 对于 Hypervisor，选择 VMware ESXi，然后选择下载映像。

代理在包含 .ova 图像 .zip 文件的文件中下载。

4. 要最大限度地减少网络延迟，请将代理部署在尽可能靠近DataSync需要访问的存储系统附近（如果可能，使用相同的本地网络）。有关更多信息，请参阅[AWS DataSync网络要求](#)：

如果需要，请参阅您的虚拟机管理程序文档，了解如何在 VMware .ova 主机中部署文件。

5. 开启虚拟机管理程序，登录代理虚拟机，获取代理的 IP 地址。您需要此 IP 地址来激活代理。

代理 VM 的默认凭证是登录名 **admin** 和密码 **password**。如果需要，通过[虚拟机的本地控制台](#)更改密码。

## 在 KVM 上部署您的代理

您可以从DataSync控制台下载代理并将其部署到您的 KVM 环境中。

开始之前：确保您的存储环境可以支持代DataSync理。有关更多信息，请参阅[虚拟机要求](#)：

要在 KVM 上部署该代理

1. [通过 https://console.aws.amazon.com/datasync/](https://console.aws.amazon.com/datasync/) 打开AWS DataSync主机。
2. 在左侧导航窗格中，选择 Agents。
3. 对于 Hypervisor，选择基于内核的虚拟机 (KVM)，然后选择下载映像。

代理在包含 .qcow2 图像 .zip 文件的文件中下载。

4. 要最大限度地减少网络延迟，请将代理部署在尽可能靠近DataSync需要访问的存储系统附近（如果可能，使用相同的本地网络）。有关更多信息，请参阅[AWS DataSync网络要求](#)：
5. 运行以下命令以安装您的 .qcow2 镜像。

```
virt-install \  
  --name "datasync" \  
  --description "DataSync agent" \  
  --os-type=generic \  
  --ram=32768 \  
  --vcpus=4 \  
  --disk path=datasync-yyyyymmdd-x86_64.qcow2,bus=virtio,size=80 \  
  --network default,model=virtio \  
  --graphics none \  
  --import
```

有关如何管理此 VM 和 KVM 主机的信息，请参阅您的虚拟机管理程序文档。

6. 打开管理程序电源，登录虚拟机并获取代理的 IP 地址。您需要此 IP 地址来激活代理。

代理 VM 的默认凭证是登录名 **admin** 和密码 **password**。如果需要，通过[虚拟机的本地控制台](#)更改密码。

## 在微软 Hyper-V 上部署您的代理

你可以从DataSync控制台下载代理并将其部署到你的 Microsoft Hyper-V 环境中。

开始之前：确保您的存储环境可以支持代DataSync理。有关更多信息，请参阅[虚拟机要求](#)：

### 在 Hyper-V 上部署代理

1. 通过 <https://console.aws.amazon.com/datasync/> 打开AWS DataSync主机。
2. 在左侧导航窗格中，选择 Agents。
3. 对于 Hypervisor，选择微软 Hyper-V，然后选择下载映像。

代理在包含 .vhdx 图像 .zip 文件的文件中下载。

4. 要最大限度地减少网络延迟，请将代理部署在尽可能靠近DataSync需要访问的存储系统附近（如果可能，使用相同的本地网络）。有关更多信息，请参阅[AWS DataSync网络要求](#)：

如果需要，请参阅您的虚拟机管理程序文档，了解如何在 Hyper-V .vhdx 主机中部署文件。

#### Warning

如果您在使用 Broadcom 网络适配器的 Hyper-V 主机上启用虚拟机队列 (VMQ)，您可能会注意到网络性能不佳。有关解决方法的信息，请参阅 [Microsoft 文档](#)。

5. 打开管理程序电源，登录虚拟机并获取代理的 IP 地址。您需要此 IP 地址来激活代理。

代理 VM 的默认凭证是登录名 **admin** 和密码 **password**。如果需要，通过[虚拟机的本地控制台](#)更改密码。

## 在 Amazon EC2 上部署您的代理

在以下之间复制数据时，您可以将DataSync代理部署为 Amazon EC2 实例：

- 自我管理的云端存储系统和AWS存储服务。

有关这些用例的更多信息，包括高级架构图，请参阅[将您的AWS DataSync代理部署在 AWS 区域](#)。

- [Amazon S3 开](#)启AWS Outposts和AWS存储服务。

#### Warning

我们不建议使用 Amazon EC2 代理访问您的本地存储，因为网络延迟会增加。取而代之的是，将代理作为 VMware、KVM 或 Hyper-V 虚拟机部署在数据中心中，尽可能靠近本地存储。

为你选择代理 AMI AWS 区域

- 使用以下 CLI 命令获取指定的最新 DataSync Amazon 系统映像 (AMI) ID AWS 区域。

```
aws ssm get-parameter --name /aws/service/datasync/ami --region region
```

Example 命令和输出示例

```
aws ssm get-parameter --name /aws/service/datasync/ami --region us-east-1

{
  "Parameter": {
    "Name": "/aws/service/datasync/ami",
    "Type": "String",
    "Value": "ami-id",
    "Version": 6,
    "LastModifiedDate": 1569946277.996,
    "ARN": "arn:aws:ssm:us-east-1::parameter/aws/service/datasync/ami"
  }
}
```

要将您的DataSync代理部署为 Amazon EC2 实例

#### Important

为避免收费，请以不要求可用区之间网络流量的方式部署代理。例如，在您的自我管理文件系统所在的可用区中部署您的代理。

要详细了解所有AWS 区域人的数据传输价格，请参阅 [Amazon EC2 按需定价](#)。

1. 从源文件系统所在AWS账户的位置，使用 Amazon EC2 启动向导中的 AMI 启动代理。使用以下 URL 启动 AMI。

```
https://console.aws.amazon.com/ec2/v2/home?region=source-file-system-region#LaunchInstanceWizard:ami=ami-id
```

在 URL 中，将 *source-file-system-region* 和 *ami-id* 替换为您自己的来源 AWS 区域和 AMI ID。选择一个实例类型页面将显示在 Amazon EC2 控制台上。

2. 为您的使用案例选择一种建议的实例类型，然后选择 Next: Configure Instance Details (下一步: 配置实例详细信息)。有关建议的实例类型，请参阅 [Amazon EC2 实例要求](#)。
3. 在 Configure Instance Details (配置实例详细信息) 页面中，执行以下操作：
  - a. 对于 Networks，请选择您的源 Amazon EFS 或 NFS 文件系统所在的虚拟私有云 (VPC)。
  - b. 对于 Auto-assign Public IP (自动分配公有 IP)，选择一个值。要使您的实例可从公共 Internet 中进行访问，请将自动分配公有 IP 设置为启用。否则，请将自动分配公有 IP 设置为禁用。如果未分配公有 IP 地址，请使用其私有 IP 地址在您的 VPC 中激活代理。

当您从云端文件系统传输文件时，为了提高性能，我们建议您选择 NFS 服务器所在的置放群组值。

4. 选择下一步：添加存储。代理不需要额外的存储空间，因此，您可以跳过此步骤，然后选择 Next: Add tags (下一步：添加标签)。
5. (可选) 在添加标签页面上，您可以向您的 Amazon EC2 实例添加标签。当您在页面上完成后，选择 Next: Configure Security Group (下一步：配置安全组)。
6. 在 Configure Security Group (配置安全组) 页面上，执行以下操作：
  - a. 确保所选安全组允许从您计划用于激活代理的 Web 浏览器到 HTTP 端口 80 的入站访问。
  - b. 确保源文件系统的安全组允许来自该代理。此外，请确保代理允许流向源文件系统的出站流量。如果您使用 VPC 终端节点部署代理，则需要允许其他端口。有关更多信息，请参阅 [DataSync 代理如何使用 VPC 终端节点](#)：

有关 DataSync 的一组完整的网络要求，请参阅 [AWS DataSync 网络要求](#)。

7. 选择审核和启动以检查您的配置，然后选择启动以启动您的实例。请记住使用您有权访问的密钥对。将显示一个确认页面，指出您的实例正在启动。
8. 选择 View Instances 以关闭确认页面并返回 Amazon EC2 实例屏幕。当您启动实例时，其初始状态为 pending (待处理)。实例启动后，其状态将更改为 running (正在运行)。此时，已为其分配了公共域名系统 (DNS) 名称和 IP 地址，您可以在“描述”选项卡中找到。



9. 如果您将自动分配公有 IP 设置为启用，则选择您的实例，然后在描述选项卡中记下公有 IP 地址。您之后将使用该 IP 地址连接到您的同步代理。

如果您将自动分配公有 IP 设置为禁用，则启动或使用 VPC 中的现有实例以激活代理。在这种情况下，您可以使用同步代理的私有 IP 地址从 VPC 中的该实例激活代理。

## 在上部署您的代理 AWS Snowcone

DataSync 代理 AMI 预安装在 Snowcone 设备上。使用以下工具之一启动代理：

- [AWS OpsHub](#)
- [Snowball 边缘客户端](#)

## 在上部署您的代理 AWS Outposts

你可以在你的前哨基地启动 A DataSync mazon EC2 实例。要了解有关在上启动 AMI 的更多信息 AWS Outposts，请参阅 AWS Outposts 用户指南中的在 [Outpost 上启动实例](#)。

使用 DataSync 在 Outposts 上访问 Amazon S3 时，您必须在允许访问您的 Amazon S3 接入点的 VPC 中启动代理，并在前哨的父区域中激活代理。代理还必须能够路由到存储桶的 Amazon S3 on Outposts 终端节点。要了解有关在 Outposts 终端节点上使用 Amazon S3 的更多信息，请参阅 Amazon S3 用户指南中的 [使用 Amazon S3 在 Outposts 站上使用 Amazon](#)

## 为您的 AWS DataSync 代理选择服务终端节点

您的 AWS DataSync 代理使用服务端点与之通信 AWS。代理可以连接到以下类型的端点：

- 虚拟私有云 (VPC) 终端节点 — 数据通过您的 VPC 而不是通过公共互联网发送，从而提高了传输数据的安全性。
- 公共终端节点 — 数据通过公共互联网发送。
- 联邦信息处理标准 (FIPS) 端点-使用符合 FIPS 的流程通过公共互联网发送数据。

当选择服务终端节点时，请记住以下事项：

- 代理只能使用一种类型的端点。如果您需要使用不同的端点类型传输数据，请为每种类型创建代理。
- 对于 DataSync Discovery，目前您只能使用公共终端节点。

有关更多信息，请参阅《AWS 一般参考》中的 [AWS 服务端点](#)。

## 主题

- [使用 VPC 终端节点](#)
- [使用公共终端节点](#)
- [使用 FIPS 端点](#)

## 使用 VPC 终端节点

您的DataSync代理可以使用提供的 VPC 终端节点与AWS进行通信AWS PrivateLink。此方法在您的存储系统、VPC 和之间提供私有连接AWS 服务。

有关更多信息，请参阅[将AWS DataSync代理与 VPC 终端节点一起使用](#)：

使用DataSync控制台指定 VPC 终端节点

1. [创建 VPC 终端节点](#)并记下终端节点 ID。

您还可以在当前终端节点中使用现有 VPC 终端节点AWS 区域。

2. [通过 https://console.aws.amazon.com/datasync/](https://console.aws.amazon.com/datasync/) 打开AWS DataSync主机。
3. 转到“代理”页面，然后选择“创建代理”。
4. 对于虚拟机管理程序，请选择 Amazon EC2。
5. 在服务终端节点部分中，选择使用的 VPC 终端节点AWS PrivateLink。

这是该代理有权访问的 VPC 终端节点。

6. 对于 VPC 终端节点，请选择您希望您的代理连接到的 VPC 终端节点。

记下您创建 VPC 终端节点时的终端节点 ID。

### Important

您必须选择包含DataSync服务名称的 VPC 终端节点（例如，`com.amazonaws.us-east-2.datasync`）。

7. 对于子网，选择要在其中运行DataSync任务的子网。

这是为您的传输DataSync创建和管理[网络接口](#)的子网。

8. 对于安全组，为您的DataSync任务选择一个安全组。

这是保护您的传输网络接口的安全组。

有关在 VPC DataSync 中使用的更多信息，请参阅[将AWS DataSync代理与 VPC 终端节点一起使用](#)。

下一步：[the section called “激活您的代理”](#)

## 使用公共终端节点

如果您使用公共终端节点，则您的DataSync代理和之间的所有通信都通过公共互联网AWS进行。

使用DataSync控制台指定公有终端节点

1. [通过 https://console.aws.amazon.com/datasync/](https://console.aws.amazon.com/datasync/) 打开AWS DataSync主机。
2. 转到“代理”页面，然后选择“创建代理”。
3. 在服务端点部分中，在**AWS #####**中选择公共服务端点。有关支持的AWS区域列表，请参阅[AWS DataSync](#)中的AWS 一般参考。

下一步：[the section called “激活您的代理”](#)

## 使用 FIPS 端点

查看[使用的 FIPS 端点列表。DataSync](#)

使用控制台指定 FIPS 终端节点 DataSync

1. [通过 https://console.aws.amazon.com/datasync/](https://console.aws.amazon.com/datasync/) 打开AWS DataSync主机。
2. 对于 Hypervisor，选择您部署的代理类型。
3. 在 Service endpoint (服务终端节点) 部分中，选择所需的 FIPS 终端节点。

下一步：[the section called “激活您的代理”](#)

## 激活您的AWS DataSync代理

部署AWS DataSync代理并指定其服务终端节点后，您可以激活代理以将其与您的AWS 账户。

### Note

一次只能激活一个AWS 账户AWS 区域代理。

## 使用DataSync控制台激活您的代理

1. 在同一创建代理页面上，转到激活密钥部分。
2. 请选择以下选项之一来激活您的代理：
  - 自动从代理获取激活密钥-此选项要求您的浏览器使用端口 80 访问代理。激活后，代理会关闭端口。
  - 对于代理地址，输入代理的 IP 地址或域名，然后选择获取密钥。

您的浏览器连接到此 IP 地址，并从代理中获取唯一的激活密钥。如果激活失败，[请检查您的网络配置](#)。

- 手动输入代理的激活密钥-如果您不想在浏览器和代理之间建立连接，请使用此选项。
  - 从[代理的本地控制台](#)获取密钥。
  - 回到DataSync控制台，在激活密钥字段中输入密钥。

### Note

如果未使用，代理激活密钥将在 30 分钟后过期。

3. ( 可选 ) 对于 Agent name (代理名称)，输入代理的名称。
4. ( 可选 ) 对于“标签”，在“密钥”和“值”字段中输入值以标记您的代理。

标签可帮助您管理、筛选和搜索AWS资源。

5. 选择 Create agent (创建代理)。
6. 在代理页面上，验证您的服务终端节点是否正确。

### Note

此时，您可能会注意到您的代理处于离线状态。激活代理后，不久之后就会脱机

您已创建完代理，该代理[可以AWS完全为您管理](#)。

下一步：[为其创建源位置 AWS DataSync](#)

# 通过以下方式发现您的存储空间 AWS DataSync 数据查找服务

要了解本地存储系统的使用和配置方式，您可以快速创建和运行发现作业。

## 主题

- [将您的本地存储系统添加到 AWS DataSync 数据查找服务](#)
- [开始你的AWS DataSync探索工作](#)

## 将您的本地存储系统添加到 AWS DataSync 数据查找服务

在控制台中，配置AWS DataSync 数据查找服务为与您的本地存储系统配合使用。

使用DataSync控制台添加本地存储系统

1. [通过 https://console.aws.amazon.com/datasync/](https://console.aws.amazon.com/datasync/) 打开AWS DataSync主机。
2. 在左侧导航窗格中，选择 Discover ger，然后选择添加存储系统。
3. 对于 St orages 类型，请选择要添加的存储系统的类型。

### Note

DataSyncDiscovery 目前支持运行 ONTAP 9.7 或更高版本的NetApp结构连接存储 (FAS) 和全闪存 FAS (AFF) 系统。

4. 在存储名称中，为您的存储系统输入一个熟悉的名称。
5. 在管理接口中，输入存储系统管理接口的域名或 IP 地址。
6. 对于服务器端口，输入访问存储系统管理接口所需的网络端口。
7. 在“凭据”中，输入访问存储系统管理界面所需的用户名和密码。

这些凭证应提供对管理接口的读取权限。有关更多信息，请参阅[访问您的本地存储系统](#)：

8. 对于代理，选择您刚刚创建的DataSync代理。

代理连接到您的存储系统的管理接口。

9. 选择添加存储系统。

下一步：[开始你的AWS DataSync探索工作](#)

## 开始你的AWS DataSync探索工作

添加本地存储系统后，即可运行DataSync发现任务来收集有关该系统的信息。

使用DataSync控制台启动发现作业

1. 在控制台的左侧导航窗格中，选择 Discoverger。
2. 选择要在其上运行发现作业的存储系统。
3. 选择操作，然后选择开始。
4. 在“持续时间”中，选择您希望发现作业运行多长时间。

### Tip

为了获得更准确的建议，我们建议持续时间至少为 14 天。较长的持续时间允许有时间收集足够数量的数据点，并真实地呈现存储性能和利用率。

5. 选择“开始发现作业”。

当发现任务收集数据时，您开始在控制台中看到有关存储系统资源的信息。

6. 发现任务完成后，请执行以下操作以获取数据的AWS存储建议：
  - a. 选择您想要推荐的存储资源（例如，卷）。
  - b. 如果存储资源处于“准备生成建议”状态，请选择存储资源名称。
  - c. 在存储资源页面上，转到“推荐”选项卡，然后选择“获取建议”。

推荐一旦可用，就会显示在同一个选项卡上。

有关更多信息，请参阅 [查看收集的存储资源信息 AWS DataSync 数据查找服务](#) 和 [从中获得推荐 AWS DataSync 数据查找服务](#)。

## 通过以下方式传输您的数据 AWS DataSync

要快速开始AWS DataSync传输，您可以使用默认设置创建任务。

主题

- [为其创建源位置 AWS DataSync](#)
- [为其创建目的地位置 AWS DataSync](#)

- [创建并开始AWS DataSync任务](#)

## 为其创建源位置 AWS DataSync

源位置定义了AWS DataSync要从中传输数据的存储系统或服务。

以下说明介绍如何为网络文件系统 (NFS) 共享创建源位置。

如果要创建其他类型的源位置，请参阅以下主题：

- [配置来自 SMB 文件服务器的AWS DataSync传输](#)
- [配置来自 HDFS 的AWS DataSync传输](#)
- [配置来自对象存储系统的AWS DataSync传输](#)
- [使用 Amazon S3 配置AWS DataSync传输](#)
- [使用亚马逊 EFS 配置AWS DataSync传输](#)
- [使用Amazon FSx for Windows File Server AWS DataSync](#)
- [使用 Amazon FSx for Lustre 配置AWS DataSync传](#)
- [使用 Amazon FSx for OpenZFS 配置AWS DataSync传输](#)
- [使用适用NetApp于 ON AWS DataSync TAP 的亚马逊 FSx 配置传输](#)

### 使用控制台创建源 NFS 位置

1. 通过 <https://console.aws.amazon.com/datasync/> 打开AWS DataSync主机。
2. 在左侧导航窗格中，展开“数据传输”，然后选择“任务”和“创建任务”。
3. 在“配置源位置”页面上，选择“创建新位置”。

或者，如果您已经在自己的位置创建了地点，请选择“选择现有地点”AWS 区域。

4. 对于 L o c a t e s ( 网络文件系统 ) 。
5. 对于代理，选择将读取您的 NFS 服务器的代理。
6. 对于 NFS server (NFS 服务器)，输入 NFS 服务器的 IP 地址或域名。

在本地部署的代理使用它来挂载 NFS 服务器。

7. 对于 M o u n t 路径，请输入 NFS 服务器导出的路径，或者是网络中的其他 NFS 客户端挂载。

DataSync使用此路径从 NFS 服务器读取数据。

8. 选择“下一步”创建您的目的地位置。

## 为其创建目的地位置 AWS DataSync

目标位置定义了要将数据传输AWS DataSync到的存储系统或服务。

以下说明介绍如何为 Amazon S3 存储桶创建目标位置。

### Note

如果您要转移到AWS Outposts资源上的 S3 存储桶，请参阅[使用 Amazon S3 配置AWS DataSync传输](#)。

如果要创建其他类型的目标位置，请参阅以下主题：

- [配置来自 NFS 文件服务器的AWS DataSync传输](#)
- [配置来自 SMB 文件服务器的AWS DataSync传输](#)
- [配置来自 HDFS 的AWS DataSync传输](#)
- [配置来自对象存储系统的AWS DataSync传输](#)
- [使用亚马逊 EFS 配置AWS DataSync传输](#)
- [使用Amazon FSx for Windows File Server AWS DataSync](#)
- [使用 Amazon FSx for Lustre 配置AWS DataSync传](#)
- [使用 Amazon FSx for OpenZFS 配置AWS DataSync传输](#)
- [使用适用NetApp于 ON AWS DataSync TAP 的亚马逊 FSx 配置传输](#)

要通过控制台创建 Amazon S3 目标位置

1. 在配置目标位置页面上，选择创建新位置。

或者，如果您已经在自己的位置创建了地点，请选择“选择现有地点”AWS 区域。

2. 对于位置类型，选择 Amazon S3。

3. 对于 S3 存储桶，请选择要用作目标位置的存储桶。

如果您的 S3 存储桶位于AWS Outposts资源上，则必须指定 Amazon S3 接入点。有关更多信息，请参阅[Amazon S3 用户指南中的使用 Amazon S3 访问点管理数据访问](#)。



4. 对于 S3 存储类别，请选择您希望对象使用的存储类别。

有关更多信息，请查看 [S3 存储类注意事项](#) 和 [请求成本](#)。

5. 对于 Folder，在 S3 存储桶中输入一个用于 DataSync 读取或写入的前缀（取决于该存储桶是源位置还是目标位置）。

#### Note

前缀不能以斜杠（例如 /photos）开头，也不能包含连续的斜杠，例如 photos//2006/January。

6. 对于 IAM 角色，选择自动生成。

DataSync 自动创建具有访问 S3 存储桶所需权限的 AWS Identity and Access Management (IAM) 角色。

7. 选择创建地点。

## 创建并开始 AWS DataSync 任务

创建源位置和目标位置后，您可以完成 AWS DataSync 任务设置并开始移动数据。

#### Important

如果您计划向 Amazon S3 地点传输数据或从 Amazon S3 站点传输数据，请在开始之前查看 [DataSync 如何影响您的 S3 请求费用](#) 和 [DataSync 定价页面](#)。

## 创建您的任务

1. 在“查看”页面上，查看并根据需要更改您的任务设置。
2. 选择创建任务。

## 开始您的任务

1. 当任务的状态显示为“可用”时，选择“开始”。
2. 选择“以默认值开始”。

## 清除 AWS 资源

如果您曾AWS DataSync用于测试或不需要创建的AWS资源，请将其删除，这样您就不会为不打算使用的资源付费。

1. 删除不需要的DataSync任务。有关说明，请参阅 [删除你的AWS DataSync任务](#)。
2. 删除不需要DataSync的地点。有关说明，请参阅 [移除传AWS DataSync送地点](#)。
3. 删除不需要的DataSync代理。有关说明，请参阅 [删除代AWS DataSync理](#)。

# 与AWS DataSync代理合作

代理是一种虚拟机 (VM) 设备，AWS DataSync用于存储发现和某些数据传输（尤其是本地存储传输）。

## 主题

- [创建代AWS DataSync理](#)
- [将AWS DataSync代理与 VPC 终端节点一起使用](#)
- [将您的AWS DataSync代理部署在 AWS 区域](#)
- [使用多个AWS DataSync代理进行转账](#)
- [为多个 NIC 配置AWS DataSync代理](#)
- [管理您的AWS DataSync代理](#)
- [编辑您的AWS DataSync代理的属性](#)
- [AWS DataSync代理状态](#)
- [使用AWS DataSync代理的本地控制台](#)
- [删除代AWS DataSync理](#)

## 创建代AWS DataSync理

AWS DataSync为不同的存储环境提供多种类型的代理。例如，您可以使用 VMware ESXi 代理来处理本地文件系统。

创建一个代理，涉及以下步骤：

1. [配置您的网络](#)，以便您的代理可以与存储系统通信，以及AWS。
2. 在尽可能靠近存储系统的地方@@ [部署代理](#)。
3. [选择您的代理用来与之通信的服务端点](#)AWS。
4. [激活您的代理](#)。

## 将AWS DataSync代理与 VPC 终端节点一起使用

使用虚拟公共 (VPC) 终端节点，您无需通过公共 Internet 将您的数据用于。AWS DataSync可AWS通过基于 Amazon VPC 服务的 VPC，将数据传输您的。

## DataSync代理如何使用 VPC 终端节点

VPC 终端节点由提供AWS PrivateLink。这些类型的终端节点使您可以私下AWS 服务连接到支持的 VPC。当您把 VPC 终端节点与一起使用时DataSync，DataSync代理和之间的所有通信都将AWS保留在您的 VPC 中。

如果您要从本地存储系统进行传输，则必须将您的 VPC 扩展到存储所在的本地网络。您可以使用AWS Direct Connect或 Virtual Private Nprivate Nets (VPC) 执行此操作AWS Site-to-Site VPN。这包括从您的本地网络设置路由表以访问 VPC 终端节点。有关更多信息，请参阅AWS PrivateLink指南中的[网关终端节点路由](#)。

部署并激活代理后，您可以为转移创建任务。执行此操作时，DataSync会[为数据流量创建网络接口](#)。这些接口是私有 IP 地址，只能从您的 VPC 内部访问。

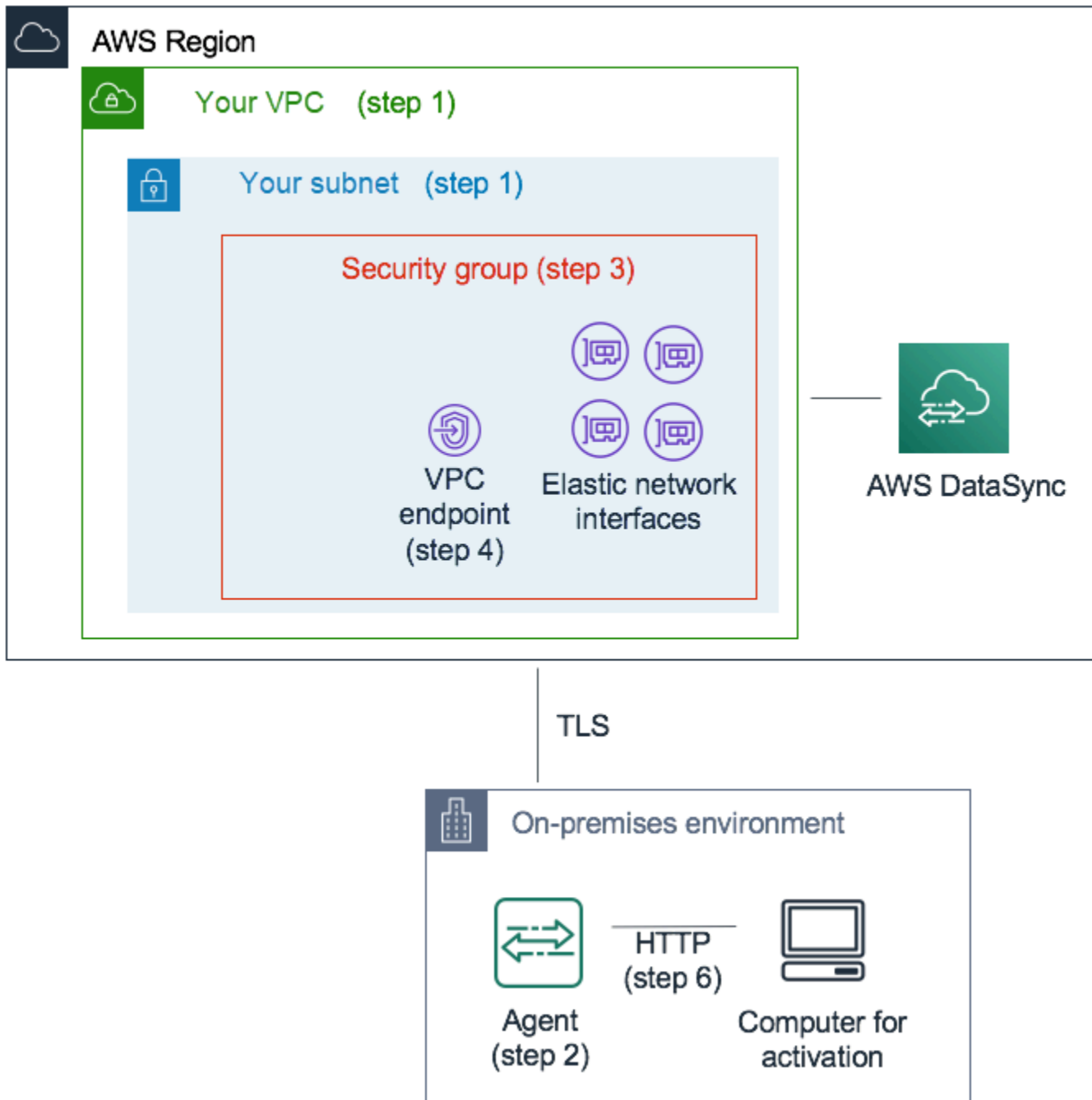
### DataSyncVPC 的局限性

- 与您一起使用的 VPCDataSync 必须具有默认租期。不支持具有专用租约的 VPC。有关更多信息，请参阅[使用上的](#)。
- DataSync不支持[共享 VPC](#)。

### 将您的DataSync代理配置为使用 VPC 终端节点

在以下步骤中，学习如何将DataSync代理配置为使用 VPC 终端节点。

下图说明了设置过程。



## 配置DataSync代理以使用 VPC 终端节点与AWS之通信

1. 选择要在其中设置 DataSync 私有 IP 地址的 VPC 和子网。

VPC 应通过使用路由规则AWS Direct Connect或 VPN 扩展到您的本地环境（您的自管理对象存储所在的位置）。

2. 在存储DataSync区附近部署代理。

代理必须能够使用 NFS、SMB 或 Amazon S3 API 访问您的源存储位置。您可以从DataSync控制台下载DataSync代理的 .ova 文件。该代理不需要公有 IP 地址。有关下载和部署 .ova 映像的更多信息，请参阅[使用创建AWS DataSync代理AWS CLI](#)。

#### Note

只能将代理用于一种类型的终端节点，即私有、公共或联邦信息处理标准 (FIPS)。如果您已将代理配置为通过公共 Internet 传输数据，请部署新代理以将数据传输到私有 DataSync 终端节点。有关详细说明，请参阅[部署您的AWS DataSync代理](#)。

3. 在您在步骤 1 中选择的 VPC 中，创建安全组以确保访问 DataSync 使用的私有 IP 地址。

这些地址包括一个用于控制流量的 VPC 终端节点和四个用于数据传输流量的[网络接口](#)。可以使用此安全组管理对这些私有 IP 地址的访问，并确保代理可以路由到这些私有 IP 地址。

代理必须能够与这些 IP 地址建立连接。在连接到端点的安全组中，配置入站规则，以允许代理的私有 IP 地址连接到这些端点。

4. 为 DataSync 服务创建 VPC 终端节点。

要执行此操作，请打开 Amazon VPC 控制台 <https://console.aws.amazon.com/vpc/>，然后从左侧的导航窗格中选择终端节点。选择 Create endpoint (创建端点)。

对于 Service category (服务类别)，选择 AWS 服务。在服务名称DataSync中，选择您的AWS区域 (例如，com.amazonaws.us-east-1.datasync)。然后，选择您在步骤 1 和步骤 3 中选择的 VPC 和安全组。确保清除 Enable Private DNS Name (启用私有 DNS 名称) 复选框。

#### Important

如果您已在 Amazon EC2 实例上部署了DataSync代理，请选择代理所在的可用区，以避免为可用区域之间的网络流量收费。


要了解有关所有数据传输价格的更多信息AWS 区域，请参阅 [Amazon EC2 按需定价](#)。

有关创建 VPC 终端节点的更多详细信息，请参阅 Amazon VPC 用户指南中的[创建接口终端节点](#)。

5. 当您的新 VPC 终端节点可用时，请确保存储环境的网络配置允许激活代理。

激活是一次性操作，可将代理与您的安全关联起来AWS 账户。要激活代理，请使用可通过端口 80 访问代理的计算机。激活后，您可以撤消此访问权限。代理必须能够访问您在步骤 4 中创建的 VPC 终端节点的私有 IP 地址。

要找到此 IP 地址，请打开 Amazon VPC 控制台 <https://console.aws.amazon.com/vpc/>，然后从左侧的导航窗格中选择终端节点。选择 DataSync 终端节点，然后在 Subnets (子网) 列表中查找所选子网的私有 IP 地址。这是 VPC 终端节点的 IP 地址。

 Note

确保使用端口 443、1024—1064 和端口 22 允许从代理到 VPC 终端节点的出站流量。端口 22 是可选的，用于信AWS Support道。

6. 激活代理。如果您的计算机可以使用端口 80 路由到代理并且可以访问控制台，请打开DataSync控制台，在左侧导航窗格中选择代理，然后选择创建代理。在服务终端节点部分中，使用选择 VPC 终端节点AWS PrivateLink。

从步骤 4 中选择 VPC 终端节点，从步骤 1 中选择子网，并从步骤 3 中选择安全组。输入代理的 IP 地址。

如果您无法使用同一台计算机访问代理和DataSync控制台，请在可以访问代理端口 80 的计算机上使用命令行激活代理。有关更多信息，请参阅[使用创建AWS DataSync代理AWS CLI](#)：

7. 选择获取密钥，（可选）输入代理名称和标签，然后选择创建代理。

您的新代理出现在DataSync控制台的“代理”选项卡上。绿色 VPC 终端节点状态表示使用此代理执行的所有任务均使用私有终端节点，无需通过公共互联网。

8. 通过配置传输的来源和目标位置来创建任务。

有关更多信息，请参阅[我可以哪里传输我的数据AWS DataSync？](#)：

为了使用私有 IP 地址简化传输，您的任务在您选择的 VPC 和子网中创建了四个网络接口。

9. 确保您的代理可以访问任务创建的四个网络接口和相关 IP 地址。

要找到这些 IP 地址，请打开 Amazon EC2 控制台 <https://console.aws.amazon.com/ec2/>，然后在控制面板上选择网络接口。在搜索过滤器中输入任务 ID 以查看任务的四个网络接口。这些是您的 VPC 终端节点使用的网络接口。请务必使用端口 443 以允许从代理到这些接口的出站流量。

现在可以开始您的任务。对于使用此代理的其他每个任务，重复步骤 9 以允许任务的通过端口 443 的流量。

## 将您的AWS DataSync代理部署在 AWS 区域

如果您在中部署AWS DataSync代理，以下指南可以帮助解决常见情况AWS 区域。如果没有上述必备项，请参阅[在 Amazon EC2 上部署您的代理](#)。

### 将数据从云文件系统传输到另一个云文件系统或 Amazon S3

要在云文件系统之间AWS 账户或从云文件系统传输数据，DataSync代理必须位于源文件系统AWS 账户所在的相同AWS 区域位置。这种类型的传输包括以下内容：

- 在 Amazon EFS 或 FSx for Windows 文件服务器文件系统之间传输到其他AWSAWS 账户存储系统。
- 从自我管理的文件系统传输到AWS存储服务。

#### Important

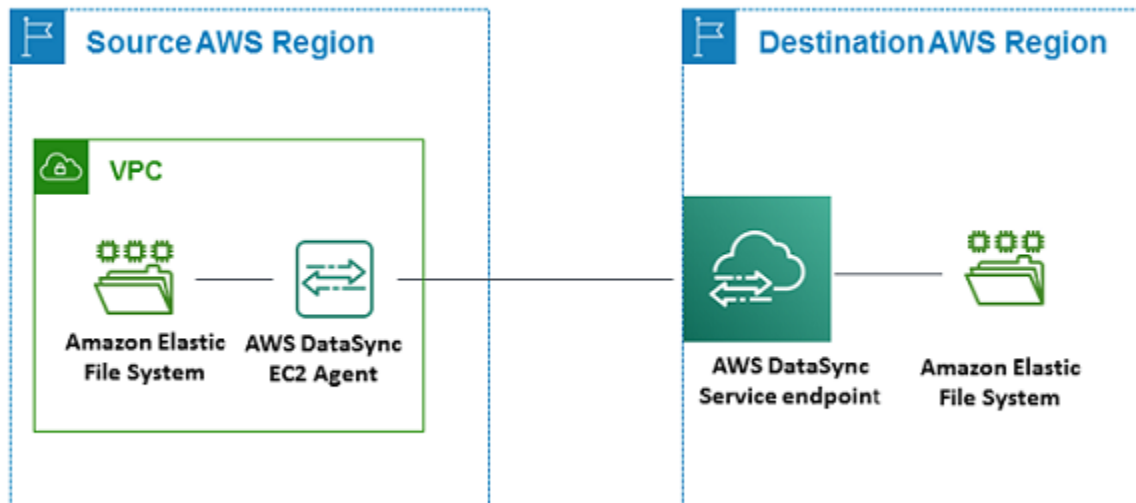
部署您的代理，使其不需要在可用区之间传输网络流量（以避免为此类流量付费）。

- 要访问您的 Amazon EFS 或 FSx for Windows 文件服务器文件系统，请在装载目标指向您的文件系统的可用区中部署代理。
- 对于自管文件系统，在文件系统所在的可用区中部署代理。

要详细了解所有AWS 区域人的数据传输价格，请参阅[Amazon EC2 按需定价](#)。

例如，下图显示了将数据从云内网络文件系统 (NFS) 传输到云中 NFS 或 Amazon S3 的DataSync架构的高级视图。





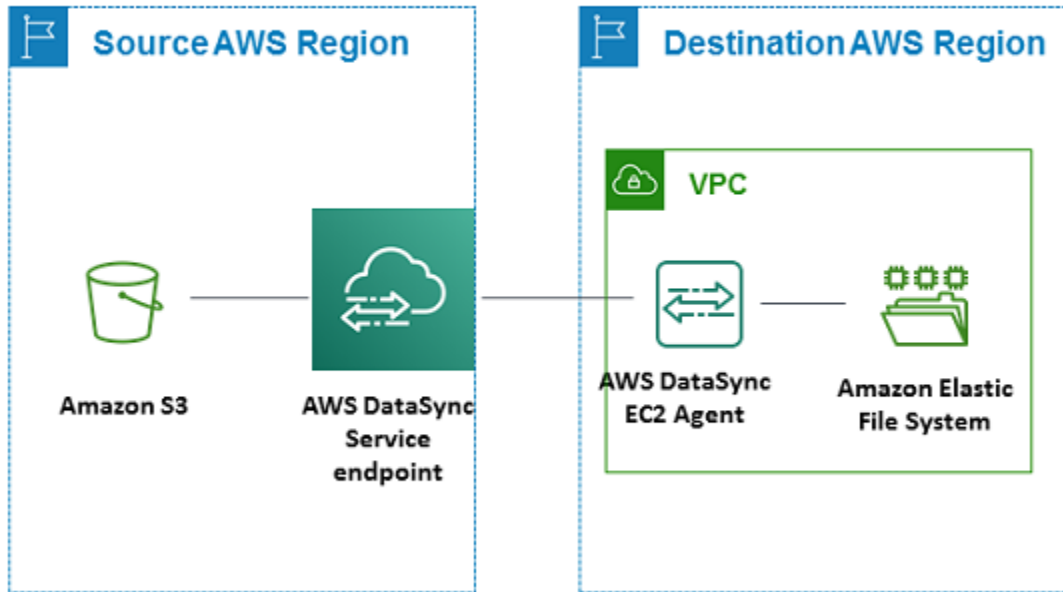
### Note

在源文件系统所在的AWS 区域和AWS 账户位置部署代理。

- 当您在两个位于不同AWS 账户的 Amazon EFS 文件系统之间进行复制时，我们建议您使用 NFS (源) 到 EFS (目标) 的传输。
- 当您在两个不同AWS 账户的 Amazon FSx 文件系统之间进行复制时，我们建议您使用服务器消息块 (SMB) (源) 到 Amazon FSx (目标) 的传输。

## 将数据从 Amazon S3 AWS

下图提供了将数据从 Amazon S3 传输到AWS文件系统 (例如 Amazon EFS 或 Amazon FSx) 的 DataSync架构的高级视图。您可以使用此架构将数据从一个传输AWS 账户到另一个架构，或者将数据从 Amazon S3 传输到自我管理的云内文件系统。



## 使用多个AWS DataSync代理进行转账

您最多可以在一个转移地点使用四个AWS DataSync代理。虽然我们建议在大多数传输中只使用一个代理，但有些传输可能包含数千万个小文件。在这种情况下，使用多个代理是有意义的。

如果您正在考虑使用多个代理，请记住以下几点：

- 在开始传输任务之前，与某个地点关联的所有代理都必须处于联机状态。如果其中一个代理处于[脱机状态](#)，则无法运行任务。
- 如果您[使用虚拟私有云 \(VPC\) 终端节点](#)进行通信AWS，则所有代理必须使用相同的终端节点和子网。
- 使用 DataSync Discovery，每个存储系统只能使用一个代理。

## 为多个 NIC 配置AWS DataSync代理

如果您将代理配置为使用多个网络适配器 (NIC)，则可以通过多个 IP 地址访问代理。您可能希望在以下情况下执行此操作：

- 最大化吞吐量-当网络适配器成为瓶颈时，您可能需要最大限度地提高代理的吞吐量。
- 网络隔离 — 出于安全原因，您的网络文件系统 (NFS)、服务器消息块 (SMB)、Hadoop 分布式文件系统 (HDFS) 或对象存储服务器可能位于缺少互联网连接的虚拟 LAN (VLAN) 上。

在典型的多适配器用例中，一个适配器被配置为代理与之通信的路由AWS（作为默认代理）。除此适配器外，NFS、SMB、HDFS 或自管对象存储位置必须与连接它们的适配器位于同一个子网中。否则，可能无法与预期的 NFS、SMB、HDFS 或对象存储位置进行通信。在某些情况下，您可以在与之通信的同一个适配器上配置 NFS、SMB、HDFS 或对象存储位置。AWS在这些情况下，该服务器的 NFS、SMB、HDFS 或对象存储流量和AWS流量流经同一个适配器。

在某些情况下，您可以将一个适配器配置为连接到 AWS DataSync 控制台，然后添加另一个适配器。在此类情况下，DataSync 会自动将路由表配置为使用第二个适配器作为首选路由。

## 管理您的AWS DataSync代理

在存储环境中部署并激活AWS DataSync代理后，即可为您AWS管理虚拟机 (VM) 设备。

### 代理软件更新

AWS自动更新您的代理的软件，包括底层操作系统和相关DataSync软件包。

DataSync仅在代理处于空闲状态时才更新代理。例如，在您的转账完成之前，您的代理不会更新。

更新后，代理可能会短暂离线。例如，这可能会在[代理激活](#)后不久在AWS更新代理时发生。

#### Warning

DataSync不支持使用云初始化指令手动更新 Amazon EC2 代理。如果您以这种方式更新代理，则可能会DataSync遇到无法激活或使用代理的互操作性问题。

### 代理故障排除

在为您AWS管理DataSync代理时，在某些情况下，您可能需要再次直接使用代理。例如，如果您的代理离线或断开与本地存储系统的连接，则可以尝试在[代理的本地控制台](#)中解决这些问题。

有关更多信息，请参阅[DataSync代理故障排除](#)。

## 编辑您的AWS DataSync代理的属性

您可以对现有AWS DataSync代理进行一些更改。如果您需要为不同的虚拟机管理程序或服务端点使用代理，则必须[创建新的](#)代理。

## 使用DataSync控制台编辑代理的属性

1. 通过 <https://console.aws.amazon.com/datasync/> 打开AWS DataSync主机。
2. 在左侧导航窗格中，选择该代理。
3. 选择要编辑的代理。
4. 选择编辑并执行所需的更改。

## AWS DataSync代理状态

下表介绍了AWS DataSync代理的状态。

客服状态	意义
在线	该代理，配置正确，可以使用。这是代理的正常运行状态。
离线	代理的虚拟机 (VM) 已关闭或代理处于不健康状态，并且已与服务断开联系五分钟或更长时间。当导致不健康状态的问题得到解决后，代理将返回联机状态。

## 使用AWS DataSync代理的本地控制台

虽然在部署AWS DataSync代理后即可对其进行AWS全面管理，但在某些情况下，您可能需要更改代理的设置或解决问题。以下是一些你为什么要通过代理的本地控制台与代理合作的示例：

- 手动将 IP 地址分配给代理的 IP 地址。
- 测试您的代理与AWS存储系统的连接。
- 向您的代理提供AWS Support访问权限以帮助解决问题（例如防火墙配置错误）。

### Important

您无需使用代理的本地控制台即可使用标准DataSync功能。

## 访问代理的本地控制台

访问本地控制台的方式取决于您使用的代理类型。

访问本地控制台 ( VMware ESXi、Linux KVM 或微软 Hyper-V )

出于安全原因，您无法远程连接到DataSync代理虚拟机 (VM) 的本地控制台。

- 如果这是您第一次使用本地控制台，请使用默认凭据登录。默认用户名为 **admin**，密码为 **password**。否则，请使用您的凭证登录。

### Note

我们建议您更改默认密码，您可以通过运行本地控制台菜单中的passwd命令来完成此操作。（主菜单5上的项目会打开命令提示符。对于 VMware 虚拟机，请选择项目6。）有关如何运行该命令的信息，请参阅[配置其他代理设置](#)。

访问本地控制台 (Amazon EC2)

要连接到 Amazon EC2 代理，您必须使用 SSH 和以下加密算法：

- SSH 密码：aes128-ctr
- 密钥交换：diffie-hellman-group14-sha1

## 获取代理激活密钥

如果您的代理尚未激活，则可以从本地控制台获取其激活密钥。只有在激活代理之前，才会显示此选项。

从本地控制台获取代理的激活密钥

1. 登录到代理的本地控制台。
2. 在“AWSDataSync激活-配置”主菜单上，输入0以获取激活密钥。
3. 输入将在AWS 区域其中激活您的代理。
4. 输入您的代理将使用的服务端点类型。选项包括公共云、联邦信息处理标准 (FIPS) 和虚拟私有云 (VPC)AWS PrivateLink。
5. 激活密钥自动生成并显示在屏幕上。选择并复制此值。

6. 使用上一步中复制的激活密钥，使用以下 `create-agent` CLI 命令创建和激活代理：

```
$ aws datasync create-agent --agent-name your-new-agent-name --activation-key generated-activation-key
```

成功激活后，此命令将返回类似以下内容的输出。

```
{  
  "AgentArn": "arn:aws:datasync:us-west-1:1234567890A:agent/agent-ID"  
}
```

您也可以使用代理创建向导在 DataSync 控制台中插入激活密钥。

激活代理后，控制台菜单显示代理 ID 和 AWS 区域。获取激活密钥的选项不再在控制台菜单中显示。


## 配置代理的网络设置

代理的默认网络配置是动态主机配置协议 (DHCP)。借助 DHCP，您的代理可获得自动分配的 IP 地址。在某些情况中，您可能需要手动将代理的 IP 地址分配为静态 IP 地址，如下所示。

将您的代理配置为使用静态 IP 地址

1. 登录到代理的本地控制台。
2. 在“AWSDataSync激活-配置”主菜单上**1**，输入开始配置网络。
3. 在 Network Configuration (网络配置) 菜单上，选择下列选项之一。

To	请执行该操作
获取有关网络适配器的信息	<p>输入 <b>1</b>。</p> <p>将显示适配器名称列表，并提示您输入适配器名称，例如 <b>eth0</b>。如果您指定的适配器正在使用中，有关该适配器的下列信息就会显示：</p> <ul style="list-style-type: none"><li>• 媒体访问控制 (MAC) 地址</li></ul>

To	请执行该操作
	<ul style="list-style-type: none"><li>• IP 地址</li><li>• 网络掩码</li><li>• 代理 IP 地址</li><li>• DHCP 启用状态</li></ul> <p>配置静态 IP 地址 ( 选项 3 ) 或设置代理的默认路由适配器 ( 选项 5 ) 时, 您可使用相同的适配器名称。</p>
配置 DHCP	<p>输入 <b>2</b>。</p> <p>系统将提示您将网络接口配置为使用 DHCP。</p>
为代理配置静态 IP 地址	<p>输入 <b>3</b>。</p> <p>系统将提示您输入网络适配器名称。</p> <div data-bbox="829 1255 1507 1528" style="border: 1px solid #f08080; border-radius: 10px; padding: 10px;"><p> <b>Important</b></p><p>如果已激活代理, 您必须从 DataSync 控制台将其关闭并重启, 然后设置才会生效。</p></div>

To	请执行该操作
将代理的所有网络配置重置为 DHCP	<p>输入 <b>4</b>。</p> <p>所有网络接口均设置为使用 DHCP。</p> <div data-bbox="829 495 1507 762" style="border: 1px solid #f08080; border-radius: 10px; padding: 10px;"><p> <b>Important</b></p><p>如果已激活代理，则必须从 DataSync 控制台关闭并重启您的代理，然后设置才会生效。</p></div>
设置代理的默认路由适配器	<p>输入 <b>5</b>。</p> <p>将显示您的代理的可用适配器，并提示您选择其中一个适配器，例如 <b>eth0</b>。</p>
编辑您的代理的域名系统 (DNS) 配置	<p>输入 <b>6</b>。</p> <p>这将显示主 DNS 和备用 DNS 服务器的可用适配器。系统将提示您提供新的 IP 地址。</p>
查看代理的 DNS 配置	<p>输入 <b>7</b>。</p> <p>这将显示主 DNS 和备用 DNS 服务器的可用适配器。</p> <div data-bbox="829 1518 1507 1738" style="border: 1px solid #add8e6; border-radius: 10px; padding: 10px;"><p> <b>Note</b></p><p>对于某些版本的 VMware 管理程序，您可以在此菜单中编辑适配器配置。</p></div>



To	请执行该操作
查看路由表	<p>输入 <b>8</b>。</p> <p>代理的默认路由将会显示。</p>

## 测试您的代理与的连接AWS

可使用代理的本地控制台测试 Internet 连接。当排查代理的网络问题时，此测试可能会很有用。

### 测试您的代理与AWS DataSync终端的连接

1. 登录到代理的本地控制台。
2. 在AWSDataSync激活-配置主菜单上，输入**2**开始测试网络连接。
3. 输入您的代理所连接的服务端点类型。有效的终端节点类型包括正在使用的公共终端节点、FIPS 和 VPC 终端节点AWS PrivateLink。

激活代理后，无需任何其他用户输入即可启动“测试网络连接”选项，因为区域和端点类型取自激活的代理信息。

- a. 要测试公共端点连接，请输入**1**，然后输入，然后AWS 区域在其中激活代理。将显示针对代理所在区域的正确终端节点的连接测试结果。有关AWS 区域和终端节点的信息，请参阅[我可以哪里使用DataSync？](#)。

所选端点中的每个端点都AWS 区域显示“通过”或“失败”消息。

- b. 要测试 FIPS 端点连接**2**，请输AWS 区域入，然后输入激活代理的。将显示针对代理所在区域的正确终端节点的连接测试结果。有关AWS 区域和终端节点的信息，请参阅[我可以哪里使用DataSync？](#)。

所选端点中的每个端点都AWS 区域显示“通过”或“失败”消息。

- c. 要测试 VPC 连接，请输入**3**。将显示您的代理的 VPC 终端节点的网络连接测试结果。

每个 VPC 终端节点都显示“通过”或“失败”消息。

有关网络和防火墙要求的信息，请参阅[AWS DataSync网络要求](#)。

## 测试代理与存储系统的连接

您可以使用控制台测试与传输中涉及的存储系统的连接，包括网络文件系统 (NFS)、服务器消息块 (SMB)、Hadoop 分布式文件系统 (HDFS) 或对象存储服务器。

### 测试与存储系统的连接

1. 登录到代理的本地控制台。
2. 在AWSDataSync激活-配置主菜单上，输入**3**开始网络测试。
3. 使用以下选项之一选择您要测试的位置类型。
  - a. 输入**1**以测试 NFS 服务器连接。
  - b. 输入**2**以测试 SMB 服务器连接。
  - c. 输入**3**，测试对象存储服务器连接。
  - d. 输入**4**以测试 HDFS 连接。
4. 输入存储服务器的 IP 地址或服务器域名。

对于 HDFS，输入 Hadoop 集群DataNode中NameNode或的 IP 地址或主机名，然后输入 TCP 端口号。

将显示指定服务器的连接测试结果，包括通过或失败，以及测试服务器的 IP 地址和端口。

## 查看代理的系统资源状态

当您登录代理控制台时，将自动检查虚拟 CPU 内核、根卷大小和 RAM。如果有任何错误或警告，则会在控制台菜单显示屏上用横幅标记，提供有关这些错误或警告的详细信息。

如果主机启动时没有错误或警告，则菜单将显示白色文本。将显示“查看系统资源检查”选项(0 Errors)。

如果出现错误或警告，控制台菜单将在菜单顶部的横幅中分别以红色和黄色显示错误和警告的数量。例如，(1 ERROR, 1 WARNING)。

### 查看系统资源检查的状态

1. 登录到代理的本地控制台。
2. 在“AWSDataSync激活-配置”主菜单上，输入**4**以查看系统资源检查的结果。

控制台为每个资源显示 [OK]、[WARNING] 或 [FAIL] 消息，如下表中所述。

对于 Amazon EC2 实例，系统资源检查会验证该实例类型是否是推荐使用的实例之一 DataSync。如果实例类型与该列表匹配，则单个结果将以绿色文本显示，如下所示。

#### [ OK ] Instance Type Check

如果 Amazon EC2 实例不在推荐列表中，则系统资源检查将验证以下资源。

- CPU 内核检查：至少需要四个内核。
- 磁盘大小检查：至少需要 80 GB 的可用磁盘空间。
- RAM 检查：每项任务需要至少 32 GiB 的 RAM 才能传输多达 2000 万个文件。每项任务至少需要 64 GiB 的 RAM 才能传输超过 2000 万个文件。
- CPU 标志检查：代理虚拟机 CPU 必须有 SSSE3 或 SSE4 指令集标志。

如果 Amazon EC2 实例不在推荐的实例列表中 DataSync，但它有足够的资源，则系统资源检查的结果将显示四个结果，全部以绿色文本显示。

对于部署在 Hyper-V、基于 Linux 内核的虚拟机 (KVM) 和 VMware 虚拟机中的代理，经过验证的资源是相同的。

还会检查 VMware 代理是否有支持的版本；不支持的版本会导致红色横幅错误。支持的版本包括 VMware 版本 6.5 和 6.7。

## 同步您的 VMware 代理上的时间

如果您使用的是 VMware 虚拟机，则可以查看网络时间协议 (NTP) 服务器配置，并将代理上的虚拟机时间与 VMware 虚拟机管理程序主机同步。

### 管理系统时间

1. 登录到代理的本地控制台。
2. 在“AWSDataSync激活-配置”主菜单上，输入**5**以管理系统的时间。
3. 在“系统时间管理”菜单上，输入**1**以查看和同步 VM 系统时间。

To	请执行该操作
查看您的 VM 时间并将其与 NTP 服务器时间同步	<p>输入 <b>1</b>。</p> <p>这将显示代理的当前时间。您的代理确定代理虚拟机与 NTP 服务器时间之间的时差，并提示您将代理时间与 NTP 时间同步。</p> <p>部署并运行代理后，在某些情况下，代理的时间可能会出现偏差。例如，假定网络中断时间延长，并且您的管理程序主机和代理没有获取时间更新。在此情况下，代理的时间与实际时间不同。当出现时间偏差时，操作（如快照）发生的预计时间和操作发生的实际时间之间会有差异。</p>
编辑 NTP 服务器配置	<p>输入 <b>2</b>。</p> <p>系统将提示您提供首选和辅助 NTP 服务器。</p>
查看 NTP 服务器配置	<p>输入 <b>3</b>。</p> <p>这将显示您的 NTP 服务器配置。</p>

## 配置其他代理设置

在DataSync代理的本地控制台中，您可以执行一些维护任务并诊断代理的问题。

在代理的本地控制台中运行配置或诊断命令

1. 登录到代理的本地控制台。
2. 在“AWSDataSync激活-配置”主菜单上，输入**5**（对于**6** VMware VM）命令提示符。
3. 使用以下命令与您的代理执行以下任务。

命令	描述
<b>dig</b>	查找有关主机的 DNS 信息。

命令	描述
<b>diskclean</b>	执行磁盘清理。
<b>exit</b>	返回控制台配置菜单。
<b>h</b>	显示可用命令列表。
<b>ifconfig</b>	显示或配置网络接口。
<b>ip</b>	显示或配置路由、设备和隧道。
<b>iptables</b>	设置和维护 IPv4 数据包过滤和网络地址转换 (NAT)。
<b>ncport</b>	测试与特定网络 TCP 端口的连通性。
<b>nping</b>	获取解决网络问题的信息。
<b>open-support-channel</b>	将代理Connect 到AWS Support。
<b>save-iptables</b>	永久保存 IP 表防火墙规则。
<b>save-routing-table</b>	保存新添加的路由表条目。
<b>sslcheck</b>	验证 SSL 证书是否有效。
<b>tcptracert</b>	收集到达目的地的 TCP 流量的traceroute 输出。

4. 按照屏幕上的说明进行操作。

## 通过以下方式向您的代理人寻求帮助AWS Support

您可以允许AWS Support访问您的AWSDataSync代理并协助您解决代理问题。默认情况下，对的AWS Support访问DataSync处于禁用状态。您可通过主机的本地控制台启用此访问。要授予AWS Support访问权限DataSync，您首先登录到主机的本地控制台，然后连接到支持服务器。

要登录在 Amazon EC2 上运行的代理，请为实例的安全组创建规则，为安全外壳 (SSH) 打开 TCP 端口 22。

**Note**

如果将新规则添加到现有安全组，则新规则适用于使用该安全组的所有实例。有关安全组以及如何添加安全组规则的更多信息，请参阅《适用于 Linux 实例 [Amazon EC2 用户指南](#)》中的[适用于 Linux 实例的 Amazon EC2 安全组](#)。

## 启用对的AWS Support访问权限AWSDataSync

1. 登录到主机的本地控制台。

如果这是您第一次登录到本地控制台，请参阅[访问代理的本地控制台](#)。

2. 在提示符处，输入**5**以打开命令提示符（对于 VMware 虚拟机，使用**6**）。
3. 输入 **h** 以打开 AVAILABLE COMMANDS 窗口。
4. 在“可用命令”窗口中，输入以下内容进行连接AWS Support：

```
open-support-channel
```

如果您将代理与 VPC 终端节点一起使用，则必须为您的支持渠道提供 VPC 终端节点 IP 地址，如下所示：

```
open-support-channel vpc-ip-address
```

您的防火墙必须允许出站 TCP 端口 22 向其启动支持通道AWS。当您连接到时AWS Support，会为您DataSync分配一个支持号码。请记住您的支持编号。

**Note**

渠道号不是传输控制协议/用户数据报协议 (TCP/UDP) 端口号。相反，它建立到服务器的安全 Shell (SSH) (TCP 22) 连接，并为该连接提供支持渠道。

5. 建立支持渠道后，请将您的支持服务号码提供给，AWS Support以便他们可以提供故障排除帮助。
6. 支持会话完成后，按**Enter**即可结束会话。
7. 输**exit**入，注销DataSync本地控制台。
8. 按照提示操作退出本地控制台。

# 删除代AWS DataSync理

删除DataSync代理后，它将不再与您的关联AWS 账户且无法撤消。

## Note

删除不会将代理的虚拟机 (VM) 从您的环境中移除。您可以重用 VM 来创建和激活新代理。

## 删除代理

1. 通过 <https://console.aws.amazon.com/datasync/> 打开AWS DataSync主机。
2. 在左侧导航窗格中，选择该代理。
3. 选择要删除的代理。
4. 选择“删除”，**delete**在出现的文本框中输入，然后选择“删除”。

## 删除代理后在 VM 或 Amazon EC2 实例上创建和激活代理

1. 删除旧代理（有关说明，请参阅前面的步骤）。请不要删除该代代代Amazon EC2 理。
2. 等到旧代理被删除并且虚拟机准备好激活，通常大约需要三分钟。或者，您可以通过检查端口 80 的状态来验证代理是否已被删除。当 VM 准备好激活时，端口 80 将打开。
3. 在现有 VM 或 Amazon EC2 实例上创建并激活新DataSync代理。有关创建DataSync代理的信息，请参阅[创建代AWS DataSync理](#)。新代理可以在不同的位置激活AWS 区域，具体取决于网络连接。

# 使用以下方法发现您的存储空间 AWS DataSync 数据查找服务

AWS DataSync 数据查找服务帮助您加速迁移至AWS。利用 DataSync Discover，您可以执行以下操作：

- 了解本地存储的使用情况 — DataSync Discovery 提供有关您的存储系统资源的详细报告，包括利用率、容量和配置信息。
- 获取有关将数据迁移到的建议 AWS — DataSync Discovery 可以为您的数据推荐AWS存储服务（例如适用于 NetApp ONTAP 的 Amazon FSx、亚马逊 EFS 和适用于 Windows 文件服务器的 Amazon FSx）。建议包括成本估算，可帮助您了解如何配置建议的存储服务。准备就绪后，您可以使用将数据迁移DataSync到AWS。

## 主题

- [将您的本地存储系统添加到 DataSync Discovery](#)
- [从事DataSync探索工作](#)
- [查看收集的存储资源信息 AWS DataSync 数据查找服务](#)
- [从中获得推荐 AWS DataSync 数据查找服务](#)
- [AWS DataSync发现状态](#)

## 将您的本地存储系统添加到 DataSync Discovery

指定要收集相关信息并为AWS DataSync 数据查找服务其提供AWS存储迁移建议的本地存储系统。

### Note

DataSyncDiscovery 目前支持运行 ONTAP 9.7 或更高版本的NetApp结构连接存储 (FAS) 和全闪存 FAS (AFF) 系统。

## 访问您的本地存储系统

要收集有关您的本地存储系统的信息，DataSyncDiscovery 需要提供对存储系统管理界面的读取访问权限的凭证。为了安全起见，DataSyncDiscovery 会将这些证书存储在中AWS Secrets Manager。



### Important

如果您在存储系统上更新这些凭据，请确保在 DataSync Discovery 中也对其进行更新。您可以使用 DataSync 控制台或操作来执行此 [UpdateStorageSystem](#) 操作。

## DataSync 探索如何使用 AWS Secrets Manager

AWS Secrets Manager 是一种秘密存储服务，用于保护数据库凭证、API 密钥和其他机密信息。DataSync Discovery 使用 Secrets Manager 来保护您提供的用于访问本地存储系统的凭证。

Secrets Manager 使用密 AWS Key Management Service 密钥加密机密。有关更多信息，请参阅 [密钥加密和解密](#)。

您可将 Secrets Manager 配置为根据您指定的计划自动轮换密钥。这使您能够将长期密钥替换为短期密钥，这有助于显著减少泄露风险。有关更多信息，请参阅 [轮换 AWS Secrets Manager 密钥](#)。

您为存储在 SecretManager 中的凭证付费。有关更多信息，请参阅 [AWS Secrets Manager 定价](#)。

## 添加您的本地存储系统

在 DataSync Discovery 收集有关存储系统的信息之前，必须提供一些有关存储系统的信息。

### 使用 DataSync 控制台

在控制台中，将 DataSync Discovery 配置为与您的本地存储系统配合使用。

### 使用控制台添加本地存储系统

1. 通过 <https://console.aws.amazon.com/datasync/> 打开 AWS DataSync 主机。
2. 在左侧导航窗格中，选择 Discovery (发现)，然后选择添加存储系统。
3. 对于存储类型，请选择要添加的存储系统的类型。
4. 在存储名称中，为您的存储系统输入一个熟悉的名称。
5. 在管理接口中，输入存储系统管理接口的域名或 IP 地址。
6. 对于服务器端口，输入访问存储系统管理接口所需的网络端口。
7. 在“凭据”中，输入访问存储系统管理界面所需的用户名和密码。

有关更多信息，请参阅 [访问您的本地存储系统](#)。

## 8. 对于代理，请执行以下操作之一：

- 选择要连接到存储系统管理界面的DataSync代理。
- 如果您尚未创建代理，请选择部署新DataSync代理。有关说明，请参阅[创建代AWS DataSync代理](#)。

部署并激活代理后，您可以完成将存储系统添加到 DataSync Discovery 中。

## 9. (可选) 选择启用日志记录。选择现有的 Amazon CloudWatch 日志组或创建新的日志组。

我们建议您启用登录，以防需要对收集存储系统信息的发现任务进行故障排除。有关更多信息，请参阅[将DataSync发现活动记录到亚马逊 CloudWatch](#)：

## 10. (可选) 选择添加标签来标记代表您的存储系统的DataSync资源。

标签是键值对。帮助您管理、筛选和搜索DataSync资源。

## 11. 选择添加存储系统。

### 使用 AWS CLI

使用 AWS Command Line Interface (AWS CLI) 将 DataSync Discovery 配置为与您的本地存储系统配合使用。

开始之前：我们建议您[启用登录功能CloudWatch](#)。

### 使用添加本地存储系统 AWS CLI

#### 1. 复制以下add-storage-system命令：

```
aws datasync add-storage-system \  
  --server-configuration ServerHostname="domain-or-ip",ServerPort=network-port \  
  --system-type storage-system-type \  
  --credentials Username="your-management-interface-username",Password="your-  
management-interface-password" \  
  --agent-arns "agent-arn"
```

#### 2. 在命令中指定以下必需参数：

- `--server-configuration ServerHostname`— 指定存储系统管理界面的域名或 IP 地址。
- `--server-configuration ServerPort`— 指定连接系统管理接口所需的网络端口。
- `--system-type`— 指定要添加的存储系统的类型。

- `--credentials`— 包括以下选项：
  - `Username`— 指定访问存储系统的管理界面所需的用户名。
  - `Password`— 指定访问存储系统的管理界面所需的密码。

有关更多信息，请参阅 [访问您的本地存储系统](#)。

- `--agent-arns`— 指定要连接到存储系统管理接口的DataSync代理。

如果您没有代理，请参阅[创建代理](#)。

### 3. (可选) 向命令添加以下任意参数：

- `--cloud-watch-log-group-arn`— 指定您要用于记录DataSync发现活动的CloudWatch日志组的 Amazon 资源名称 (ARN)。
- `--tags`— 指定Key和Value来标记代表您的存储系统的DataSync资源。

标签是键值对。帮助您管理、筛选和搜索DataSync资源。

- `--name`— 为存储系统指定名称。

### 4. 运行 `add-storage-system` 命令。

您会收到一条向您显示刚才添加的存储系统 ARN 的响应。

```
{
  "StorageSystemArn": "arn:aws:datsync:us-east-1:123456789012:system/storage-
system-abcdef01234567890"
}
```

添加存储系统后，您可以运行发现任务来收集有关存储系统的信息。

## 移除您的本地存储系统

从 DataSync Discovery 中移除本地存储系统时，您将永久删除所有相关的发现任务、收集的数据和建议。

### 使用 DataSync 控制台

1. 通过 <https://console.aws.amazon.com/datsync/> 打开AWS DataSync主机。
2. 在左侧导航窗格中，选择 Discovery (发现)，然后选择要删除的存储系统。
3. 选择操作，然后选择删除。

#### 4. 输入 **remove**，然后选择“删除”。

#### 使用 AWS CLI

##### 1. 复制以下 `remove-storage-system` 命令：

```
aws datasync remove-storage-system --storage-system-arn "your-storage-system-arn"
```

##### 2. 对于 `--storage-system-arn`，指定存储系统的 ARN。

##### 3. 运行 `remove-storage-system` 命令。

如果成功，您会收到一个带有空 HTTP 200正文的 HTTP 200 响应。

## 将DataSync发现活动记录到亚马逊 CloudWatch

启用亚马逊登录后CloudWatch，您可以更轻松地解决 DataSync Discovery 的问题。例如，如果您的发现任务中断，则可以检查日志以找到问题。如果您在问题发生后的 12 小时内解决了问题，您的发现任务就会从中断的地方继续进行。

如果您使用控制台添加本地存储系统，则DataSync可以自动为您启用日志记录。

如果您使用配置系统AWS CLI，则必须使用允许将事件记录DataSync到[日志组的资源策略创建](#)日志组。您可以使用类似于DataSync任务的[日志组资源策略](#)，但有一些区别：

- 对于服务主体，使用 `discovery-datasync.amazonaws.com`。
- 如果您使用ArnLike条件，请指定存储系统 ARN，如下所示：

```
"ArnLike": {  
  "aws:SourceArn": [  
    "arn:aws:datasync:region:account-id:system/*"  
  ]  
},
```

## 从事DataSync探索工作

部署AWS DataSync代理并将本地存储系统添加到 DataSync Discovery 后，您可以运行发现任务来收集有关系统的信息并获取AWS迁移建议。

## 开始探索任务

您最多可以运行 31 天的发现任务。存储系统同时只能有一个活动发现任务。发现任务收集的信息在任务结束后 60 天内可用（除非您在此之前从 DataSync Discovery 中移除了相关的存储系统）。

### Tip

DataSync发现任务运行的时间越长，Discovery 可以提供更准确的建议。我们建议将发现任务运行至少 14 天。

### 使用 DataSync 控制台

使用控制台，你可以在短短一天内运行发现任务。要运行少于一天的发现任务，请使用AWS CLI。

1. 通过 <https://console.aws.amazon.com/datasync/> 打开AWS DataSync主机。
2. 在左侧导航窗格中，选择 Discovery（发现），然后选择要在其上运行发现作业的存储系统。
3. 选择操作，然后选择开始。
4. 在“持续时间”中，选择您希望发现作业运行多长时间。
5. 选择“开始发现作业”。

### 使用 AWS CLI

使用 AWS Command Line Interface (AWS CLI)，你可以在短短 1 小时内运行发现作业。

1. 复制以下start-discovery-job命令：

```
aws datasync start-discovery-job \  
  --storage-system-arn "your-storage-system-arn" \  
  --collection-duration-minutes discovery-job-duration
```

2. 在命令中指定以下参数：
  - --storage-system-arn— 指定您添加到 [DataSync Discovery 的本地存储系统的 Amazon 资源名称 \(ARN\)](#)。
  - --collection-duration-minutes— 指定您希望发现作业在分钟内运行多长时间。输入介于60（1小时）和44640（31天）之间的值。
3. 运行 start-discovery-job 命令。

你得到的回复显示了你刚刚开始地发现任务。

```
{
  "DiscoveryJobArn": "arn:aws:datsync:us-east-1:123456789012:system/storage-
system-abcdef01234567890/job/discovery-job-12345678-90ab-cdef-0abc-021345abcdef6"
}
```

开始发现任务后不久，您就可以[开始查看该任务收集的信息](#)（包括存储系统的容量和使用情况）。

## 停止发现作业

随时停止发现作业。您仍然可以[获得有关已停止工作的建议](#)。

使用 DataSync 控制台

1. [通过 https://console.aws.amazon.com/datsync/](https://console.aws.amazon.com/datsync/) 打开AWS DataSync主机。
2. 在左侧导航窗格中，选择 Discovery（发现），然后选择您正在运行发现任务的存储系统。
3. 选择操作，然后选择停止（保留数据）。

使用 AWS CLI

1. 复制以下stop-discovery-job命令：

```
aws datsync stop-discovery-job --discovery-job-arn "your-discovery-job-arn"
```

2. 对于--discovery-job-arn，指定当前正在运行的发现任务的 ARN。
3. 运行 stop-discovery-job 命令。

如果成功，您会收到一个带有空 HTTP 200正文的 HTTP 200 响应。

## 查看收集的存储资源信息 AWS DataSync 数据查找服务

AWS DataSync 数据查找服务收集有关您的本地存储系统的信息，这些信息可以帮助您了解其存储资源的配置、性能和使用情况。DataSyncDiscovery 使用这些信息来生成有关将数据迁移到的建议 AWS。

发现任务可以为您提供有关存储系统资源（例如其卷）的以下信息：

- 总存储容量、可用容量和在用存储容量
- 资源中公用 Internet 文件系统 (CIFS) 共享的数量以及资源是否可通过网络文件系统 (NFS) 获得
- 数据传输协议

#### Note

DataSyncDiscovery 目前不会表明您的资源使用互联网小型计算机系统接口 (iSCSI) 协议，即使该协议已配置为使用该协议。

- 性能 (例如 IOPS、吞吐量和延迟)

## 查看收集的有关您的存储系统的信息

开始DataSync发现任务后不久，您就可以开始查看 Discovery 正在收集有关您的本地存储系统的哪类信息。

您可以使用以下选项查看此类信息：

- DataSync控制台 — 获取有关 DataSync Discovery 可以收集的所有存储系统资源的可视化数据，包括利用率、容量和配置数据。您可以查看存储系统的资源概述，也可以重点关注单个资源。
- [DescribeStorageSystemResources](#)操作-获取有关 DataSync Discovery 可以收集的所有存储系统资源的数据，包括利用率、容量和配置数据。
- [DescribeStorageSystemResourceMetrics](#)操作-获取 DataSync Discovery 可以收集的有关存储系统中特定资源的性能和容量信息。

### 使用 DataSync 控制台

1. 通过 <https://console.aws.amazon.com/datasync/> 打开AWS DataSync主机。
2. 在左侧导航窗格中，选择 Discovery (发现)，然后选择 DataSync Discovery 正在收集有关信息的存储系统。

在“卷”面板中，您可以看到有关存储系统资源的基本指标。

3. 选择资源以在“容量和性能数据”选项卡上查看有关该资源的更多详细信息。

您可以看到显示资源容量、IOPS 峰值等信息的图表。

## 使用 AWS CLI

以下步骤说明如何将此[DescribeStorageSystemResources](#)操作与AWS CLI。

1. 复制以下describe-storage-system-resources命令：

```
aws datasync describe-storage-system-resources \  
  --discovery-job-arn "your-discovery-job-arn" \  
  --resource-type "storage-system-resource-type"
```

2. 在命令中指定以下参数：

- --discovery-job-arn— 指定您运行的[发现任务](#)的 Amazon 资源名称 ( ARN )。
- --resource-type— 根据需要有关哪种存储系统资源的信息，指定以下值之一：
  - CLUSTER
  - SVM
  - VOLUME

3. ( 可选 ) 使用要了解的存储系统资源的 ID 指定--resource-ids参数。

4. 运行 describe-storage-system-resources 命令。

以下示例响应返回发现任务收集到的有关存储系统中两个卷的信息。

请注意，NONE每个卷都RecommendationStatus是。要获得AWS存储建议，必须先运行generate-recommendations命令，然后再运行describe-storage-system-resources命令。有关更多信息，请参阅[获取建议](#)：

```
{  
  "ResourceDetails": {  
    "NetAppONTAPVolumes": [  
      {  
        "VolumeName": "vol1",  
        "ResourceId": "a1b2c3d4-5678-90ab-cdef-EXAMPLE11111",  
        "CifsShareCount": 0,  
        "SecurityStyle": "unix",  
        "SvmUuid": "a1b2c3d4-5678-90ab-cdef-EXAMPLEeaaaaa",  
        "SvmName": "my-svm",  
        "CapacityUsed": 409600,  
        "CapacityProvisioned": 1099511627776,  
        "LogicalCapacityUsed": 409600,  
        "NfsExported": true,  
      }  
    ]  
  }  
}
```



```
    "SnapshotCapacityUsed": 573440,
    "MaxP95Performance": {
      "IopsRead": 251.0,
      "IopsWrite": 44.0,
      "IopsOther": 17.0,
      "IopsTotal": 345.0,
      "ThroughputRead": 2.06,
      "ThroughputWrite": 0.88,
      "ThroughputOther": 0.11,
      "ThroughputTotal": 2.17,
      "LatencyRead": 0.06,
      "LatencyWrite": 0.07,
      "LatencyOther": 0.13
    },
    "Recommendations": [],
    "RecommendationStatus": "NONE"
  },
  {
    "VolumeName": "root_vol",
    "ResourceId": "a1b2c3d4-5678-90ab-cdef-EXAMPLE22222",
    "CifsShareCount": 0,
    "SecurityStyle": "unix",
    "SvmUuid": "a1b2c3d4-5678-90ab-cdef-EXAMPLEeaaaaa",
    "SvmName": "my-svm",
    "CapacityUsed": 462848,
    "CapacityProvisioned": 1073741824,
    "LogicalCapacityUsed": 462848,
    "NfsExported": true,
    "SnapshotCapacityUsed": 421888,
    "MaxP95Performance": {
      "IopsRead": 261.0,
      "IopsWrite": 53.0,
      "IopsOther": 23.0,
      "IopsTotal": 360.0,
      "ThroughputRead": 10.0,
      "ThroughputWrite": 2.0,
      "ThroughputOther": 4.0,
      "ThroughputTotal": 12.0,
      "LatencyRead": 0.25,
      "LatencyWrite": 0.3,
      "LatencyOther": 0.55
    },
    "Recommendations": [],
    "RecommendationStatus": "NONE"
  }
}
```

```
    }  
  ]  
}  
}
```

## 从中获得推荐 AWS DataSync 数据查找服务

在AWS DataSync 数据查找服务收集了有关您的本地存储系统的信息后，它可以建议按资源将您的数据移动到以下一项或多项AWS存储服务：

- [适用于 ONTAP 的亚马逊 FS NetApp x](#)
- [Amazon Elastic File System \(Amazon EFS\)](#)
- [Amazon FSx for Windows File Server](#)

## 建议中包含什么？

DataSync发现建议包括存储配置和成本估算，以帮助您选择适合您的数据的AWS存储服务。

### AWS存储配置

DataSyncDiscovery 提供有关您可能需要如何配置推荐的AWS存储服务的信息。存储配置旨在优化成本，同时根据发现工作期间收集的信息，帮助满足存储性能和容量需求。

存储配置只是近似值，可能不包括AWS存储服务提供的所有功能。有关更多信息，请参阅 [建议中未包含什么？](#)

### 估算费用

DataSyncDiscovery 为其推荐的每项AWS存储服务提供了估计的每月成本。费用基于标准AWS定价，仅提供AWS费用的估算值。它不包括任何可能适用的税款。您的实际费用取决于各种因素，包括您对AWS服务的使用。

估算的费用也不包括将数据迁移到的一次性或定期费用AWS。

## 建议中未包含什么？

DataSyncDiscovery 不会推荐不满足您的存储配置需求的存储服务。AWS

此外，在确定建议时，目前未考虑以下AWS存储能力：

- 适用于 NetApp ONTAP 的 Amazon FSx — 单可用区部署和备份存储
- 亚马逊 EF S — EFS 单区存储类和备份存储
- 适用于 Windows 文件服务器的 Amazon FsX — 单可用区部署和备份存储

## 获取建议

您可以在发现任务完成后、停止作业时生成AWS存储建议，有时甚至在任务完成但从存储系统收集信息时遇到了一些问题。

在某些情况下，您可能无法获得建议（例如，如果您的发现工作失败了）。有关更多信息，请参阅[推荐状态](#)：

### Tip

在开始迁移到之前AWS，请与您的AWS客户团队一起查看 DataSync Discovery 建议。

## 使用 DataSync 控制台

1. 通过 <https://console.aws.amazon.com/datasync/> 打开AWS DataSync主机。
2. 在左侧导航窗格中，选择 Discovery（发现），然后选择运行发现任务的存储系统。
3. 选择要推荐的存储资源（例如集群、SVM 或卷）。
4. 如果存储资源处于“准备生成建议”状态，请选择存储资源名称。
5. 在存储资源页面上，转到“推荐”选项卡，然后选择“获取建议”。

推荐一旦可用，就会显示在同一个选项卡上。

## 使用 AWS CLI

1. 复制以下describe-discovery-job命令：

```
aws datasync describe-discovery-job --discovery-job-arn "your-discovery-job-arn"
```

2. 对于--discovery-job-arn参数，请指定您在存储系统上运行的[发现任务](#)的 Amazon 资源名称（ARN）。
3. 运行 describe-discovery-job 命令。

如果您的回复包含不Status是FAILED，则可以继续。如果看到FAILED，则必须在存储系统上运行其他发现任务才能尝试生成建议。

4. 如果您的发现任务成功完成，请跳过此步骤。否则请执行以下操作以手动生成建议：
  - a. 复制以下generate-recommendations命令：

```
aws datasync generate-recommendations \  
  --discovery-job-arn "your-discovery-job-arn" \  
  --resource-type cluster-svm-volume \  
  --resource-ids storage-resource-UUIDs
```

- b. 对于--discovery-job-arn参数，请指定您在步骤 2 中指定的同一发现任务的 ARN。
  - c. 对于--resource-type参数，RESOURCE根据您想要推荐的资源类型指CLUSTER定SVM、或。
  - d. 对于--resource-ids参数，指定要推荐的资源的通用唯一标识符 (UUID)。
  - e. 运行 generate-recommendations 命令。
  - f. 等待直到响应中的RecommendationStatus元素处于COMPLETED状态，然后转到下一步。
5. 复制以下describe-storage-system-resources命令：

```
aws datasync describe-storage-system-resources \  
  --discovery-job-arn "your-discovery-job-arn" \  
  --resource-type cluster-svm-volume
```

6. 在命令中指定以下参数：
  - --discovery-job-arn— 指定您在步骤 2 中指定的同一发现任务的 ARN。
  - --resource-type— 指定您生成建议的资源类型（例如，VOLUME）。
7. 运行 describe-storage-system-resources 命令。

#### Note

在回复中，如果您看不到 COMPLETEDRecommendationStatus，请查看[推荐状态](#)以获取更多信息。您可能需要重试生成建议。

在此示例响应中，Recommendations元素建议了几项AWS存储服务，您可以在其中迁移特定卷、如何配置服务以及估计的每月AWS存储成本。

```
{
  "Recommendations": [{
    "StorageType": "fsxOntap",
    "StorageConfiguration": {
      "StorageCapacityGB": "1024",
      "ProvisionedIOpsMode": "AUTOMATIC",
      "CapacityPoolGB": "0",
      "TotalIOps": "0",
      "DeploymentType": "Multi-AZ",
      "ThroughputCapacity": "128"
    },
    "EstimatedMonthlyStorageCost": "410.0"
  },
  {
    "StorageType": "efs",
    "StorageConfiguration": {
      "InfrequentAccessStorageGB": "1",
      "StandardStorageGB": "1",
      "InfrequentAccessRequests": "0",
      "ProvisionedThroughputMBps": "0",
      "PerformanceMode": "General Purpose",
      "ThroughputMode": "Bursting"
    },
    "EstimatedMonthlyStorageCost": "1.0"
  }
],
  "RecommendationStatus": "COMPLETED"
}
```

## AWS DataSync发现状态

您可以检查发现任务的状态以及是否AWS DataSync 数据查找服务可以为AWS迁移提供存储建议。

### Discover 状态

使用下表来了解您的发现工作发生了什么。

控制台状态	API 状态	描述
正在进行中	RUNNING	您的发现任务正在运行。该任务在您指定的持续时间内收集有关您的本地存储系统的数据。
出错了	WARNING	您的发现任务遇到了错误，目前无法收集数据。查看亚马逊 CloudWatch 日志并在 12 小时内解决这些问题，否则工作将被终止。
Stopped (已停止)	STOPPED	在发现任务预计完成之前，您就停止了该任务。
Completed	COMPLETED	您的发现任务成功地从本地存储系统收集了所有数据。
已完成，但有问题	COMPLETED_WITH_ISSUES	在发现工作期间，Disco DataSync very 有时无法收集数据。有关详细信息，请参阅您的 CloudWatch 日志。
已终止	TERMINATED	由于未解决的问题和未收集到某些数据，您的发现任务被取消。有关详细信息，请参阅您的 CloudWatch 日志。
Failed	FAILED	您的发现任务遇到了问题，无法从本地存储系统收集数据。有关详细信息，请参阅您的 CloudWatch 日志。

## 推荐状态

使用下表了解针对特定本地存储资源的 DataSync Discovery 建议是否可供查看。

控制台状态	API 状态	描述
暂不可用	NONE	您还无法生成推荐。尝试在发现任务完成后生成建议。
已准备好生成	NONE	您的发现工作收集的数据足以让 DataSync Discovery 提供建议。如果您提前停止了发现任务或任务已完成但在数据收集方面遇到了问题，则可以生成建议。
正在生成	IN_PROGRESS	DataSyncDiscovery 正在研究你的建议。这需要多长时间取决于你为多少资源生成推荐。如果您使用的是控制台，则可能需要几分钟时间才能生成存储资源建议。
可用	COMPLETED	您可以查看建议。
Failed	FAILED	DataSyncDiscovery 无法生成推荐。您可以查看CloudWatch 日志以找出问题所在，然后尝试再次生成建议。
Unavailable	NONE	不提供建议。如果发现任务失败或存储资源出现问题，您可能会看到此状态。
不匹配	COMPLETED	DataSyncDiscovery 目前不支持满足存储资源需求的存储服务。AWS

# 通过以下方式传输您的数据 AWS DataSync

借助AWS DataSync，您可以从本地、内部、其他云端和边缘的存储中移动数据。AWS

## 主题

- [我可以在哪里传输我的数据AWS DataSync ?](#)
- [使用本地存储进行传输 AWS DataSync](#)
- [使用以下方式向存储器传输或从AWS存储器传输 AWS DataSync](#)
- [通过以下方式传输到其他云存储或从其他云存储进行传输AWS DataSync](#)
- [使用以下方式向边缘或离线存储传输或从边缘存储器传输AWS DataSync](#)
- [如何AWS DataSync处理元数据和特殊文件](#)
- [移除传AWS DataSync送地点](#)
- [处理AWS DataSync转移任务](#)

## 我可以在哪里传输我的数据AWS DataSync ?

可供集群AWS DataSync取决于以下因素：

- 您的转账来源和目的地[位置](#)
- 如果你的位置不同AWS 账户
- 如果你的位置不同AWS 区域

您可以使用以下类型的位置配置：

- 网络文件系统 (NFS)
- 服务器消息块
- Hadoop File File Systfs
- 对象存储
- Amazon S3
- Amazon Elastic File System (Amazon EFS)
- Amazon FSx for Windows File Server



- Amazon FSx for Lustre
- Amazon FSx for OpenZFS
- Amazon FSx for NetApp or ONTAP
- Microsoft Azure Blob 存储

### Note

请记住，DataSync它所支持的每个存储系统都没有位置类型。例如，要从 Google Cloud Storage 中移动数据，您需要将传输源配置为对象存储位置。

## 支持在同一时间传输AWS 账户

DataSync支持在以下与之关联的存储系统之间进行传输AWS 账户。

来源 ( 来自 )	目的地 ( 至 )
<ul style="list-style-type: none"> <li>• NFS</li> <li>• SMB</li> <li>• HDFS</li> <li>• 对象存储</li> <li>• 谷歌云存储 ( 作为对象存储位置 )</li> <li>• Microsoft ( 作为 SMB 地点 )</li> <li>• Microsoft Azure Blob 存储</li> <li>• AWS Snowcone ( 作为 NFS 位置 )</li> </ul>	<ul style="list-style-type: none"> <li>• Amazon S3 ( 英寸AWS 区域 )</li> <li>• Amazon EFS</li> <li>• FSx for Windows File Server</li> <li>• FSx for Lustre</li> <li>• FSx for OpenZFS</li> <li>• FSx for ONTAP</li> </ul>
<ul style="list-style-type: none"> <li>• Amazon S3 ( 英寸AWS 区域 )</li> <li>• Amazon EFS</li> <li>• FSx for Windows File Server</li> <li>• FSx for Lustre</li> <li>• FSx for OpenZFS</li> <li>• FSx for ONTAP</li> </ul>	<ul style="list-style-type: none"> <li>• NFS</li> <li>• SMB</li> <li>• HDFS</li> <li>• 对象存储</li> <li>• 谷歌云存储 ( 作为对象存储位置 )</li> <li>• Microsoft</li> </ul>

来源 (来自)	目的地 (至)
	( 作为 SMB 地点 ) • AWS Snowcone ( 作为 NFS 位置 )
<ul style="list-style-type: none"> <li>• Amazon S3 ( 英寸AWS 区域 )</li> <li>• Amazon EFS</li> <li>• FSx for Windows File Server</li> <li>• FSx for Lustre</li> <li>• FSx for OpenZFS</li> <li>• FSx for ONTAP</li> </ul>	<ul style="list-style-type: none"> <li>• Amazon S3 ( 英寸AWS 区域 )</li> <li>• Amazon EFS</li> <li>• FSx for Windows File Server</li> <li>• FSx for Lustre</li> <li>• FSx for OpenZFS</li> <li>• FSx for ONTAP</li> </ul>
<ul style="list-style-type: none"> <li>• Amazon S3 ( 在AWS区域中 )</li> </ul>	<ul style="list-style-type: none"> <li>• Amazon S3AWS Outposts</li> </ul>
<ul style="list-style-type: none"> <li>• Amazon S3AWS Outposts</li> </ul>	<ul style="list-style-type: none"> <li>• Amazon S3 ( 英寸AWS 区域 )</li> </ul>

## 支持跨境传输AWS 账户

DataSync支持不同存储系统之间的某些传输AWS 账户。虽然通常您不需要DataSync代理在两者之间AWS 服务进行传输，但当此类传输仅涉及 Amazon EFS 或 Amazon FSx 文件系统时，则需要代理。

来源 (来自)	目的地 (至)
<ul style="list-style-type: none"> <li>• Amazon EFS ( 作为 NFS 位置 )</li> <li>• FSx for Windows File Server ( 作为 SMB 地点 )</li> </ul>	<ul style="list-style-type: none"> <li>• Amazon S3 ( 英寸AWS 区域 )</li> <li>• Amazon EFS</li> <li>• FSx for Windows File Server</li> <li>• FSx for Lustre</li> <li>• FSx for OpenZFS</li> <li>• FSx for ONTAP</li> </ul>
<ul style="list-style-type: none"> <li>• Amazon S3 ( 英寸AWS 区域 )</li> </ul>	<ul style="list-style-type: none"> <li>• Amazon S3 ( 英寸AWS 区域 )</li> <li>• Amazon EFS</li> <li>• FSx for Windows File Server</li> <li>• FSx for Lustre</li> </ul>

来源 (来自)	目的地 (至)
	<ul style="list-style-type: none"> <li>FSx for OpenZFS</li> <li>FSx for ONTAP</li> </ul>
<ul style="list-style-type: none"> <li>Amazon S3 (英寸AWS 区域)</li> <li>Amazon EFS</li> <li>FSx for Windows File Server</li> <li>FSx for Lustre</li> <li>FSx for OpenZFS</li> <li>FSx for ONTAP</li> </ul>	<ul style="list-style-type: none"> <li>Amazon S3 (英寸AWS 区域)</li> </ul>
<ul style="list-style-type: none"> <li>NFS</li> <li>SMB</li> <li>HDFS</li> <li>对象存储</li> </ul>	<ul style="list-style-type: none"> <li>Amazon S3 (英寸AWS 区域)</li> </ul>

## 支持在同一时间传输AWS 区域

在同一区域AWS 区域 (包括[默认禁用的区域](#)) 内传输数据时没有任何限制。有关更多信息，请参阅[AWS 区域支持者DataSync](#)。

## 支持跨境传输AWS 区域

除以下情况DataSync外，您可以在[AWS 区域支持者](#)之间传输数据：

- 对于AWS GovCloud (US)区域，您只能在AWS GovCloud (美国东部) 和AWS GovCloud (美国西部) 之间转移。
- 如果[默认情况下，其中一个或两个区域处于禁用状态](#)，则无法在区域之间传输。

当您在不同的位置之间AWS 服务传输数据时AWS 区域，两个位置中的一个必须位于您所使用的区域DataSync。

**⚠ Important**

您需要为之间传输的数据付费AWS 区域。此传输按从源区域向目标区域传输数据进行计费。有关更多信息，请参阅[数据传输定价](#)。

## 使用本地存储进行传输 AWS DataSync

使用AWS DataSync，您可以将数据从多个本地存储系统传输到AWS。

**📘 Note**

您也可以使用将数据移动DataSync到本地存储系统。有关更多信息，请参阅[我可以在哪里传输我的数据AWS DataSync？](#)：

### 主题

- [配置来自 NFS 文件服务器的AWS DataSync传输](#)
- [配置来自 SMB 文件服务器的AWS DataSync传输](#)
- [配置来自 HDFS 的AWS DataSync传输](#)
- [配置来自对象存储系统的AWS DataSync传输](#)

## 配置来自 NFS 文件服务器的AWS DataSync传输

要从网络文件系统 (NFS) 文件服务器传输数据，必须创建AWS DataSync传输位置。

### 创建您的 NFS 传输位置

在开始之前，请注意以下事项：

- 您需要一个要从中传输数据的 NFS 文件服务器。
- 如果您需要复制 NFS 版本 4 访问控制列表 (ACL)，请[咨询AWS存储专家](#)。DataSync不支持自己复制这些 ACL。

### 使用控制台创建 NFS 位置

1. [通过 https://console.aws.amazon.com/datasync/](https://console.aws.amazon.com/datasync/) 打开AWS DataSync主机。

2. 在左侧导航窗格中，展开“数据传输”，然后选择“位置”和“创建位置”。
3. 对于位置类型，请选择网络文件系统 (NFS)。
4. 对于代理，选择一个或多个要连接到 NFS 文件服务器的DataSync代理。

如果您选择多个代理，请务必了解[为一个地点使用多个代理](#)。

5. 对于 NFS 服务器，请输入DataSync代理将挂载 NFS 文件服务器的域名系统 ( DNS ) 名称或 IP 地址。
6. 对于 Mount path (挂载路径)，输入 NFS 位置的挂载路径。

此路径必须是由 NFS 文件服务器导出的路径，或者是导出路径的一个子目录。此路径应使其能由网络中的其他 NFS 客户端挂载。有关如何解决挂载路径问题的信息，请参阅[我的任务状态为不可用，表示存在装载错误](#)。

要传输指定文件夹中的所有数据，DataSync必须有权读取所有数据。为此，请为 NFS 导出配置，no\_root\_squash或者确保DataSync要传输的文件权限所有用户进行读取访问。通过执行上述任一操作，将允许代理读取这些文件。要使代理能够访问目录，您还必须开启全部执行访问权限。

7. ( 可选 ) 展开其他设置并选择特定 NFS 版本DataSync，以便在访问文件服务器时使用。

默认情况下，DataSync使用 NFS 版本 4.1。DataSync还支持 NFS 3.x 和 4.0。

8. ( 可选 ) 选择添加标签来标记您的 NFS 位置。

标签是键值对，帮助您管理、筛选和搜索位置。我们建议至少为您的位置创建一个名称标签。

9. 选择创建地点。

## 配置来自 SMB 文件服务器的AWS DataSync传输

要从服务器消息块 (SMB) 文件服务器传输数据，必须创建AWS DataSync传输位置。

### 访问 SMB 文件服务器

DataSync使用 SMB 协议连接到您的文件服务器，并使用您提供的凭据进行身份验证。

#### 主题

- [支持的 SMB 协议版本](#)
- [所需的权限](#)

## 支持的 SMB 协议版本

默认情况下，DataSync会根据与 SMB 文件服务器的协商自动选择 SMB 协议版本。

您也可以配置DataSync为使用特定的 SMB 版本，但我们建议仅在无法自动与 SMB 文件服务器协商的情况下DataSync才这样做。（DataSync支持 SMB 版本 1.0 和更高版本。）

有关DataSync控制台和 API 中的选项列表，请参阅下表：

控制台选项	API 选项	描述
自动	AUTOMATIC	DataSync并且 SMB 文件服务器会协商 2.1 和 3.1.1 之间相互支持的 SMB 的最高版本。  这是默认和推荐选项。如果您改为选择文件服务器不支持的特定版本，则可能会出现 Operation Not Supported 错误。
SMB 3.0.2	SMB3	将协议协商限制为仅 SMB 版本 3.0.2。
SMB 2.1	SMB2	将协议协商限制为仅 SMB 版本 2.1。
SMB 2.0	SMB2_0	将协议协商限制为仅 SMB 版本 2.0。
SMB 1.0	SMB1	将协议协商限制为仅 SMB 版本 1.0。

## 所需的权限

DataSync需要有权挂载和访问您的 SMB 位置的用户。这可以是 Windows 文件服务器上的本地用户，也可以是在 Microsoft Active Directory 中定义的域用户。

要设置对象所有权，DataSync需要SE\_RESTORE\_NAME权限，该权限通常授予内置 Active Directory 组 Backup 操作员和域管理员的成员。向用户DataSync提供此权限还有助于确保除了 NTFS 系统访问控制列表 (SACL) 之外的文件、文件夹和文件元数据的足够权限。

复制 SACL 需要其他权限。具体而言，这需要 Windows SE\_SECURITY\_NAME 权限，该权限授予域管理员组的成员。如果将任务配置为复制 SACL，请确保用户具有所需的权限。要了解有关将任务配置为复制 SACL 的更多信息，请参阅[管理文件、对象和元数据的AWS DataSync传输方式](#)。

在 SMB 文件服务器和 Amazon FSx for Windows File Server 系统之间复制数据时，源和目标位置必须属于同一个微软 Active Directory 域，或者它们的域之间必须有 Active Directory 信任关系。

## 创建您的 SMB 传输地点

在开始之前，您需要 SMB 文件服务器，以便从中传输数据。

### 使用控制台创建 SMB 位置

1. 通过 <https://console.aws.amazon.com/datasync/> 打开AWS DataSync主机。
2. 在左侧导航窗格中，展开“数据传输”，然后选择“位置”和“创建位置”。
3. 对于 Location type (位置类型)，选择 Server Message Block (SMB) (服务器消息块 (SMB))。

您可在稍后将此位置配置为源或目标。

4. 对于代理，选择一个或多个要连接到 SMB 文件服务器的DataSync代理。

如果您选择多个代理，请务必了解[为一个地点使用多个代理](#)。

5. 对于 SMB 服务器，请输入DataSync代理要挂载 SMB 文件服务器的域名系统 ( DNS ) 名称或 IP 地址。

#### Note

您无法指定 IP 版本 6 ( IPv6 ) 地址。

6. 在共享名称中，输入您的 SMB 文件服务器导出的共享名称，该共享DataSync将在其中读取或写入数据。

可以在共享路径中包含子目录 ( 例如， /path/to/subdirectory )。确保网络中的其他 SMB 客户端也可以挂载此路径。

要复制子目录中的所有数据，DataSync必须能够挂载 SMB 共享并访问其所有数据。有关更多信息，请参阅[所需的权限](#)：

7. ( 可选 ) 展开其他设置，选择DataSync要在访问文件服务器时使用的 SMB 版本。

默认情况下，DataSync会根据与 SMB 文件服务器的协商自动选择版本。有关信息，请参阅 [支持的 SMB 协议版本](#)。

8. 对于用户，输入一个可以装载您的 SMB 文件服务器并有权访问传输中涉及的文件和文件夹的用户名。

有关更多信息，请参阅[所需的权限](#)：

9. 在“密码”中，输入可以安装您的 SMB 文件服务器并有权访问传输中涉及的文件和文件夹的用户的密码。
10. (可选) 在“域”中，输入 SMB 文件服务器所属的 Windows 域名称。
11. (可选) 选择添加标签来标记您的 SMB 位置。

标签是键值对，帮助您管理、筛选和搜索位置。我们建议至少为您的位置创建一个名称标签。

12. 选择创建地点。

## 配置来自 HDFS 的 AWS DataSync 传输

要从 Hadoop 分布式文件系统 (HDFS) 传输数据，您必须创建 AWS DataSync 传输位置。

### 访问 HDFS 集群

要连接到 HDFS 集群，请 DataSync 使用在 HDFS 集群附近部署的代理。要了解有关 DataSync 代理的更多信息，请参阅 [与 AWS DataSync 代理合作](#)。DataSync 代理充当 HDFS 客户端，与集群 DataNodes 中的 NameNodes 和通信。

启动任务时，DataSync 查询 NameNode 集群上文件和文件夹的位置。如果将 HDFS 位置配置为源，则从集群 DataNodes 中的 DataSync 读取文件和文件夹数据，并将数据复制到目标。如果将 HDFS 位置配置为目标，则 DataSync 将文件和文件夹从目标写入集群 DataNodes 中的。在运行 DataSync 任务之前，请验证代理与 HDFS 集群的连接。有关更多信息，请参阅 [测试代理与存储系统的连接](#)：

### 身份验证

连接到 HDFS 集群时，DataSync 支持简单身份验证或 Kerberos 身份验证。要使用简单身份验证，请提供具有 HDFS 集群读写权限的用户名。要使用 Kerberos 身份验证，请提供 Kerberos 配置文件、Kerberos 密钥表 (keytab) 文件和 Kerberos 主体名称。Kerberos 主体的证书必须位于提供的密钥表文件中。

### 加密

使用 Kerberos 身份验证时，DataSync 支持在 DataSync 代理和 HDFS 集群之间传输的数据进行加密。使用您的 HDFS 集群上的保护质量 (QOP) 配置设置以及在创建 HDFS 位置时指定 QOP 设置来加密您的数据。QOP 配置包括数据传输保护和远程程序调用 (RPC) 保护设置。

DataSync 支持以下 Kerberos 加密类型：

- des-cbc-crc



- des-cbc-md4
- des-cbc-md5
- des3-cbc-sha1
- arcfour-hmac
- arcfour-hmac-exp
- aes128-cts-hmac-sha1-96
- aes256-cts-hmac-sha1-96
- aes128-cts-hmac-sha256-128
- aes256-cts-hmac-sha384-192
- camellia128-cts-cmac
- camellia256-cts-cmac

您还可以为 HDFS 集群配置，使用透明数据加密 (TDE)。使用简单身份验证时，对启用 TDE 的集群进行 DataSync 读取和写入。如果您使用将数据复制 DataSync 到启用 TDE 的集群，请先在 HDFS 集群上配置加密区域。DataSync 不会创建加密区域。

## 创建您的 HDFS 传输位置

配置一个可使用 DataSync 传输源的地点。

开始之前：通过执行以下操作验证代理和 Hadoop 集群之间的网络连接：

- 测试对中列出的 TCP 端口的访问权限 [自管理存储系统的网络要求](#)。
- 测试您的本地代理与 Hadoop 集群之间的访问权限。有关说明，请参阅 [测试代理与存储系统的连接](#)。

## 使用控制台创建 HDFS 位置 DataSync

1. [通过 https://console.aws.amazon.com/datasync/](https://console.aws.amazon.com/datasync/) 打开 AWS DataSync 主机。
2. 在左侧导航窗格中，展开“数据传输”，然后选择“位置”和“创建位置”。
3. 对于位置类型，请选择 Hadoop 分布式文件系统 (HDFS)。您可以稍后将此位置配置为源位置或目标。
4. 对于代理，从可用代理列表选择一个或多个要使用的代理。代理连接到您的 HDFS 集群，以便在 HDFS 集群和之间安全地传输数据。DataSync

5. 对于 NameNode，提供 HDFS 群集主NameNode群集的域名或 IP 地址。
6. 在“文件夹”中，输入 HDFS 集群DataSync上用于数据传输的文件夹。将该位置用作任务源时，DataSync将文件复制到提供的文件夹中。将您的位置用作任务的目的地时，DataSync将所有文件写入提供的文件夹。
7. 要设置块大小或复制因子，请选择其他设置。默认块大小为 128 MiB，并且提供的任何块大小必须是 512 字节的倍数。将数据传输到 HDFS 集群DataNodes时，默认复制因子为三。
8. 在安全部分中，选择您的 HDFS 集群上使用的身份验证类型。
  - 简单 — 对于用户，在 HDFS 集群上指定具有以下权限的用户名（取决于您的用例）：
    - 如果您计划使用此位置作为源位置，请指定仅具有读取权限的用户。
    - 如果您打算将此位置用作目标位置，请指定具有读取和写入权限的用户。
  - 或者，指定 HDFS 集群的密钥管理服务器（KMS）的 URI。
  - Kerberos — 指定有权访问您的 HDFS 集群的 Kerberos 主体。接下来，提供包含所提供的 Kerberos 主体的KeyTab文件。然后，提供 Kerberos 配置文件。最后，在 RPC 保护和数据传输保护下拉列表中指定传输保护中的加密类型。
9. （可选）选择添加标签来标记您的 HDFS 位置。

标签是键值对，帮助您管理、筛选和搜索位置。我们建议至少为您的位置创建一个名称标签。

10. 选择创建地点。

## 不支持的 HDFS 功能

HDFS 的以下功能目前不支持：DataSync

- 使用 Kerberos 身份验证时的透明数据加密 (TDE)
- 配置多个 NameNodes
- 基于 HTTP 的 Hadoop HDFS (HTTPFS)
- POSIX 访问控制列表 (ACL)
- HDFS 扩展属性 (xattrs)

## 配置来自对象存储系统的AWS DataSync传输

要从对象存储系统传输数据，必须创建AWS DataSync传输位置。DataSync可以将此位置用作传输数据的源或目的地。

## 先决条件

您的对象存储系统必须与以下 [Amazon S3 API 操作](#) 兼容才能AWS DataSync连接到它：

- AbortMultipartUpload
- CompleteMultipartUpload
- CopyObject
- CreateMultipartUpload
- DeleteObject
- DeleteObjects
- DeleteObjectTagging
- GetBucketLocation
- GetObject
- GetObjectTagging
- HeadBucket
- HeadObject
- ListObjectsV2
- PutObject
- PutObjectTagging
- UploadPart

## 创建您的对象存储传输位置

在开始之前，您需要一个计划从中传输数据的对象存储系统。

使用控制台创建对象存储位置

1. [通过 https://console.aws.amazon.com/datasync/](https://console.aws.amazon.com/datasync/) 打开AWS DataSync主机。
2. 在左侧导航窗格中，展开“数据传输”，然后选择“位置”和“创建位置”。
3. 对于位置类型，选择对象存储。

您可在稍后将此位置配置为源或目标。

4. 对于代理，选择一个或多个DataSync代理。

在传输过程中，代理会安全地连接到您的对象存储服务器。

5. 对于服务器，请提供对象存储服务器的域名或 IP 地址。
6. 在存储桶名称中，输入传输中涉及的对象存储桶的名称。
7. 对于文件夹，输入对象前缀。

DataSync仅复制具有此前缀的对象。

8. 要配置与对象存储服务器的连接，请展开其他设置并执行以下操作：
  - a. 对于服务器协议，选择 HTTP 或 HTTPS。
  - b. 对于服务器端口，请使用默认端口（HTTP 为 80，HTTPS 为 443）或根据需要指定自定义端口。
  - c. 对于证书，选择选择文件以指定用于签署对象存储服务器证书的证书。

该文件最多可以是 32768 字节（在 base64 编码之前），还包括以下内容：

- 签名证书颁发机构 (CA) 的证书
- 任何中间证书
- 一个 .pem 扩展

9. 如果需要凭据才能访问对象存储服务器，请选择需要凭据，然后输入访问密钥和密钥来访问存储桶。

访问密钥和密钥可以分别是用户名和密码。

10. （可选）选择添加标签来标记您的对象存储位置。

标签是键值对，帮助您管理、筛选和搜索位置。我们建议至少为您的位置创建一个名称标签。

11. 选择创建地点。

## 使用以下方式向存储器传输或从AWS存储器传输 AWS DataSync

使用AWS DataSync，您可以将数据传输到多个存储服务或从多种AWS存储服务传输数据。有关更多信息，请参阅[我可以将数据传输到哪里DataSync？](#)。

### 主题

- [使用 Amazon S3 配置AWS DataSync传输](#)
- [使用亚马逊 EFS 配置AWS DataSync传输](#)

- [使用 Amazon FSx for Windows File Server AWS DataSync](#)
- [使用 Amazon FSx for Lustre 配置 AWS DataSync 传输](#)
- [使用 Amazon FSx for OpenZFS 配置 AWS DataSync 传输](#)
- [使用适用 NetApp 于 ON AWS DataSync TAP 的亚马逊 FSx 配置传输](#)

## 使用 Amazon S3 配置 AWS DataSync 传输

要向 S3 存储桶传输数据或从您的 S3 存储桶传输数据，您必须创建 AWS DataSync 传输位置。DataSync 可以将此位置用作传输数据的源或目的地。

### Important

在创建位置之前，请确保阅读以下部分：

- [Amazon S3 传输的存储类注意事项](#)
- [使用时评估 S3 请求成本 DataSync](#)

## 访问 S3 存储桶

DataSync 需要访问 Amazon S3 桶。为此，请扮演 DataSync 具有 IAM 策略和 AWS Identity and Access Management (IAM) 信任关系的 AWS Security Token Service (IAMAWS STS) 角色。该策略决定角色可以执行哪些操作。

DataSync 可以为您创建此角色，但在某些情况下您可能需要手动创建角色。有关更多信息，请参阅 [使用 IAM 策略访问您的 S3 存储桶](#)：

## Amazon S3 传输的存储类注意事项

DataSync 可以将对象直接传输到您在创建 [Amazon S3 位置时指定的 Amazon S3 存储类](#) 中。有些存储类别的行为会影响您的 Amazon S3 存储成本。有关更多信息，请参阅 [Amazon S3 定价](#)。

### Important

复制到 S3 存储桶的新对象使用您在创建 Amazon S3 位置时指定的存储类进行存储。DataSync 不会更改存储桶中现有对象的存储类别（即使该对象是在源位置修改的）。

Amazon S3 存储类	注意事项
S3 Standard	选择 S3 Standard 以冗余方式将您经常访问的文件存储在地理上分隔的多个可用区中。如果您没有指定存储类别，则这是默认设置。
S3 Intelligent-Tiering	<p>选择 S3 Intelligent-Tiering (S3 智能分层)，可通过自动将数据移动到最具成本效益的存储访问层来优化存储成本。</p> <p>对于存储在 S3 Intelligent-Tiering 存储类中的对象，您需要按月付费。这项 Amazon S3 费用包括监控数据访问模式和在层间移动对象。</p>
S3 Standard-IA	<p>选择 S3 Standard-IA，将您不经常访问的对象以冗余方式存储在地理上分隔的多个可用区中。</p> <p>存储在 S3 Standard-IA 存储类中的对象可能会因覆盖、删除或检索而产生额外费用。考虑这些对象的更改频率，计划保留这些对象的时间以及需要访问的频率。对对象数据或元数据的更改等同于删除一个对象并创建一个新对象来替换它。这会导致存储在 S3 Standard-IA 存储类中的对象产生额外费用。</p> <p>在 S3 Standard-IA 存储类中，小于 128 KB 的对象小于每个对象的最低容量费用。这些对象存储在 S3 标准存储类中。</p>
S3 One Zone-IA	<p>选择 S3 单区 — IA 将您不常访问的对象存储在单个可用区中。</p> <p>存储在 S3 One Zone-IA 存储类中的对象可能会因覆盖、删除或检索而产生额外费用。考虑这些对象的更改频率，计划保留这些对象的时间以及需要访问的频率。对对象数据或元数据的更改等同于删除一个对象并创建一个新对象来替换它。这会导致存储在 S3 单区 — IA 存储类中的对象产生额外费用。</p> <p>在 S3 单区 — IA 存储类中，小于 128 KB 的对象小于每个对象的最低容量费用。这些对象存储在 S3 标准存储类中。</p>
S3 Glacier 即时检索	<p>选择 S3 Glacier 即时检索可存档很少访问但需要在毫秒内检索的对象。</p> <p>与 S3 Standard-IA 存储类别相比，存储在 S3 Glacier Instant Retrieval 存储类中的数据可以节省成本。但是 S3 Glacier 即时检索的数据访问成本高于 S3 Standard-IA。</p>

Amazon S3 存储类	注意事项
	<p>存储在 S3 Glacier 即时检索中的对象可能会因覆盖、删除或检索而产生额外费用。考虑这些对象的更改频率，计划保留这些对象的时间以及需要访问的频率。对对象数据或元数据的更改等同于删除一个对象并创建一个新对象来替换它。这会导致存储在 S3 Glacier 即时检索存储类中的对象产生额外费用。</p> <p>小于 128KB 的对象小于 S3 Glacier Instant Retrieval 存储类中每个对象的最低容量费用。这些对象存储在 S3 标准存储类中。</p>
S3 Glacier Flexible Retrieval	<p>选择 S3 Glacier Flexible Retrieval</p> <p>存储在 S3 Glacier Flexible Retrieval 中的对象会因覆盖、删除或检索而产生额外费用。考虑这些对象的更改频率，计划保留这些对象的时间以及需要访问的频率。对对象数据或元数据的更改等同于删除一个对象并创建一个新对象来替换它。这会导致存储在 S3 Glacier Flexible Retrieval 存储类中的对象产生额外费用。</p> <p>小于 40 KB 的对象小于 S3 Glacier Flexible Retrieval 存储类中每个对象的最低容量费用。这些对象存储在 S3 标准存储类中。</p> <p>必须先恢复在此存储类中存档的对象，然后DataSync才能读取它们。有关信息，请参阅 Amazon S3 用户指南中的<a href="#">处理存档对象</a>。</p> <p>使用 S3 Glacier 灵活检索时，选择“仅验证传输的数据”任务选项，以便在传输结束时比较数据和元数据校验和。您不能使用此存储类的“验证目标中的所有数据”选项，因为它需要从目标检索所有现有对象。</p>

Amazon S3 存储类	注意事项
S3 Glacier Deep Archive	<p>选择 S3 Glacier Deep Archive 来存档您的对象，以实现长期数据保留和数字保存，每年访问一次或两次数据。</p> <p>存储在 S3 Glacier Deep Archive 中的对象可能会因覆盖、删除或检索而产生额外费用。考虑这些对象的更改频率，计划保留这些对象的时间以及需要访问的频率。对对象数据或元数据的更改等同于删除一个对象并创建一个新对象来替换它。这会导致存储在 S3 Glacier Deep Archive 存储类中的对象产生额外费用。</p> <p>小于 40 KB 的对象小于 S3 Glacier Deep Archive 存储类中每个对象的最低容量费用。这些对象存储在 S3 标准存储类中。</p> <p>必须先恢复在此存储类中存档的对象，然后 DataSync 才能读取它们。有关信息，请参阅 Amazon S3 用户指南中的<a href="#">处理存档对象</a>。</p> <p>使用 S3 Glacier Deep Archive 时，选择“仅验证传输的数据”任务选项，以便在传输结束时比较数据和元数据校验和。您不能使用此存储类的“验证目标中的所有数据”选项，因为它需要从目标检索所有现有对象。</p>
S3 Outposts 基地	Amazon S3 on Outposts 的存储类。

## 使用时评估 S3 请求成本 DataSync

使用 Amazon S3 站点时，您会产生与发出 S3 API 请求相关的费用 DataSync。本部分可以帮助您了解这些请求的 DataSync 使用方式以及它们可能如何影响您的 [Amazon S3 成本](#)。

### 主题

- [由 S3 发出的 S3 请求 DataSync](#)
- [与成本相关的注意事项](#)

### 由 S3 发出的 S3 请求 DataSync

下表描述了当您将数据复制到 Amazon S3 位置或从 Amazon S3 位置复制数据时 DataSync 可能发出的 S3 请求。



S3 请求	怎么DataSync用
<a href="#">ListObjectV2</a>	DataSync对每个以正斜杠 (/) 结尾的对象发出至少一个LIST请求，以列出以该前缀开头的对象。此请求是在任务的 <a href="#">准备</a> 阶段调用的。
<a href="#">HeadObject</a>	DataSync在任务的 <a href="#">准备</a> 和 <a href="#">验证</a> 阶段发出检索对象元数据的HEAD请求。每个对象可能有多个HEAD请求，具体取决于您DataSync要如何 <a href="#">验证其传输的数据的完整性</a> 。
<a href="#">GetObject</a>	DataSync在任务的 <a href="#">传输</a> 阶段发出从对象读取数据的GET请求。可以有多个对大型对象的GET请求。
<a href="#">PutObject</a>	DataSync在任务的 <a href="#">传输</a> 阶段发出在目标 S3 存储桶中创建对象的PUT请求。由于DataSync使用 <a href="#">Amazon S3 分段上传功能</a> ，因此可能会有多个对大型对象的PUT请求。
<a href="#">CopyObject</a>	DataSync只有在对象的元数据发生变化时，才会COPY请求创建该对象的副本。如果您最初使用其他未传送其元数据的服务或工具将数据复制到 S3 存储桶，则可能会发生这种情况。

## 与成本相关的注意事项

DataSync每次运行任务时都会在 S3 存储桶上发出 S3 请求。在某些情况下，这可能会导致费用增加。例如：

- 您经常将对象从 S3 存储桶转移或从 S3 存储桶传出。
- 您可能没有传输太多数据，但是您的 S3 存储桶中有很多对象。在这种情况下，您仍然会看到高额费用，因为DataSync会对存储桶的每个对象发出 S3 请求。
- 您在 S3 存储桶之间传输，在源和目标上发出 S3 请求DataSync也是如此。

为了帮助最大限度地减少与相关的 S3 请求成本DataSync，请考虑以下几点：

## 主题

- [我使用的是哪些 S3 存储类？](#)
- [我需要多久传输一次数据？](#)

### 我使用的是哪些 S3 存储类？

S3 请求费用可能因对象使用的 Amazon S3 存储类别而异，尤其是归档对象的类（例如 S3 Glacier Instant Retrieval、S3 Glacier Flexible Retrieval 和 S3 Glacier Deep Archive）

以下是存储类在使用时会影响 S3 请求费用的一些场景DataSync：

- 每次运行任务时，都会DataSync发出检索对象元数据的HEAD请求。即使你没有移动任何物体，这些请求也会产生费用。这些请求对账单的影响程度取决于您的对象使用的存储类以及DataSync扫描的对象数量。
- 如果您将对象移至 S3 Glacier 即时检索存储类（直接或通过存储分段生命周期配置），则对该类中对象的请求比其他存储类中的对象更昂贵。
- 如果您将DataSync任务配置为[验证源位置和目标位置是否完全同步](#)，则所有存储类别（S3 Glacier 灵活检索和 S3 Glacier Deep Archive 除外）中的每个对象都会GET收到请求。
- 除GET请求外，您还会为 S3 Standard-IA、S3 One Zone-IA 或 S3 Glacier Instant Retrieval 存储类中的对象产生数据检索费用。

有关更多信息，请参阅 [Amazon S3 定价](#)。

### 我需要多久传输一次数据？

如果您需要定期移动数据，请考虑运行的任务不会超过您的需求[的时间表](#)。

您也可以考虑限制您的转账范围。例如，您可以配置DataSync为专注于某些前缀中的对象或[筛选传输的数据](#)。这些选项可以帮助减少每次运行DataSync任务时发出的 S3 请求数量。

## 使用 Amazon S3 传输的其他注意事项

将 Amazon S3 与配合使用时DataSync，请记住以下几点：

- 对对象数据或元数据的更改等同于删除和替换对象。在以下情况下，这些更改会导致额外费用：
  - 使用对象版本控制时-对对象数据或元数据的更改会创建对象的新版本。
  - 当使用的存储类可能会因覆盖、删除或检索对象而产生额外费用时，更改对象数据或元数据会产生此类费用。有关更多信息，请参阅[Amazon S3 传输的存储类注意事项](#)：

- 在 Amazon S3 中使用对象版本控制时，运行一次 DataSync 任务可能会创建 Amazon S3 对象的多个版本。
- DataSync 如果对象的名称中有非标准字符，则可能不会传输该对象。有关更多信息，请参阅 Amazon S3 用户指南中的 [对象密钥命名指南](#)。
- 为了帮助最大限度地降低您的 Amazon S3 存储成本，我们建议使用生命周期配置来停止未完成的分段上传。有关更多信息，请参阅 [Amazon S3 用户指南](#)。
- 最初将数据从 S3 存储桶传输到文件系统（例如，NFS 或 Amazon FsX）后，后续运行的同一 DataSync 任务将不包括已修改但大小与第一次传输时相同的对象。

## 创建您的 Amazon S3 传输地点

要创建位置，您需要现有 S3 桶。如果您没有，请参阅 [Amazon S3 用户指南中的 Amazon S3 入门](#)。

### Tip

如果您的 S3 存储桶包含具有不同存储类别的对象，请了解 [这些存储类别是如何 DataSync 使用的](#)，以及它 [会如何影响您的 AWS 账单](#)。

## 创建 Amazon S3 位置

1. 通过 <https://console.aws.amazon.com/datasync/> 打开 AWS DataSync 主机。
2. 在左侧导航窗格中，展开“数据传输”，然后选择“位置”和“创建位置”。
3. 对于位置类型，选择 Amazon S3。
4. 对于 S3 桶，请选择要用作位置的桶。（稍后创建 DataSync 任务时，您可以指定此位置是源位置还是目标位置。）

如果您的 S3 存储桶位于 AWS Outposts 资源上，则必须指定 Amazon S3 接入点。有关更多信息，请参阅 [Amazon S3 用户指南中的使用 Amazon S3 访问点管理数据访问](#)。


5. 对于 S3 存储类别，请选择您希望对象使用的存储类别。

有关更多信息，请参阅 [Amazon S3 传输的存储类注意事项](#)：DataSync 默认情况下，Outposts 在 Amazon S3 中使用 S3 Outposts 存储类。

6. （仅限 Outposts 站上的 Amazon S3）对于代理，请指定 Outposts DataSync 上代理的 Amazon 资源名称 (ARN)。

有关更多信息，请参阅 [在上部署您的代理 AWS Outposts](#)：

- 对于文件夹，请在 S3 存储桶中输入DataSync读取或写入内容的前缀（具体视存储桶是源位置还是目标位置而定）。

 Note

前缀不能以斜杠（例如/photos）开头，也不能包含连续的斜杠，例如photos//2006/January。

- 对于 IAM role (IAM 角色)，执行以下操作之一：
  - 选择“自动生成”DataSync 以自动创建具有访问 S3 存储桶所需权限的 IM 角色。

如果DataSync之前为此 S3 存储桶创建了 IAM 角色，则默认情况下会选择该角色。
  - 选择您创建的自定义 IAM 角色。有关更多信息，请参阅[手动创建 IAM 角色以访问 Amazon S3 桶](#)：
- （可选）选择添加标签来标记您的 Amazon S3 位置。

标签 是帮助您管理、筛选和搜索位置的键值对。
- 选择创建地点。

## 使用 IAM 策略访问您的 S3 存储桶

根据您的 S3 存储桶的安全设置，您可能需要创建允许DataSync访问该存储桶的自定义 IAM 策略。

### 主题

- [手动创建 IAM 角色以访问 Amazon S3 桶](#)
- [防范跨服务混淆代理问题](#)
- [使用服务器端加密访问 S3 桶](#)

### 手动创建 IAM 角色以访问 Amazon S3 桶

虽然DataSync您可以使用所需的 S3 存储桶权限为您创建 IAM 角色，但您也可以自己配置角色。

### 手动创建 IAM 角色以访问 Amazon S3 桶

- 通过以下网址打开 IAM 控制台：<https://console.aws.amazon.com/iam/>。
- 在左侧导航窗格的“访问管理”下，选择“角色”，然后选择“创建角色”。

3. 在“选择可信实体”页面上，为“可信实体类型”选择AWS 服务。
4. 对于用例，DataSync在下拉列表中选择并选择 DataSync-S3 位置。选择下一步。
5. 在添加权限页面上，FullAccess为中的 S3 存储桶选择 Amazon S3。AWS 区域选择下一步。

您可以手动创建比 AmazonS3 FullAccess 更严格的政策。示例如下：

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Action": [
        "s3:GetBucketLocation",
        "s3:ListBucket",
        "s3:ListBucketMultipartUploads"
      ],
      "Effect": "Allow",
      "Resource": "YourS3BucketArn"
    },
    {
      "Action": [
        "s3:AbortMultipartUpload",
        "s3:DeleteObject",
        "s3:GetObject",
        "s3:ListMultipartUploadParts",
        "s3:GetObjectTagging",
        "s3:PutObjectTagging",
        "s3:PutObject"
      ],
      "Effect": "Allow",
      "Resource": "YourS3BucketArn/*"
    }
  ]
}
```

对于 Amazon S3 on Outposts，使用以下策略：

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Action": [
```

```

        "s3-outposts:ListBucket",
        "s3-outposts:ListBucketMultipartUploads"
    ],
    "Effect": "Allow",
    "Resource": [
        "s3outpostsBucketArn",
        "s3outpostsAccessPointArn"
    ],
    "Condition": {
        "StringLike": {
            "s3-outposts:DataAccessPointArn": "s3outpostsAccessPointArn"
        }
    }
},
{
    "Action": [
        "s3-outposts:AbortMultipartUpload",
        "s3-outposts>DeleteObject",
        "s3-outposts:GetObject",
        "s3-outposts:ListMultipartUploadParts",
        "s3-outposts:GetObjectTagging",
        "s3-outposts:PutObjectTagging"
    ],
    "Effect": "Allow",
    "Resource": [
        "s3outpostsBucketArn/*",
        "s3outpostsAccessPointArn"
    ],
    "Condition": {
        "StringLike": {
            "s3-outposts:DataAccessPointArn": "s3outpostsAccessPointArn"
        }
    }
},
{
    "Effect": "Allow",
    "Action": [
        "s3-outposts:GetAccessPoint"
    ],
    "Resource": "s3outpostsAccessPointArn"
}
]
}

```

6. 为您的角色命名，然后选择创建角色。
7. 通过 <https://console.aws.amazon.com/datasync/> 打开AWS DataSync主机。
8. 选择 IAM 角色设置旁边的刷新按钮，然后选择您刚刚创建的角色。

## 防范跨服务混淆代理问题

为防止[跨服务混淆副手问题](#)，我们建议在您的 IAM 角色的信任策略中使用aws:SourceArn和aws:SourceAccount全局条件上下文密钥。

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Principal": {
        "Service": "datasync.amazonaws.com"
      },
      "Action": "sts:AssumeRole",
      "Condition": {
        "StringEquals": {
          "aws:SourceAccount": "123456789012"
        },
        "StringLike": {
          "aws:SourceArn": "arn:aws:datasync:us-east-2:123456789012:*"
        }
      }
    }
  ]
}
```

## 使用服务器端加密访问 S3 桶

DataSync可以将数据复制到[使用服务器端加密的 S3 存储桶](#)或从中复制数据。存储桶使用的加密密钥类型可以决定您是否需要DataSync允许访问存储桶的自定义策略。

DataSync与使用服务器端加密的 S3 存储桶一起使用时，请记住以下几点：

- 如果您的 S3 存储桶使用AWS托管密钥加密，则在所有资源都位于相同的情况下，默认情况下DataSync可以访问该存储桶的对象AWS 账户。

- 如果您的 S3 存储桶使用客户管理 AWS Key Management Service (AWS KMS) 密钥 (SSE-KMS) 进行加密，则[密钥的策略](#)必须包含DataSync用于访问存储桶的 IAM 角色。
- 如果您的 S3 存储桶是使用客户管理的 SSE-KMS 密钥加密的，则DataSync需要权限才能访问另一个AWS 账户存储段中的存储桶。AWS 账户您可以通过执行以下操作进行设置：
  - 在使用的 IAM 角色中DataSync，[指定与目标存储桶关联的 SSE-KMS 密钥](#)。
  - 在 SSE-KMS 密钥策略中，[指定使用的 IAM 角色](#)。DataSync
- 如果您的 S3 存储桶已使用客户提供的加密密钥 (SSE-C) 进行加密，则DataSync无法访问此存储桶。

示例：对的 SE-KMS 密钥策略 DataSync

以下示例是客户管理的 SSE-KMS [密钥的密钥策略](#)。该策略与使用服务器端加密的 S3 存储桶关联。以下值特定于您的设置：

- `####-#`的 AWS 账户
- `your-admin-role`— 可以管理密钥的 IAM 角色。
- `your-datasync-role`— 允许在访问存储桶时DataSync使用密钥的 IAM 角色。

```
{
  "Id": "key-consolepolicy-3",
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "Enable IAM Permissions",
      "Effect": "Allow",
      "Principal": {
        "AWS": "arn:aws:iam::your-account:root"
      },
      "Action": "kms:*",
      "Resource": "*"
    },
    {
      "Sid": "Allow access for Key Administrators",
      "Effect": "Allow",
      "Principal": {
        "AWS": "arn:aws:iam::your-account:role/your-admin-role"
      },
      "Action": [
```



```

        "kms:Create*",
        "kms:Describe*",
        "kms:Enable*",
        "kms:List*",
        "kms:Put*",
        "kms:Update*",
        "kms:Revoke*",
        "kms:Disable*",
        "kms:Get*",
        "kms>Delete*",
        "kms:TagResource",
        "kms:UntagResource",
        "kms:ScheduleKeyDeletion",
        "kms:CancelKeyDeletion"
    ],
    "Resource": "*"
},
{
    "Sid": "Allow use of the key",
    "Effect": "Allow",
    "Principal": {
        "AWS": "arn:aws:iam::your-account:role/your-datasync-role"
    },
    "Action": [
        "kms:Encrypt",
        "kms:Decrypt",
        "kms:ReEncrypt*",
        "kms:GenerateDataKey*"
    ],
    "Resource": "*"
},
{
    "Sid": "Allow attachment of persistent resources",
    "Effect": "Allow",
    "Principal": {
        "AWS": "arn:aws:iam::your-account:role/your-datasync-role"
    },
    "Action": [
        "kms:CreateGrant",
        "kms:ListGrants",
        "kms:RevokeGrant"
    ],
    "Resource": "*",
    "Condition": {

```

```
        "Bool": {
            "kms:GrantIsForAWSResource": "true"
        }
    }
}
```

## 在不同的 S3 存储桶之间转移或从中转出 AWS 账户

借助 DataSync，您可以将数据移入或移出[不同](#)的 S3 存储桶 AWS 账户。有关更多信息，请参阅以下教程：

- [将数据从本地存储传输到另一个 Amazon S3 AWS 账户](#)
- [将数据从 Amazon S3 传输到 Amazon S3 AWS 账户](#)

## 使用亚马逊 EFS 配置 AWS DataSync 传输

要向您的 Amazon EFS 文件系统传输数据或从中传输数据，您必须创建 AWS DataSync 传输位置。DataSync 可以将此位置用作传输数据的源或目的地。

### 访问 Amazon EFS 文件系统

DataSync 使用[网络接口](#)以根用户身份从虚拟私有云 (VPC) 装载您的 Amazon EFS 文件系统。

创建位置时，您可以让 DataSync 使用网络文件系统 (NFS) 端口 2049 指定用于连接您的 Amazon EFS 文件系统的装载目标或[接入点](#)的子网和安全组。

DataSync 还可以装载配置为受限访问的 Amazon EFS 文件系统。例如，您可以指定一个 AWS Identity and Access Management (IAM) 角色，DataSync 为连接到您的文件系统提供必要的权限级别。有关更多信息，请参阅[使用 IAM 策略访问 Amazon EFS 文件系统](#)：

### 亚马逊 EFS 地点的注意事项

创建 Amazon EFS 文件系统的 DataSync 位置时，请考虑以下几点：

- 与您一起使用的 VPC DataSync 必须具有默认租期。不支持具有专用租户的 VPC。有关更多信息，请参阅[使用 VPC](#)。

- 在突增吞吐量模式下创建 Amazon EFS 文件系统时，您将获得价值 2.1 TB 的突增积分。在突增吞吐量模式下，所有 Amazon EFS 文件系统的吞吐量均可突增至每秒 100 MB。当有突发积分可用时，Amazon S3 标准类存储容量超过 1 TiB 的文件系统每 TB 每秒可以驱动 100 MiB。

DataSync 消耗文件系统突增积分。这可能会影响应用程序性能。在使用 DataSync 具有活动工作负载的文件系统时，请考虑使用 Amazon EFS 弹性吞吐量或预置吞吐量模式。

- 处于通用性能模式的 Amazon EFS 文件系统的限制为每秒 35,000 次文件系统操作。该限制可能会影响 DataSync 在复制文件时达到的最大吞吐量。

读取数据或元数据的操作消耗一个文件操作。写入数据或更新元数据的操作消耗五个文件操作。这意味着文件系统可以支持每秒 35,000 次读取操作、7,000 次写入操作或两者的某种组合。从所有连接的客户端统计文件操作数。

有关更多信息，请参阅 [Amazon EFS file System User system Guide](#) 中的

## 创建您的亚马逊 EFS 传输地点

要创建位置，您需要现有的 Amazon EFS 文件系统。如果您没有，请参阅 [Amazon Elastic File System User system Guide](#) Amazon Elastic File System id e System

### 创建 Amazon EFS 位置

1. 通过 <https://console.aws.amazon.com/datasync/> 打开 AWS DataSync 主机。
2. 在左侧导航窗格中，展开“数据传输”，然后选择“位置”和“创建位置”。
3. 对于位置类型，选择亚马逊 EFS 文件系统。

您可在稍后将此位置配置为源或目标。

4. 对于 File system ( 文件系统 ) ，请选择要用作位置的 Amazon EFS 文件系统。

您可在稍后将此位置配置为源或目标。

5. 对于“挂载路径”，输入 Amazon EFS 文件系统的挂载路径。

这可指定 DataSync 读取或写入数据的位置 ( 具体视源位置还是目标位置而定 ) 。

默认情况下，DataSync 使用根目录 ( 如果您配置了根目录，则使用接入点 ) 。您也可以使用正斜杠指定子目录 ( 例如，/path/to/directory ) 。

6. 对于子网，请选择一个子网，该子网在其中 DataSync 创建网络接口用于管理传输过程中的流量。

子网必须位于：

- 与Amazon EFS 文件系统位于同一 VPC 中。
- 至少与一个文件系统挂载目标位于同一个可用区。

**Note**

您无需指定包含文件系统挂载目标的子网。

7. 对于安全组，请选择与 Amazon EFS 文件系统的挂载目标关联的安全组。

**Note**

您指定的安全组必须允许 NFS 端口 2049 上的入站流量。有关更多信息，请参阅 [Amazon Elastic File System Guide](#) 中的 [将 VPC 安全组用于 Amazon EC 2 实例和挂载目标](#)。

8. 对于传输中加密，请选择将数据复制DataSync到文件系统或从文件系统复制数据时，您是否要使用传输层安全性协议 ( TLS ) 加密。

**Note**

如果您想使用您的位置配置接入点、IAM 角色或两者，则必须启用此设置。

9. ( 可选 ) 对于 EFS 接入点，请选择DataSync可用于装载 Amazon EFS 文件系统的接入点。
10. ( 可选 ) 对于 IAM 角色，请指定DataSync允许访问您的文件系统的角色。

有关创建此角色的信息，请参见 [使用 IAM 策略访问 Amazon EFS 文件系统](#)

11. ( 可选 ) 选择添加标签来标记您的文件系统。

标签 是帮助您管理、筛选和搜索位置的键值对。

12. 选择创建地点。

## 使用 IAM 策略访问 Amazon EFS 文件系统

您可以使用 IAM 策略为您的 Amazon EFS 文件系统配置更高的安全级别。在您的[文件系统策略](#)中，您可以指定仍DataSync允许与文件系统连接的 IAM 角色。

**Note**

要使用 IAM 角色，在为文件系统创建DataSync位置时必须启用 TLS 进行传输中加密。

有关更多信息，请参阅 Amazon Elastic File System Guide 中的使用 IAM 控制文件系统[数据访问](#)。

为创建 IAM 角色 DataSync

创建以可信实体DataSync为角色的角色。

创建 IAM 角色

1. 通过以下网址打开 IAM 控制台：<https://console.aws.amazon.com/iam/>。
2. 在左侧导航窗格的“访问管理”下，选择“角色”，然后选择“创建角色”。
3. 在选择可信实体页面上，对于“可信实体类型”，选择“自定义信任策略”。
4. 将下面的 JSON 粘贴到策略编辑器中：

```
{
  "Version": "2012-10-17",
  "Statement": [{
    "Effect": "Allow",
    "Principal": {
      "Service": "datasync.amazonaws.com"
    },
    "Action": "sts:AssumeRole"
  }]
}
```

5. 选择 Next ( 下一步 )。在 Add permissions ( 添加权限 ) 页面上，选择 Next ( 下一步 )。
6. 为您的角色命名，然后选择创建角色。

在创建 Amazon EFS 文件系统的位置时，请指定该角色。

Amazon EFS 文件系统策略示例

以下 IAM 策略示例包含有助于限制对 Amazon EFS 文件系统的访问的元素（在策略中标识为 `fs-1234567890abcdef0`）：

- **Principal**：指定一个DataSync授予连接到文件系统的权限的 IAM 角色。

- Action : 提供 DataSync root 访问权限并允许其读取和写入文件系统。
- aws:SecureTransport : 要求 NFS 客户端在连接到文件系统时使用 TLS。
- elasticfilesystem:AccessPointArn : 仅允许通过特定的接入点访问文件系统。

```
{
  "Version": "2012-10-17",
  "Id": "ExampleEFSFileSystemPolicy",
  "Statement": [{
    "Sid": "AccessEFSFileSystem",
    "Effect": "Allow",
    "Principal": {
      "AWS": "arn:aws:iam::111122223333:role/MyDataSyncRole"
    },
    "Action": [
      "elasticfilesystem:ClientMount",
      "elasticfilesystem:ClientWrite",
      "elasticfilesystem:ClientRootAccess"
    ],
    "Resource": "arn:aws:elasticfilesystem:us-east-1:111122223333:file-system/fs-1234567890abcdef0",
    "Condition": {
      "Bool": {
        "aws:SecureTransport": "true"
      },
      "StringEquals": {
        "elasticfilesystem:AccessPointArn": "arn:aws:elasticfilesystem:us-east-1:111122223333:access-point/fsap-abcdef01234567890"
      }
    }
  }]
}
```

## 使用Amazon FSx for Windows File Server AWS DataSync

要将数据传输到 Amazon FSx for Windows File Server 系统或从文件系统传输数据，您必须创建AWS DataSync传输位置。DataSync可以将此位置用作传输数据的源或目的地。

### 访问 FSx for Windows File Server 系统

DataSync使用服务器消息块 (SMB) 协议连接到 FSx for Windows 文件服务器，并使用[网络](#)接口从虚拟私有云 (VPC) 装载文件系统。

**Note**

与您一起使用的 VPC DataSync 必须具有默认租期。不支持具有专用租户的 VPC。有关更多信息，请参阅[使用 VPC](#)。

**主题**

- [所需的权限](#)
- [对象所有权和 NTFS ACL 权限](#)

**所需的权限**

DataSync需要一个对您的 FSx for Windows 文件服务器文件、文件夹和文件元数据具有足够权限的用户帐户。

我们建议您将此用户设为文件系统管理员组的成员。如果您将 File Server AWS Directory Service for Microsoft Active Directory 与 File Server 结合使用，则用户必须是AWS委托 FSx 组的成员。如果您将自行管理的 Active Directory 与 File Server File Server 结合使用，则用户必须是以下两个组中的一个组的成员：

- 域管理员组。
- 您在创建文件系统时为文件系统管理指定的自定义组。

在 SMB 文件服务器和 Windows 文件服务器文件系统的 FsX 之间传输时，或者在 Windows 文件服务器文件系统的 FsX 之间传输时，传输源和目标必须：

- 属于同一个Active Directory 域。
- 在他们的域之间建立 Active Directory 信任关系。

**对象所有权和 NTFS ACL 权限**

要设置对象所有权，DataSync需要具有权限的用户，该SE\_RESTORE\_NAME权限通常授予内置 Active Directory 组 Backup 操作员和域管理员的成员。DataSync为用户提供了此权限还有助于确保除了 NTFS 系统访问控制列表 (SACL) 之外的文件、文件夹和文件元数据的足够权限。

复制 SACL 需要其他权限，特别是授予域管理员组成员的 Windows SE\_SECURITY\_NAME 权限。DataSync 要配置 DataSync 复制 ACL 的方式，请参阅 [管理文件、对象和元数据的 AWS DataSync 传输方式](#)。

#### Warning

复制 NTFS ACL 时，请确保 SYSTEM 用户对您的源位置和目标位置中的所有文件夹具有完全控制权限。如果不这样做，DataSync 可以更改目标的权限，使您的 FSx for Windows File Server 共享无法访问。有关更多信息，请参阅 [Amazon FSx for Windows File Server 用户指南](#)。

## 创建 FSx for Windows File Server 传输位置

在您开始之前，请确保您已经有 FSx for Windows File Server。AWS 区域有关更多信息，请参阅 [Amazon FSx Server User Guide](#) 中的 Amazon File Server

使用控制台创建 Amazon FSx for Windows File Server 位置 DataSync

1. 通过 <https://console.aws.amazon.com/datasync/> 打开 AWS DataSync 主机。
2. 在左侧导航窗格中，展开“数据传输”，然后选择“位置”和“创建位置”。
3. 对于位置类型，选择亚马逊 FsX。
4. 对于 FSx 文件系统，请选择要用作位置的 FSx for Windows File Server 文件系统。
5. 在共享名称中，使用正斜杠输入 FSx for Windows File Server 的装载路径。

这指定 DataSync 读取或写入数据的路径（具体视源位置还是目标位置而定）。

您也可以包含子目录（例如 /path/to/directory）。

6. 对于安全组，最多选择五个安全组，以提供对文件系统的首选子网的访问权限。

#### Note

如果您选择不允许从内部连接的安全组，请执行以下操作之一：

- 将安全组配置为允许其内部通信。
- 选择可以与装载目标的安全组通信的其他安全组。

7. 对于用户，输入可以访问您的 FSx for Windows File Server 用户名。



有关更多信息，请参阅[访问 FSx for Windows File Server 系统](#)：

8. 在“密码”中，输入用户名的密码。
9. （可选）对于域，请输入 File Server 所属的 Windows 域的名称。
10. （可选）在“密钥”和“值”字段中输入值以标记 FSx for Windows File Server。

标签可帮助您管理、筛选和搜索AWS资源。我们建议至少为您的位置创建一个名称标签。

11. 选择创建地点。

## 使用 Amazon FSx for Lustre 配置AWS DataSync传

要将数据传输到 Amazon FSx for Lustre 文件系统或从文件系统传输数据，您必须创建AWS DataSync 传输位置。DataSync可以将此位置用作传输数据的源或目的地。

### 访问 FSx for Lustre 文件系统

DataSync使用 Lustre 客户端访问您的 FSx for Lustre 文件系统。DataSync需要访问您的 FSx for Lustre 文件系统上的所有数据。要获得此访问级别，请使用的用户 ID (UID) 和组 ID (GID) 以 root 用户身份DataSync挂载文件系统。0

DataSync使用[网络接口](#)从 Virtual Private Cloud ( VPC ) 挂载文件系统。DataSync代表您全面管理这些网络接口的创建、使用和删除。

#### Note

与您一起使用的 VPC DataSync 必须具有默认租期。不支持具有专用租户的 VPC。有关更多信息，请参阅[使用 VPC](#)。

### 创建您的 FsX for Lustre 传输地点

如果您还没有 FSx for LustreAWS 区域，请创建一个。有关更多信息，请参阅 [Amazon FSx for Lustre Ustre Ustre 用户指南](#)中的 Amazon FSx for Lu stre。

使用控制台为 Lustre 位置创建 FSx DataSync

1. [通过 https://console.aws.amazon.com/datasync/](https://console.aws.amazon.com/datasync/) 打开AWS DataSync主机。
2. 在左侧导航窗格中，展开“数据传输”，然后选择“位置”和“创建位置”。

### 3. 对于位置类型，选择亚马逊 FsX。

您可在稍后将此位置配置为源或目标。

### 4. 对于 FsX 文件系统，选择要用作位置的 FsX for Lustre 文件系统。

### 5. 在装载路径中，输入 FsX for Lustre 文件系统的装载路径。

该路径可包含子目录。将该位置用作源时，从装载路径DataSync读取数据。将该位置用作目标时，DataSync将所有数据写入装载路径。如果未提供子目录，则DataSync使用根目录 (/)。

### 6. 对于安全组，最多选择五个安全组，以提供对您的 FsX for Lustre 文件系统的访问权限。

安全组必须能够访问文件系统的端口。此外，文件系统必须允许来自安全组的访问。

有关安全组的更多信息，请参阅 Amazon FSx for Lustre 用户指南中的使用Amazon [VPC 的文件系统访问控制](#)。

### 7. (可选) 在“密钥”和“值”字段中输入值，以标记 FsX for Lustre 文件系统。

标签可帮助您管理、筛选和搜索AWS资源。我们建议至少为您的位置创建一个名称标签。

### 8. 选择创建地点。

## 使用 Amazon FSx for OpenZFS 配置AWS DataSync传输

要将数据传输到您的 Amazon FSx for OpenZFS 文件系统或从中传输数据，您必须创建一个AWS DataSync传输位置。DataSync可以将此位置用作传输数据的源或目的地。

### 访问 FSx for OpenZFS 文件系统

DataSync[使用网络接口从虚拟私有云 \(VPC\) 装载 FSx for OpenZFS 文件系统](#)。DataSync代表您全面管理这些网络接口的创建、使用和删除。

#### Note

与您一起使用的 VPC DataSync 必须具有默认租期。不支持具有专用租户的 VPC。有关更多信息，请参阅[使用 VPC](#)。

### 为 OpenZFS 文件系统授权配置 FSx fs 文件系统授权

DataSync作为 NFS 客户端访问您的 FSx for OpenZFS 文件系统，以 root 用户身份装载文件系统，用户 ID (UID) 和组 ID (GID) 为 0。

DataSync要复制所有文件元数据，必须使用no\_root\_squash在文件系统卷上配置 NFS 导出设置。但是，您可以将此访问级别限制为仅限特定DataSync任务。

有关更多信息，请参阅《Amazon FSx for OpenZFS 用户指南》中的[卷属性](#)。

### 配置特定于DataSync (推荐) 的 NFS 导出

您可以针对只能由您的DataSync任务访问的每个卷配置 NFS 导出。对您在创建 FSx for OpenZFS 位置时指定的装载路径的最新祖先卷执行此操作。

### 配置特定于 NFS 导出 DataSync

#### 1. 创建您的[DataSync任务](#)。

这将创建任务的弹性网络接口，您将在 NFS 导出设置中指定这些接口。

2. 使用 Amazon EC2 控制台或，找到任务网络接口的私有 IP 地址AWS CLI。
3. 对于您的 FSx for OpenZFS 文件系统卷，为任务的每个网络接口配置以下 NFS 导出设置：

- 客户机地址：输入网络接口的私有 IP 地址（例如10.24.34.0）。
- NFS 选项：输入rw,no\_root\_squash。

### 为所有客户端配置 NFS 导出

您可以指定允许根用户访问所有客户机的 NFS 导出。

### 为所有客户端配置 NFS 导出

- 对于您的 FSx for OpenZFS 文件系统卷，配置以下 NFS 导出设置：
  - 客户地址：输入\*。
  - NFS 选项：输入rw,no\_root\_squash。

### 为 OpenZFS 创建您的 FSx 传输位置

要创建位置，您需要现有 FSx for OpenZFS 文件系统。如果你没有，请参阅[适用于 OpenZFS 的 Amazon FSx 用户指南中的 Amazon FSx for OpenZFS 入门](#)。

### 使用控制台创建 Amazon FSx for OpenZFS 位置的 FSx DataSync

1. [通过 https://console.aws.amazon.com/datasync/](https://console.aws.amazon.com/datasync/) 打开AWS DataSync主机。

2. 在左侧导航窗格中，选择 Locates ( 地点 ) ，然后选择 Creates ( 创建地点 ) 。
3. 对于位置类型，选择亚马逊 FsX。

您可在稍后将此位置配置为源或目标。

4. 对于 FsX 文件系统，选择要用作位置的 FsX for OpenZFS 文件系统。
5. 在装载路径中，输入 FSx for OpenZFS 文件系统的装载路径。

路径必须以开头 /fsx ，可以是文件系统中任何现有的目录路径。将该位置用作源时，从装载路径 DataSync 读取数据。将该位置用作目标时，DataSync 将所有数据写入装载路径。如果未提供子目录，则 DataSync 使用根卷目录 ( 例如 /fsx ) 。

6. 对于安全组，最多选择五个安全组，为您的 FsX for OpenZFS 文件系统提供网络访问权限。

安全组必须提供对 FSx for OpenZFS 文件系统的网络端口的访问权限。文件系统必须允许来自安全组的网络访问。

有关安全组的更多信息，请参阅 Amazon FSx for OpenZFS User Guide 中的使用 Amazon [VPC 的文件系统访问控制](#)。

7. ( 可选 ) 展开其他设置，对于 NFS 版本，选择 DataSync 用于访问文件系统的 NFS 版本。

默认情况下，DataSync 使用 NFS 版本 4.1。

8. ( 可选 ) 在 “密钥” 和 “值” 字段中输入值以标记 FSx for OpenZFS 文件系统。

标签可帮助您管理、筛选和搜索位置。我们建议至少为您的位置创建一个名称标签。

9. 选择创建地点。

## 使用适用 NetApp 于 ON AWS DataSync TAP 的亚马逊 FSx 配置传输

要将数据传输到您的 Amazon FSx for NetApp ONTAP 文件系统或从中传输数据，您必须创建一个 AWS DataSync 传输位置。DataSync 可以将此位置用作传输数据的源或目的地。

### 访问 FSx for ONTAP 文件系统

要访问 FSx for ONTAP 文件系统，请使用虚拟私有云 (VPC) 中的 [网络接口](#) 在文件系统上 DataSync 安装存储虚拟机 (SVM)。DataSync 仅当您创建包含 FSx for ONTAP 位置的任务时，才会在文件系统的首选子网中创建这些网络接口。

**Note**

与您一起使用的 VPC DataSync 必须具有默认租期。不支持具有专用租户的 VPC。有关更多信息，请参阅[使用 VPC](#)。

**主题**

- [支持的协议](#)
- [不支持的协议](#)
- [选择正确的协议](#)

**支持的协议**

DataSync可以使用以下协议连接到 FSx for ONTAP 文件系统的 SVM 并复制数据：

- 网络文件系统 (NFS) — 使用 NFS 协议，DataSync使用用户 ID (UID) 和组 ID (GID) 为 AUTH\_SYS 的安全机制向您的 SVM 进行身份验证。

**Note**

DataSync目前仅支持 Amazon FS 版本 3 FSx or ONTAP 位置。如果您需要复制 NFS 版本 4 访问控制列表 (ACL)，请[咨询AWS存储专家](#)。

- 服务器消息块 (SMB) — 使用 SMB 协议，DataSync使用您提供的凭据向 SVM 进行身份验证。创建位置时，您可以在 SVM 中指定本地用户或在 Microsoft Active Directory 中指定域用户。

要使用 SMB 在 FSx for ONTAP 文件系统（或使用 SMB 的其他类型的文件系统）之间进行复制，您的源位置和目标位置必须属于同一 Active Directory 域或者它们的域之间具有 Active Directory 信任关系。

默认情况下，DataSync会根据与 SMB 文件服务器的协商自动选择 SMB 协议版本。您也可以配置 DataSync为使用特定版本，但我们建议仅在与 SMB 文件服务器自动协商时才DataSync这样做。

有关DataSync控制台和 API 中使用您的 FsX for ONTAP 位置配置 SMB 版本的选项列表，请参阅下表：

控制台选项	API 选项	描述
自动	AUTOMATIC	DataSync 并且 SMB 文件服务器会协商 2.1 和 3.1.1 之间相互支持的 SMB 的最高版本。  这是默认和推荐选项。如果您改为选择文件服务器不支持的特定版本，则可能会出现 Operation Not Supported 错误。
SMB 3.0.2	SMB3	将协议协商限制为仅 SMB 版本 3.0.2。
SMB 2.1	SMB2	将协议协商限制为仅 SMB 版本 2.1。
SMB 2.0	SMB2_0	将协议协商限制为仅 SMB 版本 2.0。

## 不支持的协议

DataSync 无法使用 iSCSI ( 互联网小型计算机系统接口 ) 协议访问 FSx for ONTAP 文件系统。

## 选择正确的协议

要在 FsX 中保留文件元数据以便 ONTAP 迁移，请将您的 DataSync 源位置和目标位置配置为使用相同的协议。在支持的协议之间，SMB 以最高保真度保留元数据 ( 有关详细信息，[如何 AWS DataSync 处理元数据和特殊文件](#) 请参阅 )。

从通过 NFS 为用户提供服务的 Unix (Linux) 服务器或网络连接存储 (NAS) 共享进行迁移时，请执行以下操作：

1. 为 [Unix \(Linux\) 服务器或 NAS 共享创建 NFS 位置](#)。( 这将是您的源位置。 )
2. 使用 Unix 安全风格为要将数据传输到的 ONTAP 卷配置 FS x。
3. 为配置为 NFS 的 FSx for ONTAP 文件系统创建一个位置。( 这将是您的目的地位置。 )

从通过 SMB 为用户提供服务的 Windows 服务器或 NAS 共享进行迁移时，请执行以下操作：

1. [为 Windows 服务器或 NAS 共享创建 SMB 位置](#)。( 这将是您的源位置。 )
2. 使用 [NTFS 安全风格为要将数据传输到的 ONTAP 卷配置 FS x](#)。
3. 为为 SMB 配置的 FSx for ONTAP 文件系统创建一个位置。( 这将是您的目的地位置。 )

如果您的 FsX for ONTAP 环境使用多个协议，我们建议与AWS存储专家合作。要了解多协议访问的最佳实践，请参阅使用 [Amazon FSx for ONTAP 启用多协议工作负载](#)。NetApp

## 为 ONTAP 创建您的 FSx 传输位置

要创建位置，您需要现有 FSx for ONTAP 文件系统。如果你没有，请参阅《[亚马逊 FSx for ONTAP 用户指南](#)》中的 [Amazon FSx for NetApp NetApp ON TAP 入门](#)。

使用控制台指定 FSx for ONTAP 文件系统 DataSync

1. [通过 https://console.aws.amazon.com/datasync/](https://console.aws.amazon.com/datasync/) 打开AWS DataSync主机。
2. 在左侧导航窗格中，展开“数据传输”，然后选择“位置”和“创建位置”。
3. 对于位置类型，选择亚马逊 FsX。

您可在稍后将此位置配置为源或目标。

4. 对于 FSx 文件系统，请选择要用作位置的 FSx for ONTAP 文件系统。
5. 对于存储虚拟机，请在文件系统中选择要将数据复制到或从中复制数据的文件系统存储虚拟机 (SVM)。
6. 对于 Mount path (挂载路径)，请指定您将在其中复制数据的指向 SVM 中的文件共享路径。

您可以指定连接路径 (也称为挂载点)、qtree 路径 (用于 NFS 文件共享) 或共享名称 (用于 SMB 文件共享)。例如，您的挂载路径可能是 /vol1、/vol1/tree1 或 /share1。

### Tip

不要在 SVM 的根卷中指定路径。有关更多信息，请参阅 Amazon FSx for ONTAP 用户指南中的管理 FSx for NetApp ON [TAP 存储虚拟机](#)。

7. 对于安全组，最多选择五个 Amazon EC2 安全组，以提供对文件系统的首选子网的访问权限。

安全组必须允许以下端口上的出站流量 (具体您使用的协议而定)：

- NFS — TCP 端口 111、635 和 2049
- SMB — TCP 端口 445

文件系统的安全组还必须允许同一端口上的入站流量。

8. 对于 Protocol (协议)，请选择DataSync用于访问文件系统 SVM 的数据传输协议。



有关更多信息，请参阅[选择正确的协议](#)：

## NFS

DataSync使用 NFS 版本 3。

## SMB

配置 SMB 版本、用户名、密码和 Active Directory 域名（如果需要）以访问 SVM。

- （可选）展开其他设置并选择DataSync要在访问您的 SVM 时使用的 SMB 版本。

默认情况下，DataSync会根据与 SMB 文件服务器的协商自动选择版本。有关更多信息，请参阅[支持的协议](#)：

- 对于用户，请输入一个用户名，该用户名可以挂载位置并访问 SVM 中所需的文件、文件夹和元数据。

如果您在 Active Directory 中提供用户，请注意以下事项：

- 如果您使用AWS Directory Service for Microsoft Active Directory，则用户必须是AWS委托 FSx 管理员组的成员。
- 如果您使用的是自行管理的 Active Directory，则用户必须是域管理员组或在创建文件系统时为文件系统管理指定的自定义组的成员。

确保用户具有复制所需数据所需的权限：

- SE\_TCB\_NAME— 设置对象所有权和文件元数据所必需的。使用此权限，您还可以复制 NTFS 自由访问列表（DACL）。
- SE\_SECURITY\_NAME— 可能需要复制 NTFS 系统访问控制列表（SACL）。此操作特别需要 Windows 权限，该权限授予域管理员组的成员。如果将任务配置为复制 SACL，请确保用户具有所需的权限。有关复制 SACL 的信息，请参阅[管理文件、对象和元数据的AWS DataSync传输方式](#)。

- 对于密码，输入您指定的可以访问 SVM 的用户的密码。
- （可选）对于 Active Directory 域名，请输入 SVM 所属的 Active Directory 的完全限定域名（FQDN）。

9. （可选）在“密钥”和“值”字段中输入值以标记 FSx for ONTAP 文件系统。

标签可帮助您管理、筛选和搜索AWS资源。我们建议至少为您的位置创建一个名称标签。

10. 选择创建地点。



# 通过以下方式传输到其他云存储或从其他云存储进行传输AWS DataSync

使用AWS DataSync，您可以在其他一些公共云和AWS存储服务之间传输数据。有关更多信息，请参阅[我可以在哪里传输数据DataSync？](#)。

## 主题

- [使用 Google 云存储配置AWS DataSync传输](#)
- [使用 Azure Blob 存储配置AWS DataSync传输](#)
- [使用微软 Azure Files SMB 共享配置AWS DataSync传输](#)

## 使用 Google 云存储配置AWS DataSync传输

以下教程显示了AWS DataSync如何使用将对象从 Google Cloud Storage 存储桶迁移到 Amazon S3 存储桶。

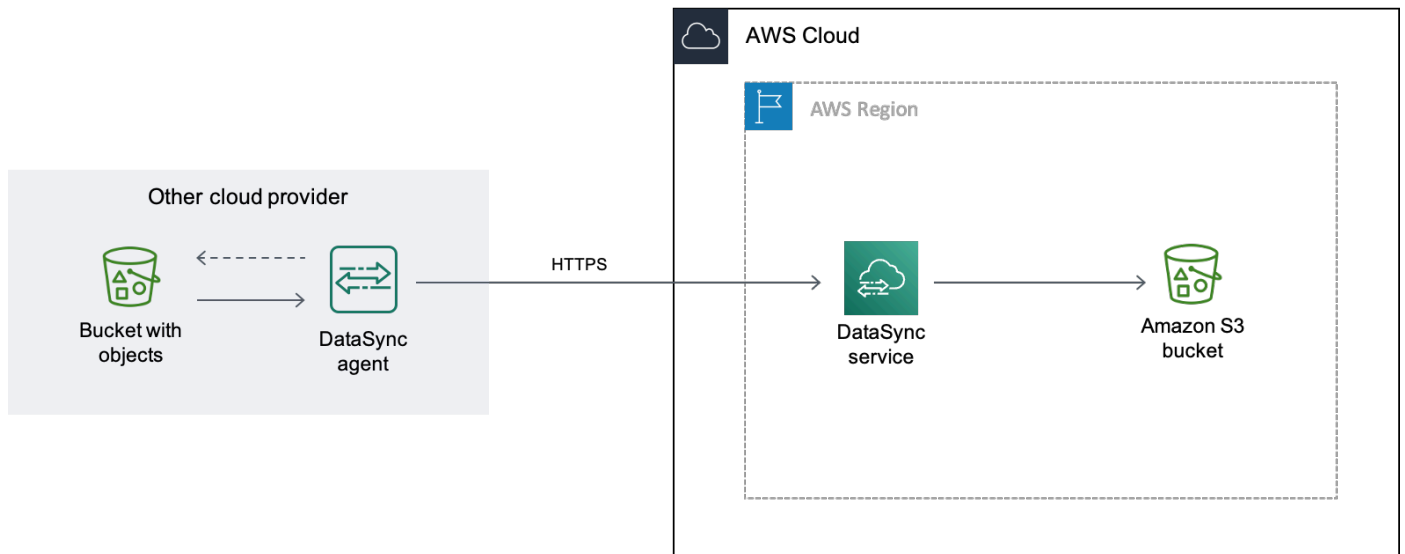
## 概览

由于DataSync与 [Google Cloud Storage XML API](#) 集成，因此您无需编写代码即可将对象复制到 Amazon S3 中。其工作原理取决于您在何处部署促进传输的DataSync代理。

### Agent in Google Cloud

1. 您在 Google 云环境中部署DataSync代理。
2. 代理使用基于哈MAC 散列消息验证码 ( HMAC ) 密钥读取 Google Cloud Storage 散列消息验证码 ( HMAC ) 密钥。
3. 您的 Google Cloud Storage 存储分区中的对象通过公共终端节点安全地通过 TLS 1.2 移入。AWS Cloud
4. 该DataSync服务会将数据写入 S3 存储桶。

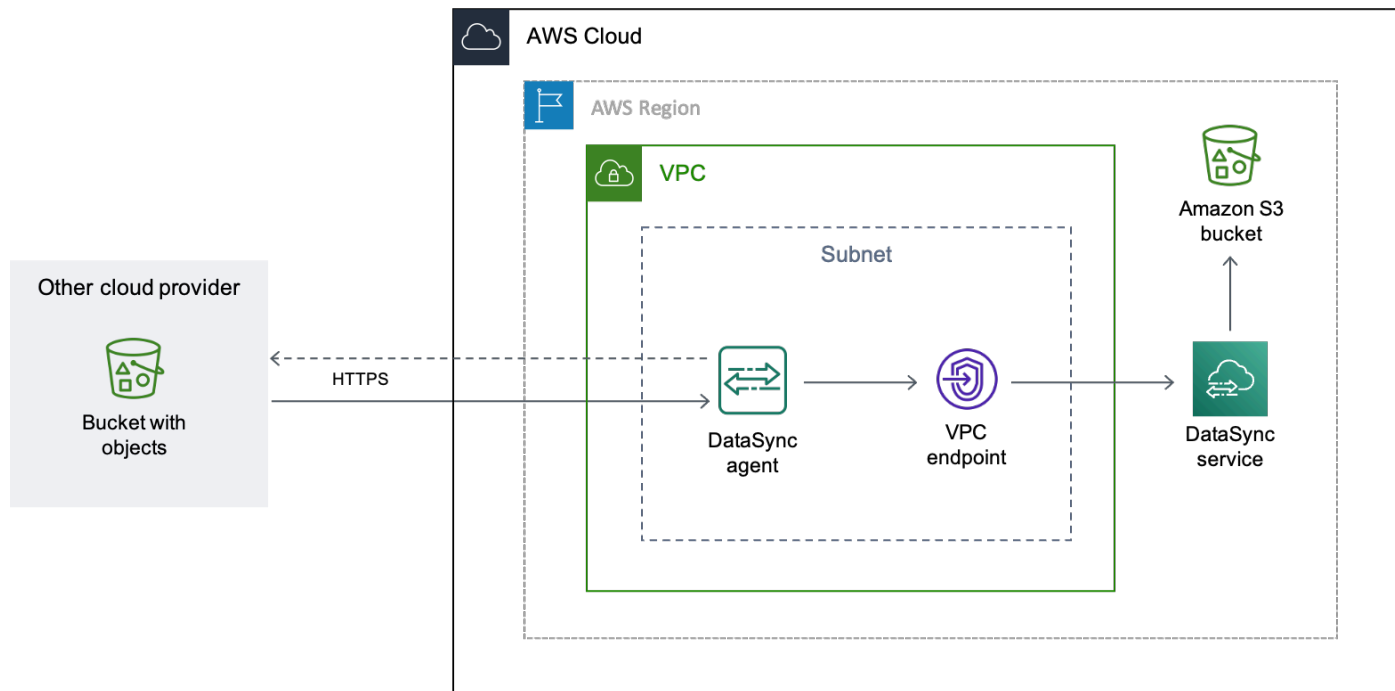
下图说明了该传输。



### Agent in your VPC

1. 您在AWS环境的虚拟私有云 ( VPC ) 中部署DataSync代理。
2. 代理使用基于哈MAC 散列消息验证码 ( HMAC ) 密钥读取 Google Cloud Storage 散列消息验证码 ( HMAC ) 密钥。
3. 通过使用私有 VPC 终端节点，您的 Google Cloud Storage 存储分区中的AWS Cloud对象通过 TLS 1.2 安全地移动到。
4. 该DataSync服务会将数据写入 S3 存储桶。

下图说明了该传输。



## 成本

与本次迁移相关的费用包括：

- 运行 Google [Compute Engine](#) 虚拟机 (VM) 实例 ( 如果您在 Google Cloud 中部署DataSync代理 )
- 运行 A [amazon EC2](#) 实例 ( 如果您在 VPC 内的 VPC 中部署DataSync代理AWS )
- 使用以下方式传输数据 [DataSync](#)
- 将数据传输出[谷歌云存储](#)
- 在 [Amazon S3](#) 中存储数据

## 先决条件

开始之前，请执行以下操作 ( 如果尚未完成 )：

- 使用您要转移到的对象@@ [创建 Google 云存储分区](#)AWS。
- [注册AWS 账户](#)。
- [创建一个 Amazon S3 存储桶](#)，用于在对象进入后对其进行存储AWS。

## 第 1 步：为 Google Cloud Storage Bucket Bucket

DataSync使用与您的 Google 服务帐号关联的 HMAC 密钥对您从中传输数据的存储分区进行身份验证和读取。（有关如何创建 HMAC 密钥的详细说明，请参阅 [Google 云存储文档](#)。）

### 创建 HMAC 密钥

1. 为您的 Google 服务帐号创建 HMAC 密钥。
2. 确保您的 Google 服务帐号至少具有 Storage Object Viewer 权限。
3. 将 HMAC 密钥的访问 ID 和密钥保存在安全位置。

稍后您将需要这些物品来配置您的 DataSync 源位置。

## 步骤 2：配置网络

此迁移的网络要求取决于您要如何部署 DataSync 代理。

对于谷歌云中的 DataSync 代理来说

如果您想在 Google Cloud 中托管 DataSync 代理，请将您的网络配置为 [允许通过公共终端节点进行 DataSync 传输](#)。

对于您的 VPC 中的 DataSync 代理

如果您想在其中托管代理 AWS，则需要一个带有接口终端节点的 VPC。DataSync 使用 VPC 终端节点来促进传输。

为 VPC 终端节点配置您的网络

1. 如果还没有存储桶，请在与 S3 存储桶 AWS 区域相同的 [VPC 中创建](#) VPC。
2. [为您的 VPC 创建私有子网](#)。
3. 使用 @@ [为创建 VPC 终端节点](#) AWS PrivateLink。DataSync
4. 将您的网络配置为 [允许通过 VPC 终端节点进行 DataSync 传输](#)。

要进行必要的配置更改，您可以修改与 VPC 终端节点关联的安全组。有关安全组的更多信息，请参阅《Amazon VPC 用户指南》中的 [使用安全组控制到资源的流量](#)。

## 步骤 3：创建 DataSync 代理

您需要一个可以访问和读取您的 Google 云存储分区的数据同步代理。

## For Cloud

在这种情况下，DataSync代理在您的 Google Cloud 环境中运行。

开始之前：[安装谷歌云命令行界面](#)。

### 为谷歌云创建代理

1. 通过 <https://console.aws.amazon.com/datasync/> 打开AWS DataSync主机。
2. 在左侧导航窗格中，选择代理，然后选择创建代理。
3. 对于虚拟机管理程序，选择 VMware ESXi，然后选择下载映像以下载包含代理的.zip文件。
4. 打开终端。通过运行以下命令解压缩映像：

```
unzip AWS-DataSync-Agent-VMWare.zip
```

5. aws-datasync通过运行以下命令从开始提取代理.ova文件的内容：

```
tar -xvf aws-datasync-2.0.1655755445.1-x86_64.xfs.gpt.ova
```

6. 运行以下 Google CLI 命令，将代理.vmdk文件导入谷歌云中：

```
gcloud compute images import aws-datasync-2-test \  
  --source-file INCOMPLETE-aws-datasync-2.0.1655755445.1-x86_64.xfs.gpt-disk1.vmdk \  
  \  
  --os centos-7
```

#### Note

导入.vmdk文件最多可能需要两个小时。

7. 为刚刚导入的代理映像创建并启动虚拟机实例。

该实例需要为您的代理配置以下配置。（有关如何创建实例的详细说明，请参阅 [Google 云计算引擎文档](#)。）

- 对于机器类型，请选择以下选项之一：
  - e2-standard-8 — 适用于传输多达 2000 万个文件的DataSync任务。
  - e2-standard-16 — 适用于传输超过 2000 万个文件的DataSync任务。
- 有关启动磁盘设置，请转到自定义映像部分。然后选择刚才导入的DataSync代理映像。

- 对于服务帐号设置，请选择您的 Google 服务帐号（与您在[步骤 1](#)中使用的帐号相同）。
- 对于防火墙设置，选择允许 HTTP（端口 80）流量的选项。

要激活DataSync代理，必须在代理上打开端口 80。该端口不需要公开访问。激活后，DataSync关闭端口。

8. 虚拟机实例运行后，记录其公有 IP 地址。

您需要此 IP 地址来激活代理。

9. 返回到 DataSync 控制台。在下载代理映像的创建代理屏幕上，执行以下操作以激活代理：

- 对于端点类型，选择公共服务终端节点选项（例如，美国东俄亥俄州的公共服务终端节点）。
- 对于激活密钥，选择自动从代理获取激活密钥。
- 对于代理地址，输入您刚刚创建的代理虚拟机实例的公有 IP 地址。
- 选择“获取密钥”。

10. 为您的代理命名，然后选择“创建代理”。

您的代理已联机并准备好移动数据。

## 适用于您的 VPC

在这种情况下，代理作为 Amazon EC2 实例运行，该实例在与您关联的 VPC 中运行AWS 账户。

开始之前：[设置AWS Command Line Interface \(AWS CLI\)](#)。

## 为您的 VPC 创建代理

1. 打开 终端。确保将您的配置AWS CLI文件配置为使用与您的 S3 存储桶关联的账户。
2. 复制以下命令。*vpc-region*替换为AWS 区域您的 VPC 所在的位置（例如us-east-1）。

```
aws ssm get-parameter --name /aws/service/datasync/ami --region vpc-region
```

3. 运行 命令。在输出中，记录该"Value"属性。

该值是您指定区域的数据同步 Amazon Machine Image (AMI) ID。例如，AMI ID 可能如下所示ami-1234567890abcdef0。

4. 复制以下 URL。同样，替换*vpc-region*为您的 VPCAWS 区域 所在的位置。然后，*ami-id*替换您在上一步中记下的 AMI ID 替换。

```
https://console.aws.amazon.com/ec2/v2/home?region=vpc-region#LaunchInstanceWizard:ami=ami-id
```

5. 将该 URL 粘贴到浏览器中。

AWS Management Console 显示屏中的 Amazon EC2 实例启动页面。

6. 对于实例类型，选择为 [DataSync 代理推荐的 Amazon EC2 实例](#) 之一。
7. 对于 key pair，选择现有密钥对，或新建一个密钥对。
8. 对于网络设置，选择要在其中部署代理的 VPC 和子网。
9. 选择 Launch instance (启动实例)。
10. 在 Amazon EC2 实例运行后，[选择您的 VPC 终端节点](#)。
11. [激活您的代理](#)。

## 第 4 步：为您的 Google 云存储分区创建 DataSync 源位置

要为您的 Google Cloud Storage 存储分区设置 DataSync 位置，您需要在 [步骤 1](#) 中创建的 HMAC 密钥的访问权限 ID 和密钥。

### 创建 DataSync 源位置

1. [通过 https://console.aws.amazon.com/datasync/](https://console.aws.amazon.com/datasync/) 打开 AWS DataSync 主机。
2. 在左侧导航窗格中，展开数据传输，然后选择位置和创建位置。
3. 对于位置类型，选择对象存储。
4. 对于代理，请选择您在 [步骤 3](#) 中创建的代理。
5. 对于服务器，输入 **storage.googleapis.com**。
6. 在存储分区名称中，输入您的 Google 云存储分区的名称。
7. 展开 Additional settings (其他设置)。对于服务器协议，选择 HTTPS。对于服务器端口，选择 443。
8. 向下滚动到认证部分。确保选中“需要凭据”复选框，然后执行以下操作：
  - 对于访问密钥，请输入您的 HMAC 密钥的访问 ID。
  - 对于私有密钥，请输入 HMAC 密钥的密钥。
9. 选择“创建地点”。

## 步骤 5：为您的 S3 存储桶创建DataSync目标位置

您需要一个DataSync位置，用于存放您希望数据存放到哪里。

### 创建DataSync目标位置

1. 通过 <https://console.aws.amazon.com/datasync/> 打开AWS DataSync主机。
2. 在左侧导航窗格中，展开数据传输，然后选择位置和创建位置。
3. [为 S3 存储桶创建DataSync位置。](#)

如果您在 VPC 中部署DataSync代理，则本教程假定 S3 存储桶与您的 VPC 和DataSync代理位于 AWS 区域相同的 VPC 中。

## 步骤 6：创建并启动DataSync任务

配置好来源和目标位置后，即可开始将数据移入AWS。

### 创建和启动DataSync任务

1. 通过 <https://console.aws.amazon.com/datasync/> 打开AWS DataSync主机。
2. 在左侧导航窗格中，展开“数据传输”，然后选择“任务”，然后选择“创建任务”。
3. 在 Configure Logs 页面上，执行以下操作：
  - a. 选择“选择现有地点”。
  - b. 选择您在[步骤 4 中创建的源位置，然后选择下一步。](#)
4. 在 Configure Logs 页面上，执行以下操作：
  - a. 选择“选择现有地点”。
  - b. 选择您在[步骤 5 中创建的目标位置，然后选择下一步。](#)
5. 在 Configure Launch 页面上，执行以下操作：
  - a. 在“数据传输配置”下，展开其他设置并清除“复制对象标签”复选框。

#### Important

由于使用 Amazon S3 API与DataSync 与 Google Cloud Storage 通信，因此存在一个限制，如果您尝试复制对象标签，可能会导致DataSync任务失败。



- b. 配置所需的任何其他任务设置，然后选择“下一步”。
6. 在 Review 页面上，检查您的设置，然后选择 Continue 页面上。
  7. 在任务的详细信息页面上，选择 Start (开始)，然后选择以下选项之一：
    - 要在不修改的情况下运行任务，请选择“以默认值启动”。
    - 要在运行任务之前对其进行修改，请选择“使用替代选项开始”。

任务完成后，您将在 S3 存储分区中看到 Google Cloud Storage 存储分区中的对象。

## 使用 Azure Blob 存储配置 AWS DataSync 传输

此功能位于 AWS DataSync 的预览版中，可能会发生变化。

要从 Azure Blob 存储中传输 blob (也称为对象)，必须创建 AWS DataSync 传输位置。然后，您可以将 blob 移至 [任何 DataSync 支持的 AWS 存储服务](#)。

### 访问 Azure Blob 存储

要传输对象，DataSync 需要具有读取 Azure Blob 存储的权限。您的对象还必须位于 DataSync 可以使用的访问层中。

#### 主题

- [SAS 代币](#)
- [访问套餐](#)

### SAS 代币

共享访问签名 (SAS) 令牌指定您的 blob 存储的访问权限。(有关 SAS 的更多信息，请参阅 [Azure Blob 存储文档](#)。)

您可以在不同的访问级别生成 SAS 令牌。DataSync 支持具有以下访问级别的令牌：

- 账户
- 容器
- 文件夹

要提供DataSync转移所需的访问权限，您的 SAS 令牌至少必须具有读取和列出权限。根据访问级别，您可能还需要使用其他权限配置令牌。例如，具有账户级访问权限的令牌比具有文件夹级访问权限的令牌需要更多的权限。

## 主题

- [用于账户级访问的 SAS 令牌权限](#)
- [容器级访问的 SAS 令牌权限](#)
- [用于文件夹级访问的 SAS 令牌权限](#)
- [SAS 到期政策](#)

### 用于账户级访问的 SAS 令牌权限

DataSync要求您配置具有以下权限的账户级访问令牌：

- 允许的服务 — Blob
- 允许的资源类型-容器、对象（如果您不包含这些权限，则DataSync无法传输您的对象元数据）
- 允许的权限-读取、列出
- 允许的 blob 索引权限-读取/写入（如果DataSync要复制对象标签）

### 容器级访问的 SAS 令牌权限

DataSync要求您配置具有以下权限的容器级访问令牌：

- Read
- List
- 标签（如果你DataSync想复制对象标签）

#### Note

在 Azure 门户中生成 SAS 令牌时，无法添加可选标签权限。要添加权限，请改为使用 [Azure Storage Explorer](#) 应用程序生成[令牌或生成提供帐户级访问权限的 SAS 令牌](#)。

### 用于文件夹级访问的 SAS 令牌权限

如果您将 Azure DataSync a Lake Storage Gen2 与 blob 存储一起使用，则可以提供仅允许访问容器中特定文件夹的 SAS 令牌。

DataSync要求您配置具有以下权限的文件夹级访问令牌：

- Read
- List

有关配置此类令牌的更多信息，请参阅 [Azure Blob 存储文档](#)。

在为 [创建 Azure Blob 存储位置](#) 时，必须指定同一文件夹的完整路径。DataSync

## SAS 到期政策

在您预计完成转移之前，请确保您的 SAS 不会过期。有关配置 SAS 过期策略的信息，请参阅 [Azure Blob 存储文档](#)。

如果 SAS 在传输期间过期，DataSync将无法再访问你的 Azure Blob 存储位置（你可能会看到“无法打开目录”错误）。发生这种情况时，您必须使用新的 SAS 令牌 [更新您的位置](#) 并重新启动DataSync任务。

## 访问套餐

DataSync可以在热层或冷层中传输 Azure Blob 存储对象。它跳过在存档访问层传输对象。您必须将存档层中的对象重新水合到热层或冷层，然后才能传输它们。

## Blob 类型

DataSync可以传输任何 blob 类型，包括块、页面和附加 blob。有关更多信息，请参阅 [Azure Blob 存储文档](#)。

## 设置

目前，您只能在DataSync控制台AWS Command Line Interface、或上使用此功能AWS SDK for Java 2.x。

### Note

如果您使用的是DataSync主机，则无需进行任何设置。

## 使用进行设置AWS CLI

以下说明描述了如何设置 [AWS Command Line Interface\(AWS CLI\) 版本 2](#) 以使用此功能。

## 设置 AWS CLI

1. 下载DataSync [API 模型](#)并解压缩 JSON 文件。
2. 运行以下命令以将服务模型添加到AWS CLI配置中。

此示例创建了一个datasync-azure-blob服务名称。您可以指定其他名称。

```
aws configure add-model \  
  --service-model file:///datasync-service-model.json \  
  --service-name datasync-azure-blob
```

添加模型后，即可将此功能与AWS CLI。

## 使用进行设置AWS SDK for Java 2.x

以下说明介绍了如何设置AWS SDK for Java 2.x以使用该功能。

### 设置适用于 Java 2.x 的开发工具包

1. 请确保您有AwsJavaSdk-DataSync-2.0.jar文件。如果不是，请联系您的DataSync联系人。
2. 在您的 Maven 项目中，检查pom.xml文件中是否存在DataSync Java SDK 的显式依赖关系。如果没有，请使用以下示例创建依赖关系：

```
<dependency>  
  <groupId>software.amazon.awssdk</groupId>  
  <artifactId>datasync</artifactId>  
  <version>2.17.290</version>  
</dependency>
```

3. 运行以下命令：

```
mvn org.apache.maven.plugins:maven-install-plugin:3.0.0-M1:install-file \  
  --batch-mode -Dfile=/AwsJavaSdk-DataSync-2.0.jar \  
  -DgroupId=software.amazon.awssdk \  
  -DartifactId=datasync \  
  -Dversion=2.17.290 \  
  -Dpackaging=jar
```

请记住：

- 确保使用正确的 `AwsJavaSdk-DataSync-2.0.jar` 文件路径。
- 该 **2.17.290** 版本就是一个例子。

## 创建您的DataSync代理

您需要一个可以连接到 Azure Blob 存储容器的DataSync代理。代理读取您的容器并便于将您的对象移至AWS。

### 创建一个代理

1. [下载并部署 Virtual Cloud \(VPC\) 中您AWS的 Amazon EC2 代理](#)。
2. [指定代理用来与之通信的 VPC 服务终端节点AWS](#)。
3. 将您的网络配置为与 [VPC 终端节点](#) 配合使用。
4. [激活代理](#)。

## 创建你的 Azure Blob 存储传输位置

你可以配置DataSync为使用现有 Azure Blob 存储容器作为传输的源位置。

开始之前：确保你知道[如何DataSync访问你的 Azure Blob 存储](#)以及如何使用 blob 存储[访问层](#)。

### 使用 DataSync 控制台

1. [通过 https://console.aws.amazon.com/datasync/](https://console.aws.amazon.com/datasync/) 打开AWS DataSync主机。
2. 在左侧导航窗格中，展开“数据传输”，然后选择“位置”和“创建位置”。
3. 对于位置类型，选择微软 Azure Blob 存储。
4. 对于代理，选择可以连接你的 Azure Blob 存储容器的DataSync代理。

您可以选择多个代理。有关更多信息，请参阅[使用多个AWS DataSync代理进行转账](#)：

5. 对于容器 URL，请输入您的传输中涉及的容器的 URL。
6. 对于文件夹，如果您想将传输限制到容器中的虚拟目录（例如 `/my/images`），请输入路径段。

DataSync仅复制带有这些路径段的对象。

**⚠ Important**

如果您的 SAS 令牌仅允许访问文件夹，则必须在此处指定该文件夹的完整路径才能进行传输。

- 对于 SAS 令牌，请输入DataSync允许访问您的 BLOB 存储的 SAS 令牌。

该令牌是 SAS URI 字符串的一部分，位于存储资源 URI 和问号之后。代币如下所示：

```
sp=r&st=2023-12-20T14:54:52Z&se=2023-12-20T22:54:52Z&spr=https&sv=2021-06-08&sr=c&sig=aBBKD
%2FXTI9E%2F%2Fmq171%2BZU178wcvqU%3D
```

- ( 可选 ) 输入 “密钥” 和 “值” 字段的值以标记位置。

标签可帮助您管理、筛选和搜索AWS资源。我们建议至少为您的位置创建一个名称标签。

- 选择 “创建地点”。

## 使用 AWS CLI

**i Note**

以下说明使用该`datasync-azure-blob`命令，但您可以将此命令命名为其他名称。有关更多信息，请参阅[使用进行设置AWS CLI](#)：

- 复制以下`create-location-azure-blob`命令：

```
aws datasync-azure-blob create-location-azure-blob \
  --container-url "https://path/to/container" \
  --authentication-type "SAS" \
  --sas-configuration '{
    "Token": "your-sas-token"
  }' \
  --agent-arns my-datasync-agent-arn \
  --subdirectory "/path/to/my/data" \
  --tags [{"Key": "key1", "Value": "value1"}]
```

- 对于`--container-url`参数，请指定传输中涉及的 Azure Blob 存储容器的 URL。
- 对于`--authentication-type`参数，指定SAS。

- 对于 `--sas-configuration` 参数的 `Token` 选项，请指定允许 DataSync 访问您的 BLOB 存储的 SAS 令牌。

该令牌是 SAS URI 字符串的一部分，位于存储资源 URI 和问号之后。代币如下所示：

```
sp=r&st=2023-12-20T14:54:52Z&se=2023-12-20T22:54:52Z&spr=https&sv=2021-06-08&sr=c&sig=aBBKD%2FXTI9E%2F%2Fmq171%2BZU178wqwU%3D
```

- 对于 `--agent-arns` 参数，指定可帮助您连接到容器的 DataSync 代理的 Amazon 资源名称 (ARN)。

以下为示例如代理 ARN：`arn:aws:datsync:us-east-1:123456789012:agent/agent-01234567890deadfb`

您可以指定多个代理。有关更多信息，请参阅[使用多个 AWS DataSync 代理进行转账](#)：

- 对于 `--subdirectory` 参数，如果您想将传输限制到容器中的虚拟目录（例如 `/my/images`），请指定路径段。

DataSync 仅复制带有这些路径段的对象。

#### Important

如果您的 SAS 令牌仅允许访问子目录（或文件夹），则必须在此处指定该子目录的完整路径才能进行传输。

- （可帮助您管理、`--tags` 筛选和搜索位置的键值对，帮助您管理、筛选和搜索位置。

我们建议为您的位置创建一个名称标签。

- 运行 `create-location-azure-blob` 命令。

您会收到一条回复，其中显示了您创建的位置的 ARN。例如：

```
{
  "LocationArn": "arn:aws:datsync:us-east-1:123456789012:location/loc-12345678abcdefgh"
}
```

## 查看你的 Azure Blob 存储传输位置

你可以获取有关现有 Azure Blob 存储位置的详细信息。

### 使用 DataSync 控制台

1. 通过 <https://console.aws.amazon.com/datasync/> 打开AWS DataSync主机。
2. 在左侧导航窗格中，选择位置。
3. 选择你的 Azure Blob 存储位置。

您可以查看有关您的位置的详细信息，包括使用该位置的任何DataSync任务。

### 使用 AWS CLI

#### Note

以下说明使用该`datasync-azure-blob`命令，但您可以将此命令命名为其他名称。有关更多信息，请参阅[使用进行设置AWS CLI](#)：

1. 复制以下`describe-location-azure-blob`命令：

```
aws datasync-azure-blob describe-location-azure-blob \  
  --location-arn "your-azure-blob-location-arn"
```

2. 对于`--location-arn`参数，指定您创建的 Azure Blob 存储位置的 ARN。
3. 运行 `describe-location-azure-blob` 命令。

您会收到一条回复，其中显示了有关您的位置的详细信息。例如：

```
{  
  "LocationArn": "arn:aws:datsync:us-east-1:123456789012:location/  
loc-12345678abcdefgh",  
  "LocationUri": "azure-blob://my-user.blob.core.windows.net/container-1",  
  "AuthenticationType": "SAS",  
  "Subdirectory": "/my/images",  
  "AgentArns": ["arn:aws:datsync:us-east-1:123456789012:agent/  
agent-01234567890deadfb"],  
}
```



## 更新你的 Azure Blob 存储传输位置

如果需要，您可以修改您的位置配置。

### Note

你不能用DataSync控制台做到这一点。

## 使用 AWS CLI

### Note

以下说明使用该`datasync-azure-blob`命令，但您可以将此命令命名为其他名称。有关更多信息，请参阅[使用进行设置AWS CLI](#)：

1. 复制以下`update-location-azure-blob`命令：

```
aws datasync-azure-blob update-location-azure-blob \  
  --location-arn "your-azure-blob-location-arn" \  
  --authentication-type "SAS" \  
  --sas-configuration '{  
    "Token": "your-sas-token"  
  }' \  
  --agent-arns my-datasync-agent-arn \  
  --subdirectory "/path/to/my/data"
```

2. 对于`--location-arn`参数，为正在更新的 Azure Blob 存储位置指定 ARN。
3. 对于`--authentication-type`参数，指定SAS。
4. 对于`--sas-configuration`参数的Token选项，请指定允许DataSync访问您的 BLOB 存储的 SAS 令牌。

该令牌是 SAS URI 字符串的一部分，位于存储资源 URI 和问号之后。代币如下所示：

```
sp=r&st=2022-12-20T14:54:52Z&se=2022-12-20T22:54:52Z&spr=https&sv=2021-06-08&sr=c&sig=qCBKD  
%2FXTI9E%2F%2Fmq171%2BZU178wcwqU%3D
```

5. 对于`--agent-arns`参数，指定要连接到容器的DataSync代理的 Amazon 资源名称 (ARN)。

以下为示例代理 ARN：`arn:aws:datsync:us-east-1:123456789012:agent/agent-01234567890deadfb`

您可以指定多个代理。有关更多信息，请参阅[使用多个AWS DataSync代理进行转账](#)：

- 对于`--subdirectory`参数，如果您想将传输限制到容器中的虚拟目录（例如`/my/images`），请指定路径段。

DataSync仅复制带有这些路径段的对象。

### Important

如果您的 SAS 令牌仅允许访问子目录（或文件夹），则必须在此处指定该子目录的完整路径才能进行传输。

## 元数据处理

从 Azure Blob 存储传输时如何DataSync处理元数据取决于你的目标位置。

当你的目的地是	DataSync能够
<ul style="list-style-type: none"> <li>Amazon S3</li> </ul>	<p>将以下元数据复制到目标：</p> <ul style="list-style-type: none"> <li>用户定义的对象元数据</li> <li>对象标签</li> <li>以下系统定义的对象元数据： <ul style="list-style-type: none"> <li>Content-Disposition</li> <li>Content-Encoding</li> <li>Content-Language</li> <li>Content-Type</li> </ul> </li> </ul> <p>注意：在初始传输期间DataSync复制所有对象的系统元数据。如果您<a href="#">将任务配置为仅传输已更改的数据</a>，则除非对象的内容或用户元数据也已修改，否则DataSync不会在后续传输中复制系统元数据。</p>

当你的目的地是	DataSync能够
<ul style="list-style-type: none"> <li>• Amazon EFS</li> <li>• FSx for Lustre</li> <li>• FSx for OpenZFS</li> <li>• 适用于 ONTAP 的 FSx ( 使用 NFS )</li> </ul>	<p>对目标文件和文件夹应用<a href="#">默认 POSIX 元数据</a>。该方法包括使用默认 POSIX 用户 ID 和组 ID 值。</p>
<ul style="list-style-type: none"> <li>• FSx for Windows File Server</li> <li>• 适用于 ONTAP 的 FSx ( 使用 SMB )</li> </ul>	<p>将以下元数据应用于目标：</p> <ul style="list-style-type: none"> <li>• 文件和文件夹时间戳</li> <li>• 所有权取决于您在创建DataSync目标位置时指定的 Windows 用户</li> <li>• 权限继承自父目录</li> </ul>

## 限制

当使用以下方式从 Azure Blob 存储中移动数据时，当前存在以下限制DataSync：

- 你可以在DataSync除AWS GovCloud (US)和中国区域之外的任何[AWS 区域支持](#)的区域中创建 Azure Blob 存储位置。
- 你只能使用 Azure Blob 存储作为源位置。
- 传输到 Amazon S3 时，DataSync不会传输大于 5 TB 的 Azure Blob 存储对象。
- 传输到 Amazon S3 时，DataSync不会传输元数据超过 2 KB 的 Azure Blob 存储对象。

## 使用微软 Azure Files SMB 共享配置AWS DataSync传输

你可以配置为AWS DataSync将数据传输到微软 Azure 文件服务器消息块 (SMB) 共享或从中传输数据。

### Tip

有关将数据从 Azure Files SMB 共享移动到的完整演练AWS，请参阅[AWS存储博客](#)。

## 访问 SMB 共享

DataSync使用 SMB 协议连接到您的 SMB 共享，并使用您提供的凭据进行身份验证。

### 主题

- [支持的 SMB 协议版本](#)
- [所需的权限](#)

### 支持的 SMB 协议版本

默认情况下，DataSync会根据与 SMB 文件服务器的协商自动选择 SMB 协议的版本。

您也可以DataSync将配置为使用特定的 SMB 版本，但我们建议仅DataSync在无法自动与 SMB 文件服务器协商的情况下才这样做。（DataSync支持 SMB 版本 1.0 和更高版本。）

有关DataSync控制台和 API 中的选项列表，请参阅下表：

控制台选项	API 选项	描述
自动	AUTOMATIC	DataSync并且 SMB 文件服务器在 2.1 和 3.1.1 之间协商它们相互支持的 SMB 的最高版本。  这是默认和推荐的选项。如果您改为选择文件服务器不支持的特定版本，则可能会出现 Operation Not Supported 错误。
SMB 3.0.2	SMB3	将协议协商限制为仅 SMB 版本 3.0.2。
SMB 2.1	SMB2	将协议协商限制为仅 SMB 版本 2.1。
SMB 2.0	SMB2_0	将协议协商限制为仅 SMB 版本 2.0。
SMB 1.0	SMB1	将协议协商限制为仅 SMB 版本 1.0。

### 所需的权限

DataSync需要具有挂载和访问您的 SMB 位置的权限的用户。这可以是你的 Windows 文件服务器上的本地用户，也可以是在 Microsoft Active Directory 中定义的域用户。

要设置对象所有权，DataSync需要SE\_RESTORE\_NAME权限，该权限通常授予内置 Active Directory 群组、Backup 操作员和域管理员的成员。为用户DataSync提供此权限还有助于确保对文件、文件夹和文件元数据有足够的权限，NTFS 系统访问控制列表 (SACL) 除外。

复制 SACL 需要其他权限。具体而言，这需要 WindowsSE\_SECURITY\_NAME 权限，该权限授予域管理员组的成员。如果将任务配置为复制 SACL，请确保用户具有所需的权限。要了解有关将任务配置为复制 SACL 的更多信息，请参阅[管理文件、对象和元数据的AWS DataSync传输方式](#)。

当您在 SMB 文件服务器和适用于 Windows 文件服务器的 Amazon FsX 文件系统之间复制数据时，源位置和目标位置必须属于同一个 Microsoft Active Directory 域或者它们的域之间具有 Active Directory 信任关系。

## 使用控制台创建 Azure 文件传输位置

1. 通过 <https://console.aws.amazon.com/datasync/> 打开AWS DataSync主机。
2. 在左侧导航窗格中，展开“数据传输”，然后选择“位置”和“创建位置”。
3. 对于 Location type (位置类型)，选择 Server Message Block (SMB) (服务器消息块 (SMB))。

您可在稍后将此位置配置为源或目标。

4. 对于代理，选择一个或多个要连接到 SMB 共享的DataSync代理。

如果您选择多个代理，请确保您了解[在一个地点使用多个代理](#)。

5. 对于 SMB 服务器，输入要挂载DataSync代理的 SMB 共享的域名系统 ( DNS ) 名称或 IP 地址。

### Note

您无法指定 IP 版本 6 ( IPv6 ) 地址。

6. 在共享名称中，输入由您的 SMB 共享导出的共享名称，该共享DataSync将在其中读取或写入数据。

可以在共享路径中包含子目录 ( 例如/path/to/subdirectory )。确保网络中的其他 SMB 客户端也可以挂载此路径。

要复制子目录中的所有数据，DataSync必须能够装载 SMB 共享并访问其所有数据。有关更多信息，请参阅[所需的权限](#)：

7. ( 可选 ) 展开其他设置并选择DataSync要在访问 SMB 共享时使用的 SMB 版本。

默认情况下，DataSync会根据与 SMB 共享的协商自动选择版本。有关信息，请参阅 [支持的 SMB 协议版本](#)。

8. 对于“用户”，输入可以挂载您的 SMB 共享并有权访问传输中涉及的文件和文件夹的用户名。

有关更多信息，请参阅[所需的权限](#)：

9. 在 Password 中，输入可以挂载您的 SMB 共享和访问文件和文件夹的用户的密码。

10. ( 可选 ) 对于域，输入您的 SMB 共享所属的 Windows 域名称。

11. ( 可选 ) 选择添加标签来标记您的位置。

标签是键值对，帮助您管理、筛选和搜索位置。我们建议至少为您的位置创建一个名称标签。

12. 选择“创建地点”。

## 使用以下方式向边缘或离线存储传输或从边缘存储器传输AWS DataSync

使用AWS DataSync，您可以将数据传输到某些设备或从某些AWS Snow Family设备传输数据。有关更多信息，请参阅[我可以在哪里传输数据DataSync？](#)。

### 主题

- [在 Snowball Edge 上使用兼容 S3 的存储器配置](#)
- [使用以下AWS DataSync方式配置传输AWS Snowcone](#)

## 在 Snowball Edge 上使用兼容 S3 的存储器配置

使用AWS DataSync，您可以在AWS Snowball Edge设备或集群上的[Amazon S3 兼容存储](#)与以下任何AWS存储服务之间传输对象：

- [Amazon S3](#)
- [Amazon Elastic File System \(Amazon EFS\)](#)
- [Amazon FSx for Windows File Server](#)
- [Amazon FSx for Lustre](#)
- [Amazon FSx for ONTAP](#)
- [适用于NetApp ONTAP 的亚马逊 FSx](#)

## 先决条件

在您开始之前，请确保您已完成以下操作：

- 在您计划将AWS 区域数据传输到或从中传输数据的AWS存储资源中创建。例如，这可能是美国东部（弗吉尼亚州北部）的 S3 存储桶或 Amazon EFS 文件系统。
- 为进出本地存储环境的流量建立了广域网 (WAN) 连接。例如，您可以与建立这种连接[AWS Direct Connect](#)。

[创建DataSync代理](#)时，您将配置此 WAN 连接，以便DataSync可以在本地的 Amazon S3 兼容存储和中的存储资源之间传输数据AWS。

- 已下载并安装了 [Snowball Edge 客户端](#)。

## 访问您的 Amazon S3 兼容存储

要访问您的 Amazon S3 兼容存储桶，DataSync需要满足以下条件：

- 您的 Snowball Edge 设备或集群上的用户证书，这些凭证可以访问您向其传输数据或从中传输数据的存储桶。
- 一个 HTTPS 证书，DataSync允许验证DataSync代理与设备或集群上的s3api终端节点之间连接的真实性。

### 主题

- [获取用户凭证以访问您的 S3 存储桶](#)
- [获取s3api端点连接的证书](#)

## 获取用户凭证以访问您的 S3 存储桶

DataSync需要访问您在 Snowball Edge 设备或集群上使用的存储桶的用户的访问密钥和密钥。

### 获取用户凭证以访问您的存储桶

1. 打开终端并运行 Snowball Edge 客户端。

有关运行 Snowball Edge 客户端的更多信息，请参阅AWS Snowball Edge开发者指南[中的使用 Snowball Edge 客户端](#)。

2. 要获取与您的设备或集群关联的访问密钥，请运行以下snowballEdge命令：

```
snowballEdge list-access-keys
```

3. 在输出中，找到要使用的存储桶的DataSync访问密钥（例如，AKIAIOSFODNN7EXAMPLE）。
4. 要获得私有访问密钥，请运行以下snowballEdge命令。*access-key-for-datasync*替换为您在上一步骤中找到的访问密钥。

```
snowballEdge get-secret-access-key --access-key-id access-key-for-datasync
```

输出包括访问密钥的相应密钥（例如wJalrXUtnFEMI/K7MDENG/bPxrFiCYEXAMPLEKEY）。

5. 将访问密钥和私有密钥保存在您能记住的地方。

在为传输[配置DataSync源位置时](#)，您将需要这些密钥。

### 获取s3api端点连接的证书

您需要一个 HTTPS 证书来验证您的DataSync代理与 Snowball Edge 设备或集群上的s3api终端节点之间连接的真实性的。

### 获取s3api端点连接的证书

1. 在 Snowball Edge 客户端中，运行以下snowballEdge命令：

```
snowballEdge get-certificate
```

2. 将输出保存到 base64 编码 .pem 的文件中。

在为传输[配置DataSync源位置时](#)，您将指定此文件。

## 在本地存储环境中创建DataSync代理

在传输期间，DataSync使用[代理](#)读取或写入您的 Snowball Edge 设备或集群上的 Amazon S3 兼容存储设备。

此代理必须部署在您的本地存储环境中，在那里它可以通过您的网络连接到您的设备或集群。例如，您可以在具有本地网络访问集群的 VMware ESXi 虚拟机管理程序上运行代理。



## 在本地存储环境中创建DataSync代理

1. 确保[DataSync代理可以在您的虚拟机管理程序上运行](#)，并且您为代理分配了足够的虚拟机 (VM) 资源。
2. 在您的本地环境中部署代理。

有关说明，请根据您的部署代理的虚拟机管理程序类型参阅下列主题中的一个：

- [在 VMware 上部署您的代理](#)
- [在基于 Linux 内核的计算机 \(KVM\) 上部署代理](#)
- [在微软 Hyper-V 上部署代理](#)
- [在 Amazon EC2 上部署您的代理](#)

### Warning

我们不建议在 Amazon EC2 代理上部署代理来访问本地存储，因为网络延迟会增加。

3. 配置您的网络，允许代理与 Amazon S3 兼容存储器之间传输以下流量：

From	目的	协议和端口
DataSync 代理	设备或集群上s3api终端节点的虚拟网络接口 (VNI)。如果您有集群，则可以是任何s3api终端节点 VNI。	TCP 443 (HTTPS)

如果您需要在设备或集群上查找 VNI，请参阅在 Snowball Edge 上[描述您的虚拟网络接口](#)。

4. [选择代理将用于与之通信的服务端点](#)AWS。
5. [激活您的代理](#)。

## 为传输配置源位置

创建代理后，您可以配置DataSync传输的源位置。

**Note**

以下说明假设您是从 Amazon S3 兼容存储中传输的，但您也可以将此位置用作传输目的地。

### 使用DataSync控制台配置传输的源位置

1. 通过 <https://console.aws.amazon.com/datasync/> 打开AWS DataSync主机。
2. 在左侧导航窗格中，展开“数据传输”。选择“任务”，然后选择“创建任务”。
3. 在 Configure source location (配置源位置) 页面中，选择 Create a new location (创建新位置)。
4. 对于位置类型，选择对象存储。
5. 对于代理，请选择您在本地存储环境中创建的代理。
6. 对于服务器，输入您的 Amazon S3 兼容存储所使用的s3api终端节点的 VNI。

如果您有 Snowball Edge 集群而不是单个设备，则可以指定集群的任何s3api终端节点 VNI。

7. 在存储段名称中，输入您要从中传输对象的 Amazon S3 兼容存储桶的名称。
8. 对于“文件夹”，输入对象前缀。

DataSync仅传输带有此前缀的对象。

9. 要配置与 Snowball Edge 设备或集群的DataSync连接，请展开其他设置并执行以下操作：
  - a. 对于服务器协议，选择 HTTPS。
  - b. 对于服务器端口，输入**443**。
  - c. 对于证书，为[s3api端点连接](#)选择证书文件。
10. 选择需要证书，然后输入访问密钥和密钥以访问[您的 Snowball Edge 设备或集群上与 Amazon S3 兼容的存储桶](#)。
11. 选择下一步。

### 配置传输的目的地位置

您的转账目的地位置必须与您创建代理AWS账户的地点相同AWS 区域。

开始之前：确保您已经[配置了传输的源位置](#)。

### 使用DataSync控制台配置传输的目的地位置

1. 在配置目标位置页面上，选择创建新位置或为要将对象传输到的AWS存储资源选择现有位置。

如果您要创建新实例启动点，请参阅以下任一主题：

- [Amazon S3](#)
- [Amazon E](#)
- [FSx for Windows File Server](#)
- [FSx for Lustre](#)
- [FSx rrrrrrrrr](#)
- [适用于 ONTAP 的 FSx](#)

2. 配置完目标位置后，选择 Next ( 下一步 )。

## 配置您的传输设置

使用DataSync，您可以指定传输时间表、自定义验证数据完整性的方式，以及指定是否只传输对象子集等。

开始之前：确保您已经[配置了转移的目的地位置](#)。

使用DataSync控制台配置您的传输设置

1. 在配置设置页面上，更改传输设置或使用默认值。

有关这些设置的更多信息，请参阅[处理AWS DataSync转移任务](#)。

2. 选择下一步。

3. 查看传送详细信息，然后选择 Create task ( 创建任务 )。

## 开始传输

创建传送任务后，您就可以开始传输。有关使用DataSync控制台启动任务的说明，或者AWS CLI，请参阅[启动任务](#)。

## 使用以下AWS DataSync方式配置传输AWS Snowcone

要将数据传输到AWS Snowcone设备或从设备传输数据，必须创建AWS DataSync传输位置。DataSync可以将此位置用作传输数据的源或目标。

## 创建您的 Snowcone 传送地点

在开始之前，您需要在Snowcone设备上启用网络文件系统（NFS）。有关更多信息，请参阅 [AWS Snowcone 用户指南](#)。

### 使用控制台创建位置

1. 通过 <https://console.aws.amazon.com/datasync/> 打开AWS DataSync主机。
2. 在左侧导航窗格中，展开“数据传输”，然后选择“位置”和“创建位置”。
3. 对于位置类型，选择网络文件系统 (NFS)。
4. 对于代理，请选择您在 Snowcone 设备上启动的 Amazon EC2 代理。

有关代理的更多信息，请参阅 [《AWS Snowcone用户指南》](#)。

5. 对于 NFS 服务器，请输入您连接到 Snowcone 设备上的 NFS 服务器的虚拟 IP 地址。
6. 对于装载路径，输入要将数据传输到或从中传输数据的 S3 存储桶的 NFS 导出路径。

存储段的导出路径的格式为 `/buckets/bucket-name`。有关更多信息，请参阅 [《AWS Snowcone用户指南》](#) 中的“[使用 NFS 文件共享管理文件存储](#)”。

7. （可选）展开其他设置并选择特定 NFS 版本供DataSync访问文件服务器时使用。

默认情况下，DataSync使用 NFS 版本 4.1。DataSync还支持 NFS 3.x 和 4.0。

8. （可选）选择添加标签来标记您的位置。

标签是键值对，帮助您管理、筛选和搜索。我们建议至少为您的位置创建一个名称标签。

9. 选择“创建地点”。

## 如何AWS DataSync处理元数据和特殊文件

AWS DataSync在存储系统和服务之间复制数据时保存元数据和特殊文件（链接和目录）。

### 主题

- [元数据复制者 AWS DataSync](#)
- [复制的链接和目录 AWS DataSync](#)

## 元数据复制者 AWS DataSync

在传输过程中如何AWS DataSync处理文件或对象元数据取决于您使用的存储系统。

**Note**

DataSync不复制系统级设置。例如，在复制对象时，DataSync不会复制存储系统的加密设置。如果您从 SMB 共享进行复制，则DataSync不会复制您在文件系统级别配置的权限。

## 在具有相似元数据结构的系统之间复制的元数据

DataSync在具有相似元数据结构的存储系统之间保留元数据。

### 主题

- [NFS 传输](#)
- [中小型企业转账](#)
- [HDFS](#)
- [Amazon S3](#)

### NFS 传输

下表描述了在使用网络文件系统 (NFS) 的位置之间DataSync可以复制哪些元数据。

在这些位置之间复制时	DataSync可以复制
<ul style="list-style-type: none"> <li>• NFS</li> <li>• Amazon EFS</li> <li>• Amazon FSx for Lustre</li> <li>• Amazon FSx for OpenZFS</li> <li>• 适用于 NetApp ONTAP 的亚马逊 FSx ( 使用 NFS )</li> </ul>	<ul style="list-style-type: none"> <li>• 文件和文件夹修改时间戳</li> <li>• 文件和文件夹访问时间戳 ( DataSync只能在尽最大努力的基础上做到这一点 )</li> <li>• 用户 ID (UID) 和组 ID (GID)</li> <li>• POSIX 权限</li> </ul>

### 中小型企业转账

下表描述了在使用服务器消息块 (SMB) 的位置之间DataSync可以复制哪些元数据。

在这些位置之间复制时	DataSync可以复制
<ul style="list-style-type: none"> <li>• SMB</li> </ul>	<ul style="list-style-type: none"> <li>• 文件时间戳：访问时间、修改时间和创建时间</li> </ul>

在这些位置之间复制时	DataSync可以复制
<ul style="list-style-type: none"> <li>Amazon FSx for Windows File Server</li> <li>适用于 ONTAP 的 FSx ( 使用 SMB )</li> </ul>	<ul style="list-style-type: none"> <li>文件所有者安全标识符 (SID)</li> <li>标准文件属性：只读 (R)、存档 (A)、系统 (S)、隐藏 (H)、压缩 (C)、未索引内容 (I)、加密 (E)、临时 (T)、脱机 (O) 和稀疏文件 (P)</li> </ul> <p>DataSync尝试复制存档、压缩和稀疏属性。如果这些属性未应用于目的地，则在任务验证期间会被忽略。</p> <p>注意：适用于 ONTAP 的 FSx 不支持索引。传输到这种类型的文件系统时，DataSync会忽略“未编入内容索引 (I)”属性。</p> <ul style="list-style-type: none"> <li>NTFS 自由访问列表 ( DACL ) ，用于确定是否向对象授予访问权限。</li> <li>NTFS 系统访问控制列表 ( SACL ) ，管理员将使用这些列表记录对受保护对象的访问尝试。</li> </ul> <p>注意：如果您使用 SMB 版本 1.0，则不会复制 SACL。</p> <p>复制 DACL 需要向使用 SACL 需要向DataSync使用 SMB 的 Windows 用户授予其他权限。有关更多信息，请参阅为 <a href="#">SMB 创建位置</a>、<a href="#">FS x for Windows File Server</a> 或 <a href="#">ONTAP 的 FsX</a> ( 取决于传输中的位置类型 ) 。</p>

## HDFS

下表描述了当传输涉及 Hadoop Distributed File System ( HDFS ) 位置的。

从此位置复制时	到这些地点之一	DataSync可以复制
<ul style="list-style-type: none"> <li>HDFS</li> </ul>	<ul style="list-style-type: none"> <li>Amazon EFS</li> <li>FSx for Lustre</li> </ul>	<ul style="list-style-type: none"> <li>文件和文件夹修改时间戳</li> </ul>

从此位置复制时	到这些地点之一	DataSync可以复制
	<ul style="list-style-type: none"> <li>FSx</li> <li>适用于 ONTAP 的 FSx ( 使用 NFS )</li> </ul>	<ul style="list-style-type: none"> <li>文件和文件夹访问时间戳 ( DataSync只能在尽最大努力的基础上做到这一点 )</li> <li>POSIX 权限</li> </ul> <p>HDFS 使用字符串来存储文件和文件夹的用户和组所有权，而不是数字标识符 ( 例如 UID 和 GID )。UID 和 GID 的默认值应用于目标文件系统。有关默认值的更多信息，请参阅<a href="#">了解何时以及如何DataSync应用默认 POSIX 元数据</a>。</p>

## Amazon S3

下表描述了传输涉及 Amazon S3 位置时DataSync可以复制哪些元数据。

### 主题

- [到Amazon S3](#)
- [在 HDAmazon S3](#)
- [在对象存储和 Amazon S3 之间](#)

### 到Amazon S3

从这些位置之一进行复制时	到这个位置	DataSync可以复制
<ul style="list-style-type: none"> <li>NFS</li> <li>Amazon EFS</li> <li>FSx for Lustre</li> <li>FSx</li> <li>适用于 ONTAP 的 FSx ( 使用 NFS )</li> </ul>	<ul style="list-style-type: none"> <li>Amazon S3</li> </ul>	<p>以下是 Amazon S3 用户元数据：</p> <ul style="list-style-type: none"> <li>文件和文件夹修改时间戳</li> </ul>

从这些位置之一进行复制时	到这个位置	DataSync可以复制
		<ul style="list-style-type: none"> <li>• 文件和文件夹访问时间戳 ( DataSync只能在尽最大努力的基础上做到这一点 )</li> <li>• 用户 ID 和组 ID</li> <li>• POSIX 权限</li> </ul> <p>存储在 Amazon S3 用户元数据中的文件元数据可通过以下方式与文件网关上的 NFS 共享进行互操作。AWS Storage Gateway 文件网关支持从本地网络对复制到 Amazon S3 的数据进行低延迟访问。DataSync 此元数据还可以与 Lustre 的 FsX 互操作。</p> <p>当 DataSync 将包含此元数据的对象复制回 NFS 服务器时，系统将还原文件元数据。恢复元数据需要向 NFS 服务器授予更高的权限。有关更多信息，请参阅<a href="#">配置来自 NFS 文件服务器的 AWS DataSync 传输</a>：</p>

## 在 HDAmazon S3

在这些位置之间复制时	DataSync可以复制
<ul style="list-style-type: none"> <li>• Hadoop Distributed File System ( HDFS )</li> <li>• Amazon S3</li> </ul>	<p>以下是 Amazon S3 用户元数据：</p> <ul style="list-style-type: none"> <li>• 文件和文件夹修改时间戳</li> <li>• 文件和文件夹访问时间戳 ( DataSync只能在尽最大努力的基础上做到这一点 )</li> <li>• 用户 ID 和组 ID</li> </ul>



在这些位置之间复制时	DataSync可以复制
	<ul style="list-style-type: none"> <li>• POSIX 权限</li> </ul> <p>HDFS 使用字符串来存储文件和文件夹的用户和组所有权，而不是数字标识符，例如 UID 和 GID。</p>

### 在对象存储和 Amazon S3 之间

在这些位置之间复制时	DataSync可以复制
<ul style="list-style-type: none"> <li>• 对象存储</li> <li>• Amazon S3</li> </ul>	<ul style="list-style-type: none"> <li>• 用户定义的对象元数据</li> <li>• 对象标签</li> <li>• 以下系统定义的对象元数据： <ul style="list-style-type: none"> <li>• Content-Disposition</li> <li>• Content-Encoding</li> <li>• Content-Language</li> <li>• Content-Type</li> </ul> </li> </ul> <p>注意：在初始传输期间DataSync复制所有对象的系统元数据。如果您<a href="#">将任务配置为仅传输已更改的数据</a>，则除非对象的内容或用户元数据也已修改，否则DataSync不会在后续传输中复制系统元数据。</p> <p>DataSync不复制其他对象元数据，例如对象访问控制列表 (ACL) 或以前的对象版本。</p> <p>重要提示：如果您要从 Google Cloud Storage 存储分区转移对象，则复制对象标签可能会导致 DataSync 任务失败。为防止出现这种情况，请在配置任务设置时取消选择“复制对象标签”选项。有关更多信息，请参阅<a href="#">管理文件、对象和元数据的AWS DataSync传输方式</a>：</p>

## 在具有不同元数据结构的系统之间复制的元数据

在不具有相似元数据结构的存储系统之间进行复制时，使用以下规则DataSync处理元数据。

从这些位置之一进行复制时	到这些地点之一	DataSync可以复制
<ul style="list-style-type: none"> <li>• SMB</li> <li>• FSx for Windows File Server</li> <li>• 适用于 ONTAP 的 FSx ( 使用 SMB )</li> </ul>	<ul style="list-style-type: none"> <li>• Amazon EFS</li> <li>• FSx for Lustre</li> <li>• FSx</li> <li>• 适用于 ONTAP 的 FSx ( 使用 NFS )</li> <li>• Amazon S3</li> <li>• 对象存储</li> <li>• NFS</li> </ul>	<p>目标文件系统上的所有文件和文件夹或目标 S3 存储桶中对象的@@ <a href="#">默认 POSIX 元数据</a>。该方法包括使用默认 POSIX 用户 ID 和组 ID 值。</p> <p>不保留基于 Windows 的元数据 ( 例如 ACL )。</p>
<ul style="list-style-type: none"> <li>• FSx for Windows File Server</li> <li>• 适用于 ONTAP 的 FSx ( 使用 SMB )</li> </ul>	<ul style="list-style-type: none"> <li>• HDFS</li> </ul>	<p>源位置的文件和文件夹时间戳。文件或文件夹所有者是根据您在创建 HDFS 位置时指定的 HDFS 用户或 Kerberos 主体设置的。Hadoop 集群上的组映射配置决定了群组。</p>
<ul style="list-style-type: none"> <li>• Amazon EFS</li> <li>• FSx for Lustre</li> <li>• FSx</li> <li>• 适用于 ONTAP 的 FSx ( 使用 NFS )</li> <li>• Amazon S3</li> <li>• 对象存储</li> <li>• NFS</li> <li>• HDFS</li> </ul>	<ul style="list-style-type: none"> <li>• SMB</li> <li>• FSx for Windows File Server</li> <li>• 适用于 ONTAP 的 FSx ( 使用 SMB )</li> </ul>	<p>源位置的文件和文件夹时间戳。所有权是根据在中DataSync指定访问亚马逊 FSx 或 SMB 共享的 Windows 用户设置的。权限继承自父目录。</p>

## 了解何时以及如何DataSync应用默认 POSIX 元数据

DataSync在以下情况下应用默认 POSIX 元数据：

- 当您的转账的来源和目标位置没有相似的元数据结构时
- 当源位置缺少元数据时

下表描述了在这些类型的传输过程中如何DataSync应用默认 POSIX 元数据：

源	目标位置	文件权限	文件夹权限	UID	GID
<ul style="list-style-type: none"> <li>• Amazon S<sub>3</sub></li> <li>• 对象存储<sup>1</sup></li> </ul>	<ul style="list-style-type: none"> <li>• Amazon EFS</li> <li>• FSx for Lustre</li> <li>• FSx</li> <li>• 适用于 ONTAP 的 FSx ( 使用 NFS )</li> <li>• NFS</li> </ul>	0755	0755	65534	65534
<ul style="list-style-type: none"> <li>• SMB</li> </ul>	<ul style="list-style-type: none"> <li>• Amazon S3</li> <li>• 对象存储</li> <li>• Amazon EFS</li> <li>• FSx for Lustre</li> <li>• FSx</li> <li>• 适用于 ONTAP 的 FSx ( 使用 NFS )</li> <li>• NFS</li> </ul>	0644	0755	0	0
<ul style="list-style-type: none"> <li>• HDFS</li> </ul>	<ul style="list-style-type: none"> <li>• Amazon EFS</li> <li>• FSx for Lustre</li> </ul>	0644	0755	0	0

源	目标位置	文件权限	文件夹权限	UID	GID
	<ul style="list-style-type: none"> <li>FSx</li> <li>适用于 ONTAP 的 FSx ( 使用 NFS )</li> <li>NFS</li> </ul>				

<sup>1</sup> 如果对象没有先前应用过的元数据DataSync。

### Note

HDFS 使用字符串而不是数字标识符 ( 例如 UID 和 GID ) 存储文件和文件夹的用户和组所有权。如果源位置上没有等效的元数据，则文件和文件夹所有权将根据您在创建位置时指定的 HDFS 用户或 Kerberos 主体进行设置。DataSyncHadoop 集群上的组映射配置决定了群组。

## 复制的链接和目录 AWS DataSync

AWS DataSync根据传输所涉及的存储位置，对复制的硬链接、符号链接和目录的处理方式有所不同。

### 硬链接

在 NFS 服务器、FSx for Lustre、OpenZFS 的 FsX、ONTAP 的 fsX ( 使用 NFS ) 和亚马逊 EFS 之间进行复制时，会保留硬链接。

复制到 Amazon S3 时，每个硬链接仅传输一次。为每个副本创建单独的 Amazon S3 对象。如果硬链接在 Amazon S3 中保持不变，则在传输到 NFS 服务器、FSx for Lustre、OpenZFS 的 FsX、ONTAP 的 FsX ( 使用 NFS ) 或亚马逊 EFS 时会正确恢复。

在 SMB 文件共享、FSx for Windows File Server 和 ONTAP 的 FsX ( 使用 SMB ) 之间进行复制时，不支持硬链接。如果在这些情况下DataSync遇到硬链接，则任务完成时会出现错误。要了解更多信息，请查看您的CloudWatch日志。

复制到 HDFS 时，不支持硬链接。复制到 HDFS 时，会跳过源代码上的硬链接并登录到。CloudWatch

## 符号链接

在 NFS 服务器、FSx for Lustre、OpenZFS 的 FsX、ONTAP 的 fsX ( 使用 NFS ) 和亚马逊 EFS 之间进行复制时，会保留符号链接。

复制到 Amazon S3 时，链接目标路径存储在 Amazon S3 对象中。传输到 NFS 服务器、Lustre 的 fsX、OpenZFS 的 FsX、ONTAP 的 fsX 或 Amazon EFS 后，链接会正确恢复。

在 SMB 文件共享、FSx for Windows File Server 和 ONTAP 的 FsX ( 使用 NFS ) 之间进行复制时，不支持符号链接。如果在这些情况下 DataSync 遇到符号链接，则任务完成时会出现错误。要了解更多信息，请查看您的 CloudWatch 日志。

复制到 HDFS 时，不支持符号链接。复制到 HDFS 时，会跳过符号链接并记录到。CloudWatch

## 目录

复制到 Amazon S3 存储桶或从 Amazon S3 存储桶复制时，目录表示为以结尾的空对象/。

有关登录的信息 DataSync，请参见 [监控亚马逊的 AWS DataSync 活动 CloudWatch](#)。

## 移除传 AWS DataSync 送地点

最佳做法是移除不再需要 AWS DataSync 的地点。

使用 DataSync 控制台删除地点

1. 通过 <https://console.aws.amazon.com/datasync/> 打开 AWS DataSync 主机。
2. 在左侧导航窗格中，展开“数据传输”，然后选择“位置”。
3. 选择要删除的位置。
4. 选择删除。输入确认删除 **delete**，然后选择 Delete。

## 处理 AWS DataSync 转移任务

一项任务描述了在何处以及如何 AWS DataSync 传输数据。任务由以下内容组成：

- [源位置](#)-DataSync 传输数据的存储系统或服务。
- [目标位置](#)-将数据 DataSync 传输到的存储系统或服务。

- [任务设置](#)-用于配置任务行为方式的选项，例如如何验证数据、何时运行等。某些任务设置是可选的。例如，您不必为任务命名。
- [任务执行](#)-当你运行任务时，它被称为任务执行。

## 主题

- [配置如何AWS DataSync验证数据完整性](#)
- [配置哪些AWS DataSync传输](#)
- [为您的AWS DataSync任务设置带宽限制](#)
- [安排AWS DataSync任务](#)
- [记录您的AWS DataSync任务活动](#)
- [为你的AWS DataSync任务添加标签](#)
- [开始你的AWS DataSync任务](#)
- [取消你的任务 AWS DataSync](#)
- [删除你的AWS DataSync任务](#)
- [AWS DataSync任务状态](#)
- [AWS DataSync任务场景示例](#)

## 配置如何AWS DataSync验证数据完整性

在传输过程中，AWS DataSync始终检查数据的完整性，但您可以使用以下选项指定进行验证的方式和时间：

- 仅验证传输的数据（推荐）-DataSync 计算源位置传输的文件和元数据的校验和。在传输结束时，将DataSync此校验和与在目标位置的这些文件上计算的校验和进行比较。

在传输到 S3 Glacier Retrieval 或 S3 Glacier Deep Archive（S3）有关更多信息，请参阅[Amazon S3 传输的存储类注意事项](#)：

- 验证目标中的所有数据 — 在传输结束时，DataSync扫描整个源和目标，以验证两个位置是否完全同步。

在传输到 S3 Glacier Flexible Retrieval 或 S3 Glacier Deep Archive（有关更多信息，请参阅[Amazon S3 传输的存储类注意事项](#)）：

- 在传输期间检查完整性-在传输结束时DataSync不进行额外验证。在传输过程中，仍将通过校验和验证来检查所有数据传输的完整性。

## 使用 DataSync 控制台

以下说明介绍了如何在创建任务时配置数据验证。您也可以在编辑任务或开始执行任务时对此进行配置。

### 使用控制台配置数据验证

1. 通过 <https://console.aws.amazon.com/datasync/> 打开AWS DataSync主机。
2. 在左侧导航窗格中，展开“数据传输”，然后选择“任务”，然后选择“创建任务”。
3. 配置任务的源位置和目标位置。

有关更多信息，请参阅[我可以哪里传输我的数据AWS DataSync？](#)：

4. 对于验证数据，请选择下列选项之一：
  - 仅验证传输的数据（推荐）
  - 验证目的地中的所有数据
  - 在传输过程中检查完整性

## 使用 DataSync API

您可以通过以下操作配置如何DataSync验证数据：

- [CreateTask](#)
- [UpdateTask](#)
- [StartTaskExecution](#)

## 配置哪些AWS DataSync传输

您可以将AWS DataSync任务配置为仅复制所需的数据。例如，排除某些文件夹或仅包含特定的文件类型。如果您计划定期传输，则可以将任务设置为仅传输源位置发生变化的数据。

### 主题

- [管理文件、对象和元数据的AWS DataSync传输方式](#)
- [筛选传输的数据AWS DataSync](#)

## 管理文件、对象和元数据的AWS DataSync传输方式

您可以配置在不同位置之间复制时如何AWS DataSync处理文件、对象及其关联元数据。

例如，对于定期传输，您可能需要用源中的更改覆盖目标中的文件，以保持位置同步。您可以复制诸如文件和文件夹的 POSIX 权限、与对象关联的标签和访问控制列表 (ACL) 之类的属性。

开始之前：DataSync可以在传输中保留的元数据取决于所涉及的存储系统以及这些系统是否使用类似的元数据结构。在配置任务之前，请务必了解[如何DataSync处理元数据和特殊文件](#)。

### 使用 DataSync 控制台

以下说明介绍了如何在创建任务时配置DataSync传输内容。您也可以在编辑或启动任务时对此进行配置。

1. [通过 https://console.aws.amazon.com/datasync/](https://console.aws.amazon.com/datasync/) 打开AWS DataSync主机。
2. 在左侧导航窗格中，展开“数据传输”，然后选择“任务”，然后选择“创建任务”。
3. 配置任务的源位置和目标位置。

有关更多信息，请参阅[我可以哪里传输我的数据AWS DataSync？](#)：

4. 对于传输模式，请选择下列选项之一：
  - 仅传输已更改的数据-仅DataSync复制源位置和目标位置之间不同的数据和元数据。
  - 传输所有数据-DataSync 将源中的所有内容复制到目标，而不比较位置之间的差异。
5. 如果您DataSync要在目标位置维护源中不存在的文件或对象，请选择“保留已删除的文件”。

如果您不选择此选项，则您的任务删除您的 Amazon S3 桶中的对象，则您可能需要为某些存储类别支付最低存储期限费用。有关详细信息，请参阅[Amazon S3 传输的存储类注意事项](#)。

#### Warning

您无法取消选择此选项并启用“传输所有数据”。当你传输所有数据时，DataSync不会扫描你的目标位置，也不知道要删除什么。

6. 如果DataSync要在源数据或元数据发生变化时修改目标位置的数据，请选择“覆盖文件”。

如果您的任务覆盖对象，则您可能需要为某些存储类别支付额外费用（例如，用于检索或提前删除）。有关详细信息，请参阅[Amazon S3 传输的存储类注意事项](#)。

如果您不选择此选项，则即使源数据不同，目标数据也不会被覆盖。



## 7. 在数据传输配置下，选择DataSync要如何复制元数据。

### Important

您在控制台中看到的设置取决于任务的源位置和目标位置。在某些情况下，您可能需要展开“其他设置”才能看到这些选项。

在做出选择之前，请务必了解[DataSync将如何处理您的场景的元数据和特殊文件](#)。

设置	描述
复制所有权	DataSync复制 POSIX 文件和文件夹所有权，例如文件所有者的群组 ID 和文件所有者的用户 ID。
复制权限	DataSync将文件和文件夹的 POSIX 权限从源复制到目标。
复制时间戳	DataSync将时间戳元数据从源复制到目标。
复制对象标签	DataSync在对象存储系统之间传输时保留与您的对象相关的标签。

设置	描述
复制所有权、DACL 和 SACL	<p>DataSync复制以下内容：</p> <ul style="list-style-type: none"> <li>• 对象所有者。</li> <li>• NTFS 自由访问列表 ( DACL ) ，用于确定是否向对象授予访问权限。</li> <li>• NTFS 系统访问控制列表 ( SACL ) ，管理员将使用这些列表记录对受保护对象的访问尝试。</li> </ul> <p>注意：如果您使用 SMB 版本 1.0 ，则不会复制 SACL。</p> <p>复制 DACL 和 SACL 需要向DataSync用于使用 SMB 访问您的位置的 Windows 用户授予特定权限。有关更多信息，请参阅为 <a href="#">SMB 创建位置</a>、<a href="#">FS x for Windows File Server</a> 或 <a href="#">ONTAP 的 FsX</a> ( 取决于传输中的位置类型 ) 。</p>
复制所有权和 DACL	<p>DataSync复制以下内容：</p> <ul style="list-style-type: none"> <li>• 对象所有者。</li> <li>• DACL ，用于确定是否向对象授予访问权限。</li> </ul> <p>DataSync选择此选项时不会复制 SACL。</p>
不要复制所有权或 ACL	<p>DataSync不复制任何所有权或权限数据。DataSync写入目标位置的对象归用户所有，该用户需要提供DataSync访问目标的凭证。目标对象权限是根据目标服务器上配置的权限确定的。</p>

## 使用 DataSync API

您可以通过使用Options参数进行以下任何操作来配置这些任务设置：

- [CreateTask](#)
- [StartTaskExecution](#)
- [UpdateTask](#)

## 筛选传输的数据AWS DataSync

AWS DataSync如果您只想传输数据子集（例如特定文件、文件夹或对象），则允许您应用过滤器。

例如，如果您的源位置包含以结尾的临时文件.tmp，则可以创建排除过滤器，防止这些文件进入目标位置。您还可以在同一任务中组合使用排除和包含过滤器。

### 主题

- [筛选术语、定义和语法](#)
- [从传输中排除数据](#)
- [在传输中包括数据](#)
- [过滤器示例](#)

### 筛选术语、定义和语法

以下是一些用于筛选的术语和定义：

#### 筛选条件

构成特定过滤器的整个字符串（例如：`*.tmp|*.temp或/folderA|/folderB`）。

过滤器由以|（管道）分隔的模式组成。在控制台上添加模式时不需要分隔符，因为每个模式都是单独添加的。

#### Note

过滤器区分大小写。例如，过滤器将/folderA不匹配/FolderA。

## Pattern

筛选条件中的模式。例如，\*.tmp是作为\*.tmp|\*.temp过滤器一部分的模式。

### 文件夹

- 所有筛选条件都相对于源位置路径。例如，假设您在创建源位置和任务时指定/my\_source/为源路径并指定包含过滤器/transfer\_this/。在这种情况下，仅DataSync传输目录/my\_source/transfer\_this/及其内容。
- 要直接在源位置下指定文件夹，请在文件夹名称前加上正斜杠 (/)。在上述示例中，模式使用 /transfer\_this，而不是 transfer\_this。
- DataSync以相同的方式解释以下模式并匹配文件夹及其内容。


/dir

/dir/

- 当您从 Amazon S3 存储桶传输数据或向 Amazon S3 存储桶传输数据时，会将对象密钥中的/字符DataSync视为等效文件系统上的文件夹。

### 特殊字符

以下是用于筛选的特殊字符。

特殊字符	描述
* (通配符)	一个用于匹配零个或多个字符的字符。例如，/movies_folder* 匹配 /movies_folder 和 /movies_folder1 这两者。
(竖线分隔符)	用作模式之间分隔符的字符。它支持指定多个模式，其中的任何模式都可以与筛选条件匹配。例如，*.tmp *.temp 匹配以 tmp 或 temp 结尾的文件。  <div data-bbox="625 1596 1510 1816" style="border: 1px solid #0070C0; border-radius: 10px; padding: 10px;"> <p> <b>Note</b> 当您在控制台上添加模式时，不需要这个分隔符，因为您在单独的行中添加了每个模式。</p> </div>
\ (反斜杠)	用于转义文件或对象名称中的特殊字符 (*、 、\ ) 的字符。

特殊字符	描述
	当反斜杠是文件名的组成部分时，需要使用双反斜杠 (\)。同样，\\ 代表文件名中两个连续的反斜杠。
	当竖线是文件名的组成部分时，需要使用反斜杠后跟竖线 ( ) 字符。
	忽略后面加上任何其他字符或模式末尾的反斜杠 (\)。

## 从传输中排除数据

**排除筛选条件** 定义在将文件从源复制到目标位置时要排除的文件、文件夹和对象。您可以在创建、编辑或启动任务时配置这些筛选器。

要在DataSync控制台使用排除筛选器创建任务，请在排除模式下的“数据传输配置”部分中指定模式列表。例如，要排除名为temp或的临时文件夹tmp，可以在排除模式文本框\*/temp中指定，选择添加模式，然后在第二个文本框\*/tmp中指定。要向筛选条件添加更多模式，请选择 Add pattern (添加模式)。使用AWS Command Line Interface (AWS CLI) 时，过滤器周围需要使用单引号 (')，并使用 | (管道) 作为分隔符。在本示例中，应指定 '\* /temp | \* /tmp'。

创建任务后，您可以编辑任务配置以在排除过滤器中添加或删除模式。您的更改将应用于future 执行该任务。

运行任务时，可以使用“以覆盖模式开始”选项修改排除筛选模式。您所做的任何更改仅应用于该任务的执行。

您还可以使用AWS CLI，创建或编辑排除过滤器。以下示例显示这样一个 CLI 命令。

```
aws datasync create-task
  --source-location-arn 'arn:aws:datasync:region:account-id:location/location-id'
  --destination-location-arn 'arn:aws:datasync:region:account-id:location/location-id'
  --cloud-watch-log-group-arn 'arn:aws:logs:region:account-id:log-group:your-log-group'
  --name your-task-name
  --excludes FilterType=SIMPLE_PATTERN,Value='*/temp|*/tmp'
```

**Note**

如果您要从NetApp系统迁移文件，我们建议您通过在排除筛选器中指定\*/.snapshot为模式来排除NetApp备份文件夹。

## 在传输中包括数据

包括筛选器定义运行任务时DataSync传输的文件、文件夹和对象。可以在创建、编辑或启动任务时配置包含过滤器。

要使用包含过滤器创建任务，请选择“特定文件和文件夹”选项，然后在“包含模式”下指定要包含的模式列表。

DataSync仅扫描和传输与包含过滤器匹配的文件和文件夹。例如，要包含源文件夹的子集，可以指定/`important_folder_1` |/`important_folder_2`。

创建任务后，您可以编辑任务配置以在包含过滤器中添加或删除模式。您所做的任何更改都将应用于该任务的future 执行。

运行任务时，可以使用“以覆盖模式开始”选项修改包含筛选模式。您所做的任何更改仅应用于该任务的执行。

您还可以使用 cludAWS CLI e 过滤器创建或编辑包含过滤器。以下示例显示了此 CLI 命令。注意过滤器周围的引号 (') 和用作分隔符的 | (管道)。

```
aws datasync start-task-execution
  --task-arn 'arn:aws:datasync:region:account-id:task/task-id'
  --includes FilterType=SIMPLE_PATTERN,Value='/important_folder1|/important_folder2'
```

**Note**

Include 过滤器仅支持将通配符 (\*) 作为模式中最右边的字符。例如，支持/`documents*` |/`code*`，但`*.txt`不支持。

## 过滤器示例

以下示例说明您可以与结合使用的常用过滤器DataSync。

**Note**

在过滤器中可以使用的字符数有限制。有关更多信息，请参阅[DataSync任务配额](#)：

### 排除源位置中的某些文件夹

在某些情况下，您可以排除源位置中的文件夹，以免将它们复制到目标位置。例如，您可能有临时 work-in-progress 文件夹。或者，您可能使用 NetApp 系统并希望排除备份文件夹。在这些情况下，您可以使用以下过滤器。

```
*/.snapshot
```

要排除内容相似（例如 /reports2021 和 /reports2022）的文件夹，可以使用如下所示的排除过滤器：

```
/reports*
```

要排除文件层次结构中任何级别的文件夹，可以使用如下所示的排除过滤器。

```
*/folder-to-exclude-1|*/folder-to-exclude-2
```

要排除源位置顶层的文件夹，可以使用如下所示的排除过滤器。

```
/top-level-folder-to-exclude-1|top-level-folder-to-exclude-2
```

### 包括源位置的文件夹子集

在某些情况下，您的源位置可能会占很大的份额，您需要传输根目录下的一部分文件夹。要包含特定文件夹，使用如下所示的包含筛选条件开始任务执行。

```
/folder-to-transfer/*
```

### 排除特定文件类型

要从传输中排除某些文件类型，您可以使用诸如 \*.temp 的排除筛选条件创建一个任务执行。

### 传输您指定的单个文件

要传输单个文件的列表，请使用包含过滤器开始执行任务，如下所示：“/folder/subfolder/file1.txt|/folder/subfolder/file2.txt |/folder/subfolder/file2.txt”

## 为您的AWS DataSync任务设置带宽限制

您可以为AWS DataSync任务及其每次执行配置网络带宽限制。

### 限制任务的带宽

在创建、编辑或启动任务时设置带宽限制。

#### 使用 DataSync 控制台

以下说明描述了如何在创建任务时为任务配置带宽限制。

1. 通过 <https://console.aws.amazon.com/datasync/> 打开AWS DataSync主机。
2. 在左侧导航窗格中，展开“数据传输”，然后选择“任务”，然后选择“创建任务”。
3. 配置任务的源位置和目标位置。

有关更多信息，请参阅[我可以哪里传输我的数据AWS DataSync？](#)：

4. 对于设置带宽限制，请选择下列选项之一：
  - 选择 Use ( 使用可用 ) 以使用所有可用的网络带宽以执行任务。
  - 选择设置带宽限制 (miB/s)，然后输入每次执行任务时DataSync要使用的最大带宽。

#### 使用 DataSync API

您可以通过将BytesPerSecond参数与以下任何操作一起使用来配置任务的带宽限制：

- [CreateTask](#)
- [UpdateTask](#)
- [StartTaskExecution](#)

### 限制任务执行的带宽

您可以修改正在运行或排队执行的任务的带宽限制。

#### 使用 DataSync 控制台

1. 通过 <https://console.aws.amazon.com/datasync/> 打开AWS DataSync主机。
2. 在导航窗格中，展开“数据传输”。然后选择“任务”。
3. 选择任务，然后选择“历史记录”以查看任务的执行情况。



4. 选择要修改的任务执行，然后选择 Edisting ( 编辑 )。
5. 在此对话框中，选择下列选项之一：
  - 选择 Use ( 使用可用 ) 以使用所有可用的网络带宽以执行任务。
  - 选择设置带宽限制 (miB/s)，然后输入DataSync要用于执行任务的最大带宽。
6. 选择保存更改。

新的带宽限制将在 60 秒内生效。

## 使用 DataSync API

您可以通过在[UpdateTaskExecution](#)操作中使用BytesPerSecond参数来修改正在运行或排队执行的任务的带宽限制。

## 安排AWS DataSync任务

您可以将AWS DataSync任务设置为在不同地点之间定期传输数据。

计划任务自动按照您配置的频率运行，最小间隔为 1 小时。例如，以下屏幕截图显示了一个配置，该配置每个星期日和星期三在 UTC 时间晚上 12:00 运行任务。

### Schedule - optional

Execute this task on a schedule by specifying the time and frequency.

Frequency

Days of the week ▼

on

Sun  Mon  Tue  Wed  Thu  Fri  Sat

at

12:00

您也可以使用 UTC 时间指定的 cron 表达式执行任务计划。例如，使用以下 cron 表达式将任务配置为在每个星期日和星期三下午 12:00 运行。

```
0 12 ? * SUN,WED *
```


### Schedule - optional

Execute this task on a schedule by specifying the time and frequency.

Frequency

Custom ▼

Cron expression

Use a cron expression for complex schedules. [Learn more and see syntax examples](#) 

0 12 ? \* SUN,WED \*

#### Important

即使使用 cron 表达式，您也无法将任务安排为以超过 1 小时的间隔运行。

有关调度表达式语法的详细信息，请参阅 [《亚马逊CloudWatch用户指南》中的规则调度表达式](#)。

## 配置任务时间表

您可以使用 DataSync 控制台或 API 配置任务执行的频率。在您创建或编辑任务时，在控制台的 Frequency (频率) 中可使用以下选项：

- 如果您不希望计划任务定期运行，请选择 Not Scheduled (未计划)。
- 选择 Hourly (每小时)，然后选择任务应在每小时的多少分钟开始运行。任务每小时在指定的分钟运行一次。
- 选择每天，然后输入您希望任务运行的 UTC 时间，格式为 HH: MM。此任务每天在指定时间运行。
- 选择每周和星期几并输入任务应运行的 UTC 时间，格式为 HH: MM。此任务每周在指定日期的指定时间运行。
- 选择一周中的几天，选择一个或多个特定的日期，然后以 HH: MM 格式输入任务应运行的 UTC 时间。此任务在您指定的日期和时间运行。
- 如果要使用自定义 cron 表达式运行任务，最小间隔为 1 小时，请选择“自定义”。然后在 Cron expression (Cron 表达式) 框中输入您的表达式。

有关调度表达式的详细信息，请参阅 [《亚马逊CloudWatch用户指南》中的规则调度表达式](#)。

## 更改任务时间表

您可以在[初始创建任务](#)时配置计划，也可以在创建任务后编辑任务计划。在您创建任务后，请使用以下过程配置计划。

### 更改任务时间表

1. 通过 <https://console.aws.amazon.com/datasync/> 打开AWS DataSync主机。
2. 在导航窗格上，展开“数据传输”，然后选择“任务”。
3. 选择要更改的任务。
4. 对于“操作”，选择“编辑”打开“编辑任务”页面并展开“计划”（可选）。
5. 在计划（可选）部分中，将您的任务配置为按新的计划运行。
6. 对于频率，配置您希望任务运行的频率，最小间隔为 1 小时。

有关频率配置选项，请参阅[配置任务时间表](#)。

## 记录您的AWS DataSync任务活动

您可以将有关您的AWS DataSync任务的详细信息发布到 Amazon Lo CloudWatch gs 日志组。

### 所需的权限

DataSync需要权限才能将日志上传到CloudWatch日志组。您可以通过几种不同的方式通过 IAM 资源策略进行设置：

- 当您使用控制台创建任务时，DataSync可以为您创建日志组和关联的资源策略。
- 如果您想使用现有的日志组，请查看[如何自己创建资源策略的示例](#)。

### 配置任务日志

您可以使用控制台或 DataSync API 为您的任务设置CloudWatch日志记录。

#### 使用 DataSync 控制台

以下说明介绍了如何在创建任务时配置CloudWatch日志记录。您也可以在编辑任务时对此进行配置。

1. 通过 <https://console.aws.amazon.com/datasync/> 打开AWS DataSync主机。

2. 在左侧导航窗格中，展开“数据传输”，然后选择“任务”，然后选择“创建任务”。
3. 配置任务的源位置和目标位置。

有关更多信息，请参阅[我可以哪里传输我的数据AWS DataSync？](#)：

4. 对于 Log ( 日志级别 ) ，请选择下列选项之一：
  - 记录基本信息，例如传输错误-仅将基本信息 ( 例如传输错误 ) 发布到CloudWatch。
  - 记录所有传输的对象和文件-将任务复制和完整性检查的所有文件或对象的日志记录发布到CloudWatch日志。
  - 不要将日志发送到 CloudWatch
5. 对于CloudWatch日志组，通过执行以下操作之一指定DataSync有权将日志上传到的日志组：
  - 选择自动生成可自动创建允许DataSync向其上传日志的日志组。
  - 选择您当前日志组中的现有日志组AWS 区域。

确保您的[资源策略](#)允许将日志上传DataSync到日志组。

## 使用 DataSync API

您可以使用CloudWatchLogGroupArn参数进行以下任何操作，为您的任务配置CloudWatch日志记录：

- [CreateTask](#)
- [UpdateTask](#)

## 为你的AWS DataSync任务添加标签

标签是键值对，帮助您管理、筛选和搜索AWS DataSync资源。您最多可以为每个DataSync任务和任务执行添加 50 个标签。

例如，您可以为大型数据迁移创建任务，并使用密钥**Project**和值标记该任务**Large Migration**。要进一步组织迁移，您可以使用键**Transfer Date**和值标记一次运行的任务**May 2021** ( 后续任务执行可能会被标记**June 2021****July 2021**、等等 ) 。

## 为任务添加标签

您只能在创建DataSync任务时为任务添加标签。

## 使用 DataSync 控制台

1. 通过 <https://console.aws.amazon.com/datasync/> 打开AWS DataSync主机。
2. 在左侧导航窗格中，展开“数据传输”，然后选择“任务”，然后选择“创建任务”。
3. 配置任务的源位置和目标位置。

有关更多信息，请参阅[我可以哪里传输我的数据AWS DataSync？](#)：

4. 在配置设置页面上，选择添加标签来标记您的任务。

## 使用 AWS CLI

1. 复制以下create-task命令：

```
aws datasync create-task \  
    --source-location-arn 'arn:aws:datasync:region:account-id:location/source-  
location-id' \  
    --destination-location-arn 'arn:aws:datasync:region:account-  
id:location/destination-location-id' \  
    --tags Key=tag-key,Value=tag-value
```

2. 在此命令中指定以下参数：
  - --source-location-arn— 指定传输中源位置的 Amazon 资源名称 (ARN)。
  - --destination-location-arn— 在传输中指定目标位置的 ARN。
  - --tags— 指定要应用于任务的标签。

对于多个标签，用空格分隔每个键值对。

3. (可选) 指定对您的传输方案有意义的其他参数。

有关列表，请参见 [create-task 命令](#)。--options

4. 运行 create-task 命令。

您会收到显示您刚创建的任务的响应。

```
{  
    "TaskArn": "arn:aws:datasync:us-east-2:123456789012:task/task-  
abcdef01234567890"  
}
```

要查看您添加到此任务的标签，可以使用[list-tags-for-resource](#)命令。

## 标记任务执行

您可以为DataSync任务的每次运行添加标签。

如果您的任务已经有标签，请记住以下关于在任务执行中使用标签的注意事项：

- 如果您使用控制台启动任务，则其用户创建的标签将自动应用于任务执行。但是，系统创建的以开头的标签aws:不适用。
- 如果您使用 DataSync API 或启动任务AWS CLI，则其标签不会自动应用于任务执行。

### 使用 DataSync 控制台

要在任务执行中添加、编辑或删除标签，必须使用重写选项启动任务。

1. 通过 <https://console.aws.amazon.com/datasync/> 打开AWS DataSync主机。
2. 在左侧导航窗格中，展开 Data Transs，然后选择 Tasks ( 任务 )。
3. 选择任务。
4. 选择 Start ( 开始 )，然后选择下列选项之一：
  - 从@@ 默认值开始-应用与您的任务相关的任何标签。
  - 从@@ 重写选项开始-允许您为执行此特定任务添加、编辑或删除标签。

### 使用 AWS CLI

1. 复制以下start-task-execution命令：

```
aws datasync start-task-execution \  
  --task-arn 'arn:aws:datsync:region:account-id:task/task-id' \  
  --tags Key=tag-key,Value=tag-value
```

2. 在此命令中指定以下参数：

- --task-arn— 指定要启动的任务的 ARN。
- --tags— 指定要应用于此任务的特定运行的标签。

对于多个标签，用空格分隔每个键值对。

3. (可选) 指定对您的情况有意义的其他参数。

有关更多信息，请参阅 [start-task-execution](#) 命令。

4. 运行 `start-task-execution` 命令。

你会得到一个响应，显示你刚刚开始的任务执行情况。

```
{
  "TaskExecutionArn": "arn:aws:datasync:us-east-2:123456789012:task/task-
  abcdef01234567890"
}
```

要查看您添加到此任务的标签，可以使用 [list-tags-for-resource](#) 命令。

## 开始你的AWS DataSync任务

创建AWS DataSync传输任务后，即可开始移动数据。任务的每次运行都称为任务执行。

启动任务时，DataSync执行递归目录列表以发现传输源和目标位置中的所有文件、对象和关联元数据。DataSync识别位置之间的差异并确定要复制的内容，该过程通常需要几分钟到几个小时。

### Important

如果您计划向 Amazon S3 地点传输数据或从 Amazon S3 站点传输数据，请在开始之前查看 [DataSync如何影响您的 S3 请求费用](#) 和 [DataSync定价页面](#)。

## 运行多个任务

在某些情况下，您可能需要运行多个DataSync任务。例如，你可能会定期复制数据，或者以增量方式迁移大量文件（你可以通过[任务筛选来实现](#)）。

请记住以下有关运行多个任务的内容：

- 任务按顺序运行（先进先出）。如果您多次运行同一个任务，则默认情况下，每次执行都会进入队列。例如，如果你两次启动一个任务，那么第二次任务的执行要等到第一个任务完成后才会开始。
- 您可以同时运行不同的任务，但是如果这些任务使用的是同一个DataSync代理，则这些任务将被排入队列。

## 开始任务

创建任务后，您可以立即开始移动数据。

### 使用 DataSync 控制台

1. 通过 <https://console.aws.amazon.com/datasync/> 打开AWS DataSync主机。
2. 在左侧导航窗格中，展开 Data Transs，然后选择 Tasks（任务）。
3. 选择要运行的任务。

确保任务的状态为“可用”。您也可以选择多个任务。

4. 选择 Acistons（操作），然后选择下列选项之一：
  - 开始-运行任务（如果选择了多个，则运行任务）。
  - 从@@ 重写选项开始-允许您在开始移动数据之前修改某些任务设置。准备就绪后，选择 Start（开始）。
5. 选择“查看执行详细信息”以查看有关正在运行的任务执行的详细信息。

### 使用 DataSync API

您可以使用[StartTaskExecution](#)操作启动任务。使用[DescribeTaskExecution](#)操作获取有关正在运行的任务执行的详细信息。

启动后，您可以在DataSync移动数据时[检查任务的执行状态](#)。如果需要，您还可以[限制任务执行的带宽](#)。

## 取消你的任务 AWS DataSync

您可以停止任何正在运行或排队的AWS DataSync任务。

### 使用控制台取消任务

1. 通过 <https://console.aws.amazon.com/datasync/> 打开AWS DataSync主机。
2. 在左侧导航窗格中，展开 Data Transs，然后选择 Tasks（任务）。
3. 选择要监控的正在运行的任务的 Task ID。

任务状态应为“正在运行”。

4. 选择“历史记录”以查看任务的执行情况。



5. 选择要停止的任务执行，然后选择“停止”。
6. 在此对话框中，选择 Stop ( 停止 )。

要使用 DataSync API 取消正在运行或排队的任务，请参阅[CancelTaskExecution](#)。

## 删除你的AWS DataSync任务

如果不再需要某AWS DataSync项任务，则可删除该任务及其相关AWS资源。

### 先决条件

运行任务时，DataSync会自动创建和管理用于数据传输流量的网络接口（有关更多信息，请参阅[网络接口要求](#)）。

要确保在删除任务时删除这些网络接口，您必须具有以下 IAM 权限：

- ec2:DeleteNetworkInterface
- ec2:DescribeNetworkInterfaces
- ec2:ModifyNetworkInterfaceAttribute

这些权限在AWS托管策略中可用AWSDataSyncFullAccess。有关更多信息，请参阅[AWS适用于AWS DataSync 的托管策略](#)：

### 删除任务

任务删除后，将无法恢复。

#### 使用 DataSync 控制台

1. 通过 <https://console.aws.amazon.com/datasync/> 打开AWS DataSync主机。
2. 在左侧导航窗格中，展开 Data Transs，然后选择 Tasks ( 任务 )。
3. 选择要删除的任务。
4. 选择 Actions (操作)，然后选择 Delete (删除)。
5. 在对话框中，选择删除。

#### 使用 AWS CLI

1. 复制以下delete-task命令：

```
aws datasync delete-task \
  --task-arn "task-to-delete"
```

- 对于--task-arn参数，请指定您删除任务的 Amazon 资源名称 (ARN) (例如arn:aws:datasync:us-east-2:123456789012:task/task-012345678abcd0123)。
- 运行 delete-task 命令。

## AWS DataSync任务状态

有一些状态可以帮助您了解您的AWS DataSync任务是准备运行、正在进行还是有问题。

### 任务状态

创建AWS DataSync任务时，您可能会看到这些状态。

控制台状态	API 状态	描述
可用	AVAILABLE	任务已准备就绪，可以开始移动数据了。
正在运行	RUNNING	数据传输正在进行中。
Unavailable	UNAVAILABLE	任务使用的DataSync代理处于脱机状态。
已排队	QUEUED	另一项任务正在使用同一个代理运行。DataSync按顺序运行任务 (先进先出)。有关更多信息，请参阅 <a href="#">运行多个任务</a> ：

### 任务执行状态

当你启动DataSync任务时，你可能会看到这些状态。

控制台状态	API 状态	描述
排队	QUEUED	如果有其他任务正在运行并且正在使用同一个代理，则这是任务执行的第一阶段。有关更多信息，请参阅 <a href="#">运行多个任务</a> ：

控制台状态	API 状态	描述
正在发射	LAUNCHING	如果没有其他任务正在运行且使用相同的代理或者未启用队列，则这是任务执行的第一阶段。此时，DataSync 将初始化任务执行。此状态通常很快会经过，但可能需要几分钟的时间。
正在准备	PREPARING	此时，DataSync正在计算需要传输哪些文件。此阶段所花的时间与源位置中的文件数成比例。这通常需要几分钟到几个小时，具体取决于源和目标文件系统以及这些文件系统的性能。有关更多信息，请参阅 <a href="#">开始你的AWS DataSync任务</a> ：
正在转移	TRANSFERRING	DataSync执行实际的数据传输。在DataSync传输文件时，传输的字节数和文件数会实时更新。
正在验证	VERIFYING	如果您的DataSync任务配置为在传输结束时执行数据完整性检查，则会看到此状态。根据 <a href="#">您配置数据验证的方式</a> ，在非常大的数据量上，这可能需要大量时间。
成功	SUCCESS	如果数据传输成功，则会看到此状态。
错误	ERROR	如果数据传输失败，您将看到此状态。

## AWS DataSync任务场景示例

以下各节为您演练了如何为不同的传输方案创建AWS DataSync任务。

### 创建任务以在自管存储和之间传输数据 AWS

如果您之前创建过某一任务并要创建其他任务，请使用以下过程。

#### 创建任务

1. 通过 <https://console.aws.amazon.com/datasync/> 打开AWS DataSync主机。
2. 在左侧导航窗格中，展开“数据传输”，然后选择“任务”，然后选择“创建任务”。

3. 在 Configure source location (配置源位置) 页面上，如果要对源使用新位置，请选择 Create new location (创建新位置) 并配置新位置。提供配置设置，然后选择 Next (下一步)。有关如何创建位置的说明，请参阅[我可以在哪里传输我的数据AWS DataSync?](#)。

如果您要使用之前创建的源位置，请选择 Choose existing location (选择现有位置)，从列表中选择所需的源位置，然后选择 Next (下一步)。

有关step-by-step说明，请参见[为其创建源位置 AWS DataSync](#)。

## 创建任务以在云端位置之间传输

使用以下说明在 Amazon EC2 实例上设置DataSync代理以进行数据传输。本节中的示例涵盖以下使用案例：

- [将数据从云文件系统传输到另一个云文件系统或 Amazon S3](#)— 将数据从亚马逊 EFS 传输到亚马逊 EFS，从自我管理的 NFS 传输到 EFS，或者传输到 Amazon S3。
- [将数据从 Amazon S3 AWS](#)— 将数据从Amazon S3 传输到亚马逊 EFS，或者从Amazon S3 传输到自我管理的 NFS。

## 创建从云中 NFS 传输到云中 NFS 或 Amazon S3 的任务

使用以下说明将数据从云内 NFS 文件系统传输到。AWS要执行此传输，DataSync代理必须位于部署文件系统的相同AWS 账户位置。AWS 区域这种类型的传输包括从 EFS 到 EFS 的传输、从自我管理的 NFS 到 Amazon EFS 的传输以及向 Amazon S3 的传输。有关云中 NFS 到云中 NFS 或 Amazon S3 的工作原理的信息，请参阅[将数据从云文件系统传输到另一个云文件系统或 Amazon S3](#)

### Note

在源 EFS 或自管 NFS 文件系统AWS 账户所在的位置部署代理。AWS 区域

将您的DataSync代理部署为 Amazon EC2 实例以从云端读取文件

将DataSync代理部署为 Amazon EC2 实例

1. 从源 EFS 所在AWS 账户的位置，使用 Amazon EC2 启动向导中的亚马逊系统映像 (AMI) 启动代理。使用以下 URL 启动 AMI。

```
https://console.aws.amazon.com/ec2/v2/home?region=source-efs-or-nfs-region#LaunchInstanceWizard:ami=ami-id
```

在该 URL 中，将 *source-efs-or-nfs-region* 和 *ami-id* 替换为您自己的值。

AMI 启动后，“选择实例类型”将出现在 Amazon EC2 控制台上。有关 AMI ID 的列表AWS 区域，请参阅[在 Amazon EC2 上部署您的代理](#)。

2. 为您的使用案例选择一种建议的实例类型，然后选择 Next: Configure Instance Details (下一步: 配置实例详细信息)。有关建议的实例类型，请参阅[Amazon EC2 实例要求](#)。
3. 在 Configure Instance Details (配置实例详细信息) 页面中，执行以下操作：
  - a. 对于网络，选择您的源 EFS 或 NFS 所在的 VPC。
  - b. 为自动分配公有 IP 选择一个值。如果希望能够从公共 Internet 中访问您的实例，请将自动分配公有 IP 设置为启用。否则，请将自动分配公有 IP 设置为禁用。如果未分配公有 IP 地址，请在使用其私有 IP 地址的 VPC 中激活代理。

当您从云端 NFS 传输文件时，为了提高性能，我们建议您选择您的 NFS 服务器所在的置放群组。

4. 选择下一步：添加存储。代理不需要额外的存储空间，因此，您可以跳过此步骤，然后选择 Next: Add tags (下一步：添加标签)。
5. (可选) 在添加标签页面上，您可以向您的 Amazon EC2 实例添加标签。当您在页面上完成后，选择 Next: Configure Security Group (下一步：配置安全组)。
6. 在 Configure Security Group (配置安全组) 页面上，执行以下操作：
  - a. 确保所选安全组允许从您计划用于激活代理的 Web 浏览器到 HTTP 端口 80 的入站访问。
  - b. 确保源 EFS 或 NFS 的安全组允许来自代理的入站流量。此外，请确保代理允许流向源 EFS 或 NFS 的出站流量。流量经由标准 NFS 端口 2049。

有关 DataSync 的一组完整的网络要求，请参阅[AWS DataSync网络要求](#)。

7. 选择审核和启动以检查您的配置，然后选择启动以启动您的实例。请记住使用您有权访问的密钥对。将显示一个确认页面，指出您的实例正在启动。
8. 选择 View Instances 以关闭确认页面并返回 Amazon EC2 实例屏幕。当您启动实例时，其初始状态为 pending (待处理)。实例启动后，其状态将更改为 running (正在运行)。此时，已为其分配了公共域名系统 (DNS) 名称和 IP 地址，可在“描述”选项卡中找到。

9. 如果您将自动分配公有 IP 设置为启用，则选择您的实例，然后在描述选项卡中记下公有 IP 地址。您之后将使用该 IP 地址连接到您的同步代理。

如果您将自动分配公有 IP 设置为禁用，则启动或使用 VPC 中的现有实例以激活代理。在这种情况下，您可以使用同步代理的私有 IP 地址从 VPC 中的该实例激活代理。

## 创建从 Amazon EFS 或自管理存储传输数据的任务

接下来，您创建一个任务来传输数据。

### Note

在目标 EFS 或 Amazon S3 存储桶所在的位置创建任务。AWS 区域 AWS 账户

## 创建任务

1. 在目的地AWS 区域所在的位置打开DataSync控制台。

目标 Amazon EFS 或Amazon S3 存储桶必须位于同一存储桶中AWS 账户。

2. 在左侧导航窗格中，选择 Agents，然后选择 Create ( 创建代理 )。
3. 在创建代理向导的激活部分中，输入 Amazon EC2 实例的 IP 地址作为代理地址，然后选择获取密钥。此 IP 地址可以是私有或公有地址。有关更多详情，请参阅[将DataSync代理部署为 Amazon EC2 实例](#)的步骤 9。

您的浏览器连接到此 IP 地址，以便从代理中获取唯一的激活密钥。此密钥可安全地将您的代理与您关联AWS 账户。不需要能够从您的网络外部访问该 IP 地址，但必须能够从浏览器中访问该 IP 地址。

4. 输入便于以后轻松识别的代理名称，完成后选择 Create agent (创建代理)。( 可选 ) 您可以将标签添加到代理。
5. 在左侧导航窗格中，展开 Data Transs，然后选择 Tasks ( 任务 )。
6. 选择“本地到”AWS，然后选择“下一步”以打开“源配置”页面。
7. 在配置源位置页面上，选择创建新位置，然后为位置类型选择网络文件系统 (NFS) 或服务器消息块 (SMB)。填充以下选项：
  - 对于代理，从列表中选择您新创建的代理。
  - 如果您正在从 EFS 进行复制，请执行以下操作：

- 对于 NFS Server (NFS 服务器), 输入源 EFS 的 DNS name (DNS 名称)。
  - 在装载路径中, 输入/ (正斜杠) 并选择下一步。
  - 如果要从自管理 NFS 或 SMB 进行复制, 请执行以下操作:
    - 对于 NFS Server (NFS 服务器), 输入源 NFS 的私有 DNS 或 IP 地址。
    - 在装载路径中, 输入 NFS 服务器导出的路径, 然后选择下一步。
8. 选择 Create new location (创建新位置)。这是数据传输的目标位置。填充以下选项:
- 如果要复制到 EFS, 请执行以下操作:
    - 对于 Location type (位置类型), 选择 EFS。
    - 选择您的目标 EFS。
    - 在装载路径中, 输入/ (正斜杠)。
    - 对于子网和安全组, 请使用默认设置, 然后选择下一步。
  - 如果您要复制到 Amazon S3, 请执行以下操作:
    - 对于位置类型, 选择 Amazon S3 存储桶。
    - 对于 Amazon S3 存储桶, 选择您的源 Amazon S3 存储桶。
    - 对于 Folder (文件夹), 选择用于传输的文件夹前缀, 或将其留为空白。
    - 选择目标 Amazon S3 存储桶和可选文件夹。DataSync 可以生成 AWS Identity and Access Management (IAM) 角色来访问您的存储桶, 也可以自己创建。
9. 选择 Next (下一步), 然后 (可选) 给任务命名并添加标签。
10. 在页面底部选择或创建 Amazon Log CloudWatch s 日志组, 然后选择下一步。有关使用 CloudWatch 日志的更多信息, 请参阅 [允许 DataSync 将日志上传到 CloudWatch 日志组](#)。
11. 检查下一页上的设置, 然后选择 Create task (创建任务)。
12. 选择 Start (启动) 运行您刚刚创建的任务以开始传输数据。

## 创建要从 Amazon S3 传输到云中 NFS 的任务

使用以下说明将数据从 Amazon S3 传输到 AWS 账户与 AWS 区域部署代理相同的云内 NFS 文件系统。这种方法包括从 Amazon S3 传输到 EFS, 或者从 Amazon S3 传输到自我管理的 NFS。下图说明了这种类型的传输。有关 Amazon S3 到云中 NFS 的工作原理的信息, 请参阅 [将数据从 Amazon S3 传输到 AWS 云中 NFS](#)。



## 在 Amazon EC2 实例上部署DataSync代理以写入您的目标位置

首先，在目标 EFS 文件系统或自管 NFS 服务器AWS 账户所在的 Amazon EC2 实例上部署DataSync 代理。AWS 区域

### 部署代理

- 使用 Amazon EC2 启动向导从选定的 AMI 启动代理。为此，请使用以下 URL。

```
https://console.aws.amazon.com/ec2/v2/home?region=DESTINATION-EFS-or-NFS-REGION#LaunchInstanceWizard:ami=AMI-ID.
```

在 URL 中，将AWS 区域和 AMI ID 替换为您自己的 AMI ID。您将会被重定向到 Amazon EC2 控制台上的选择一个实例类型页面。有关 AMI ID 的列表AWS 区域，请参阅[在 Amazon EC2 上部署您的代理](#)。

### 创建任务以从 Amazon S3 传输数据

接下来，您创建一个任务来传输数据。

#### Note

在源 Amazon S3 存储桶AWS 区域所在的AWS 账户和位置中创建任务。

### 创建将数据从 Amazon S3 传输到 EFS 或自我管理的 NFS 或 SMB 的任务

1. 在您的源 Amazon S3 桶打开DataSync控制台。AWS 区域
2. 选择 Create agent (创建代理)。
3. 如果您将自动分配公有 IP 设置为启用，则选择您的实例，然后在描述选项卡中记下公有 IP 地址。您之后将使用该 IP 地址连接到您的同步代理。

如果您将自动分配公有 IP 设置为禁用，则启动或使用 VPC 中的现有实例以激活代理。在这种情况下，您可以使用同步代理的私有 IP 地址从 VPC 中的该实例激活代理。

4. 在创建代理向导中，对于代理地址，输入 Amazon EC2 实例的 IP 地址（私有或公有，如步骤 3 所述），然后选择获取密钥。



您的浏览器连接到此 IP 地址，以便从代理中获取唯一的激活密钥。此密钥可安全地将您的代理与您关联AWS 账户。不需要能够从您的网络外部访问该 IP 地址，但必须能够从浏览器中访问该 IP 地址。

5. 选择一个便于以后识别的代理名称。您可以选择性地添加标签。完成后，选择 Create agent (创建代理)。
6. 选择AWS本地，然后选择下一步。
7. 选择 Create new location (创建新位置)：
  - 对于位置类型，选择 Amazon S3 存储桶。
  - 对于 Amazon S3 存储桶，选择您的源Amazon S3 存储桶。
  - 对于 Folder (文件夹)，选择用于传输的文件夹前缀，或将其留为空白。

DataSync可以生成 IAM 角色来访问您的存储桶，也可以自己创建。

8. 选择下一步。选择创建新位置，为位置类型选择 NFS 或 SMB，然后从列表中选择刚刚创建的代理。
9. a. 如果要复制到 EFS，请执行以下操作：
  - 对于 NFS Server (NFS 服务器)，输入源 EFS 的 DNS name (DNS 名称)。
  - 在装载路径中，输入/ (正斜杠) 并选择下一步。
- b. 如果要复制到云端 NFS，请执行以下操作：
  - 对于 NFS Server (NFS 服务器)，输入源 NFS 的私有 DNS 或 IP 地址。
  - 对于 Mount path (挂载路径)，输入您的 NFS 服务器导出的路径。有关更多信息，请参阅[创建 NFS 位置](#)：
10. 选择 Next (下一步)，然后 (可选) 给任务命名并添加标签。
11. 在页面底部选择或创建CloudWatch日志日志组，然后选择下一步。有关使用CloudWatch日志的更多信息，请参阅[允许DataSync将日志上传到CloudWatch日志组](#)。
12. 检查下一页上的设置，然后选择 Create task (创建任务)。
13. 选择“开始”以运行您刚刚创建的用于传输数据的任务，然后在“启动任务”页面上再次选择“启动”。

# 监控亚马逊的AWS DataSync活动CloudWatch

您可以AWS DataSync使用监控AmazonCloudWatch，此工具可从收集原始数据，DataSync并将数据处理为易读、近乎实时的指标。这些统计数据将保留 15 个月。

默认情况下，DataSync指标数据以 5 分钟为间隔自动发送到CloudWatch。有关更多信息，请参阅[什么是 AmazonCloudWatch ?](#)。在亚马逊CloudWatch用户指南中。

## 亚马逊的CloudWatch指标DataSync

亚马逊CloudWatch提供指标，您可以使用这些指标来获取有关DataSync绩效的信息和解决问题。您可以使用以下工具查看CloudWatch指标：DataSync

- CloudWatch 控制台
- CloudWatch CLI
- CloudWatch API
- DataSync控制台（任务执行页面）

有关信息，请参阅[亚马逊CloudWatch用户指南中的使用亚马逊CloudWatch指标](#)。

DataSync 指标使用 AWS/DataSync 命名空间并提供以下维度的指标。

- AgentId— 代理的唯。
- TaskId— 任务的唯。其格式为 task-01234567890abcdef。

AWS/DataSync 命名空间包括以下指标。

指标	描述
BytesCompressed	应用压缩后通过网络传输的物理字节数。在大多数情况下，BytesTransferred 除非数据不可压缩，否则此数字会小于。  单位：字节
BytesPreparedDestination	在目标位置准备的数据的总字节数。  单位：字节

指标	描述
BytesPreparedSource	在源位置准备的数据的总字节数。 单位：字节
BytesTransferred	传输涉及的总字节数。有关通过网络发送的字节数，请参阅BytesCompressed。 单位：字节
BytesVerifiedDestination	在目标位置验证的数据的总字节数。 单位：字节
BytesVerifiedSource	在源位置验证的数据的总字节数。 单位：字节
BytesWritten	已传输到目标位置的所有文件的总逻辑大小。 单位：字节
FilesPreparedDestination	在目标位置准备的文件的总数。 单位：计数
FilesPreparedSource	在源位置准备的文件的总数。 单位：计数

指标	描述
FilesTransferred	<p>通过网络传输的文件或元数据的实际数量。该值是在该TRANSFERRING 阶段持续计算和更新的。当从源位置读取每个文件并通过网络发送时，它会定期更新。</p> <p>如果传输期间发生故障，该值可能小于 EstimatedFilesToTransfer 。在某些情况下，此值也可能大于 EstimatedFilesTransferred 。此元素是特定于某些位置类型的实施，因此不要将其用作正确文件编号的指示符或监控任务执行。</p> <p>单位：计数</p>
FilesVerifiedDestination	<p>在目标位置验证的文件的总数。</p> <p>单位：计数</p>
FilesVerifiedSource	<p>在源位置验证的文件的总数。</p> <p>单位：计数</p>

## 的亚马逊EventBridge活动DataSync

亚马逊EventBridge事件描述了DataSync资源的变化。您可以设置规则来匹配这些事件，并将这些事件路由到一个或多个目标函数或流。尽最大努力发出事件。

主题

- [DataSync转移事件](#)
- [DataSync探索活动](#)

## DataSync转移事件

以下EventBridge事件可供DataSync转移。

Agent state changes

事件	描述
----	----

## Agent state changes

Online	The agent is configured properly and is available to use. This status is the normal running status for an agent.
Offline	The agent's VM is turned off or the agent is in an unhealthy state and has been out of contact with the service for 5 minutes or longer. When the issue that caused the unhealthy state is resolved, the agent returns to ONLINE status.

## Location state changes

事件	描述
Adding	DataSync is adding a location.
Available	The location is created and is available to use.

## Task state changes

事件	描述
Available	The task was created and is ready to start.
Running	The task is in progress and functioning properly.
Unavailable	The task isn't configured properly and can't be used. You may see this when an agent associated with the task goes offline.
Queued	Another task is running and using the same agent. DataSync runs tasks in series (first in, first out).

## Task execution state changes

事件	描述
----	----

## Agent state changes

Queueing	DataSync is waiting for another task that's using the same agent to finish.
Launching	DataSync is initializing the task execution.
Preparing	DataSync is determining which files need to be transferred.
Transferring	DataSync is performing the actual transfer of your data.
Verifying	DataSync performs a full data and metadata integrity verification to ensure that the data in your destination is an exact copy of your source.
Success	The transfer is successful.
Error	The transfer failed.

## DataSync探索活动

以下EventBridge事件可用于DataSync Discovery。

### Storage system state changes

事件	描述
Storage System Connectivity Status Change	The connection between your DataSync agent and on-premises storage system changed. For details, see your CloudWatch logs.

### Discovery job state changes

事件	描述
----	----

## Storage system state changes

### Discovery Job State Change

The status of your discovery job changed. For more information, see [发现任务状态](#).

### Discovery Job Expiration Soon

Your discovery job expires soon. This includes any information the discovery job collected about your on-premises storage system. Before the job expires, you can export collected data by using the [DescribeStorageSystemResources](#) and [DescribeStorageSystemResourceMetrics](#) operations.

## 允许DataSync将日志上传到CloudWatch日志组

DataSync需要足够的权限才能将日志发送到您的CloudWatch日志组。当您使用控制台创建任务时，DataSync可以自动为您创建具有正确权限的 IAM 资源策略。

以下示例是授予这些权限的资源策略。

```
{
  "Statement": [
    {
      "Sid": "DataSyncLogsToCloudWatchLogs",
      "Effect": "Allow",
      "Action": [
        "logs:PutLogEvents",
        "logs:CreateLogStream"
      ],
      "Principal": {
        "Service": "datasync.amazonaws.com"
      },
      "Condition": {
        "ArnLike": {
          "aws:SourceArn": [
            "arn:aws:datasync:region:account-id:task/*"
          ]
        }
      },
      "StringEquals": {
        "aws:SourceAccount": "account-id"
      }
    }
  ]
}
```

```
    }
  },
  "Resource": "arn:aws:logs:region:account-id:log-group:*:*"
}
],
"Version": "2012-10-17"
}
```

该策略使用条件语句来确保只有来自指定账户的DataSync任务才能访问指定的CloudWatch日志组。我们建议在这些条件语句中使用[aws:SourceArn](#)和[aws:SourceAccount](#)全局条件上下文密钥，以防出现混乱的副手问题。有关更多信息，请参阅[防止跨服务混淆代理](#)：

要指定一个或多个任务，请`region`替换为任务AWS区域所在的区域代码，并`account-id`替换为包含这些任务的账户的AWS账户ID。DataSync要指定CloudWatch日志组，请替换相同的值。您也可以将Resource语句修改为易读、近乎实时的指标。有关使用SourceArn和的更多信息SourceAccount，请参阅IAM用户指南中的[全局条件键](#)。

要应用此策略，请将此策略语句保存到本地计算机上的文件中。然后运行以下AWS CLI命令来应用资源策略：

```
aws logs put-resource-policy --policy-name trustDataSync --policy-document file://full-path-to-policy-file
```

#### Note

使用AWS账户与AWS区域您激活DataSync代理时相同的命令运行此命令。

有关信息，请参阅《Amazon Logs Logs Logs 用户指南》中的使用CloudWatch日志[组和日志流](#)。

## 从命令行监控您的DataSync任务

您可以使用AWS Command Line Interface或标准 Unixwatch 实用程序跟踪您的DataSync任务。

### 主题

- [使用以下方法监控您的任务AWS CLI](#)
- [使用该watch实用程序监控您的任务](#)



## 使用以下方法监控您的任务AWS CLI

要使用 CLI 监控您的DataSync任务状态，请使用describe-task-execution命令。

```
aws datasync describe-task-execution \  
    --task-execution-arn 'arn:aws:datasync:region:account-id:task/task-id/  
execution/task-execution-id'
```

此命令返回有关任务执行的信息，如下所示。

```
{  
  "BytesCompressed": 0,  
  "BytesTransferred": 0,  
  "BytesWritten": 0,  
  "EstimatedFilesToTransfer": 0,  
  "EstimatedBytesToTransfer": 0,  
  "FilesTransferred": 0,  
  "Options": {  
    "VerifyMode": "POINT_IN_TIME_CONSISTENT",  
    "Atime": "BEST_EFFORT",  
    "Mtime": "PRESERVE",  
    "Uid": "INT_VALUE",  
    "Gid": "INT_VALUE",  
    "PreserveDevices": "NONE",  
    "PosixPermissions": "PRESERVE",  
    "PreserveDeletedFiles": "PRESERVE",  
    "OverwriteMode": "NEVER",  
    "TaskQueueing": "ENABLED"  
  },  
  "Result": {  
    "PrepareDuration": 4355,  
    "PrepareStatus": "Ok",  
    "TransferDuration": 5889,  
    "TransferStatus": "Ok",  
    "VerifyDuration": 4538,  
    "VerifyStatus": "Pending"  
  },  
  "StartTime": 1532658526.949,  
  "Status": "VERIFYING",  
  "TaskExecutionArn": "arn:aws:datasync:us-east-1:112233445566:task/  
task-08de6e6697796f026/execution/exec-04ce9d516d69bd52f"  
}
```

如果任务执行成功，则 Status (状态) 的值会变为 SUCCESS (成功)。如果 describe-task-execution 命令失败，结果会发送错误代码，帮助您排查问题。有关错误代码的信息，请参阅[the section called “TaskExecutionResultDetail”](#) DataSyncAPI 参考中的。

## 使用该watch实用程序监控您的任务

要通过命令行实时监控任务进度，可以使用标准 Unixwatch 实用程序。任务执行持续时间值以毫秒为单位测量。

该watch实用程序无法识别DataSync别名。以下示例说明如何直接调用 CLI。

```
# pass '-n 1' to update every second and '-d' to highlight differences
$ watch -n 1 -d \ "aws datasync describe-task-execution --task-execution-arn
'arn:aws:datasync:region:account-id:task/task-id/execution/task execution-id'"
```

# AWS DataSync 中的安全性

AWS 的云安全性的优先级最高。作为 AWS 客户，您将从专为满足大多数安全敏感型企业的要求而打造的数据中心和网络架构中受益。

安全性是 AWS 和您的共同责任。[责任共担模型](#)将其描述为云的安全性和云中的安全性：

- 云的安全性 – AWS负责保护在AWS云中运行AWS服务的基础设施。AWS还向您提供可安全使用的服务。作为 [AWS 合规性计划](#)的一部分，第三方审核人员将定期测试和验证安全性的有效性。要了解适用于 AWS DataSync 的合规性计划，请参阅[合规性计划范围内的 AWS 服务](#)。
- 云中的安全性 - 您的责任由您使用的 AWS 服务决定。您还需要对其它因素负责，包括您的数据的敏感性、您的公司的要求以及适用的法律法规。

该文档帮助您了解如何在使用 DataSync 时应用责任共担模型。以下主题说明如何配置 DataSync 以实现您的安全性和合规性目标。您还将了解如何使用其他AWS服务帮助您监控和保护您的 DataSync 资源。

## 主题

- [AWS DataSync 中的数据保护](#)
- [AWS DataSync 中的 Identity and Access Management](#)
- [使用记录AWSDataSync API 调用AWS CloudTrail](#)
- [AWS DataSync 的合规性验证](#)
- [中的故障恢复能力AWSDataSync](#)
- [中的基础设施安全性AWSDataSync](#)

# AWS DataSync 中的数据保护

AWS DataSync在自我管理的存储系统和存储服务之间以及AWS存储服务之间AWS安全地传输数据。在传输过程中如何加密您的存储数据在一定程度上取决于传输所涉及的位置。

传输完成后，存储数据的系统或服务会对数据进行静态加密（不是DataSync）。

## 主题

- [AWS DataSync传输中加密](#)
- [AWS DataSync静态加密](#)

- [互连网络流量隐私](#)

## AWS DataSync传输中加密

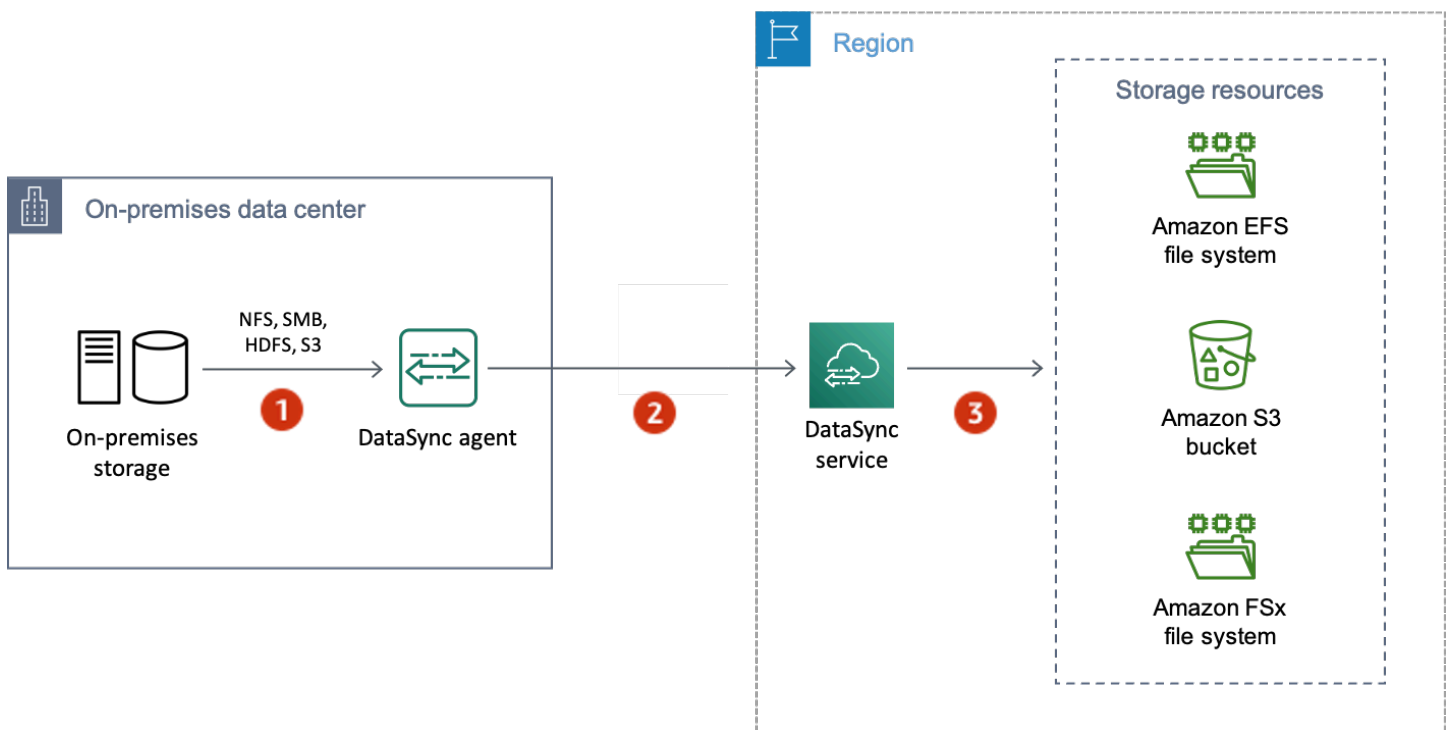
您的存储数据（包括元数据）在传输过程中经过加密，但是在整个传输过程中如何对其进行加密取决于您的来源和目标位置。

连接某个位置时，DataSync使用该位置的数据访问协议提供的最安全的选项。例如，使用服务器消息块 (SMB) 连接文件系统时，DataSync使用 SMB 提供的安全功能。

### 传输中的网络连接

DataSync需要三个网络连接才能复制数据：一个连接用于从源位置读取数据，另一个连接用于在位置之间传输数据，另外一个用于将数据写入目标位置。

下图是DataSync用于将数据从本地存储系统传输到AWS存储服务的网络连接的示例。要了解连接发生在何处以及在数据通过每个连接时如何保护数据，请使用随附的表。



参考	网络连接	描述
1	从源位置读取数据	DataSync使用存储系统的协议访问数据（例如，SMB 或 Amazon S3

参考	网络连接	描述
		API ) 进行连接。对于此连接，使用存储系统的安全功能保护数据。
2	在不同地点之间传输数据	对于此连接，DataSync使用传输层安全性协议 ( TLS ) 1.2 加密进行加密。
3	将数据写入目标位置	与源位置一样，通过使用存储系统的协议访问数据进行DataSync连接。再次使用存储系统的安全功能保护数据。

了解在DataSync连接到以下AWS存储服务时如何对传输中的数据进行加密：

- [Amazon EFS](#)
- [Amazon FSx for Windows File Server](#)
- [Amazon FSx for Lustre](#)
- [Amazon FFSXFSFSFS](#)
- [适用于NetApp ONTAP 的亚马逊 FSx](#)
- [Amazon S3](#)

## TLS 密码

在不同位置之间传输数据时，DataSync使用不同的 TLS 密码。DataSync使用的 TLS 密码取决于用于激活DataSync代理的端点类型。

### 公共或 VPC 终端节点

对于这些终端节点，DataSync使用以下 TLS 密码之一：

- TLS\_ECDHE\_RSA\_WITH\_AES\_256\_GCM\_SHA384 (ecd\_h\_x25519)
- TLS\_ECDHE\_RSA\_WITH\_CHACHA20\_POLY1305\_SHA256 (ecd\_h\_x25519)
- TLS\_ECDHE\_RSA\_WITH\_AES\_128\_GCM\_SHA256 (ecd\_h\_x25519)

## FIPS 终端节点

对于 FIPS 终端节点，DataSync使用以下 TLS 密码：

- TLS\_ECDHE\_RSA\_WITH\_AES\_128\_GCM\_SHA256 (ecd\_h\_x25519)

## AWS DataSync静态加密

因为它AWS DataSync是一项传输服务，所以它通常不管理你的静态存储数据。DataSync支持的存储服务系统和系统负责保护该状态下的数据。但是，有些与服务相关的数据是静态DataSync管理的。

### 什么是加密的？

唯一能DataSync处理静态的数据与它发现的有关您的本地存储系统的信息以及完成传输所需的详细信息有关。DataSync使用完全静态加密将以下数据存储存储在 Amazon DynamoDB 中：

- 收集的有关您的本地存储系统的信息（如果您使用DataSync Discovery）。这些信息还以完全静态加密方式存储在 Amazon S3 中。
- 任务配置（例如，有关传输中位置的详细信息）。
- 允许您的DataSync代理在某个位置进行身份验证的用户证书。这些凭证是使用代理的公钥加密的。代理可以根据需要使用其私钥解密这些密钥。

有关更多信息，请参阅 Amazon DynamoDB 开发人员指南中的 DynamoDB [静态加密](#)进行加密。

### DataSyncDiscovery 收集的信息

DataSyncDiscovery 可存储和管理其收集的有关您的本地存储系统的数据，最长可达 60 天。您可使用 AmazonEventBridge 在到期日期临近时通知您。有关更多信息，请参阅[DataSync探索活动](#)：

从DataSync Discovery 中删除本地存储系统资源时，您将永久删除所有相关的发现任务、收集的数据和建议。

### 密钥管理

您无法管理用于在 DynamoDB 中存储与运行任务相关的信息的加密密钥。DataSync此信息包括您的任务配置和代理用于对存储位置进行身份验证的凭据。

## 什么没有加密？

尽管DataSync无法控制存储数据的静态加密方式，但我们仍然建议将您的位置配置为其支持的最高安全级别。例如，您可以使用 Amazon S3 托管密钥 (SSE-S3) 或 (KMS) 密AWS钥AWS Key Management Service (SSE-KMS) 加密进行加密。

详细了解AWS存储服务如何加密静态数据：

- [Amazon EFS](#)
- [Amazon FSx for Windows File Server](#)
- [Amazon FSx for Lustre](#)
- [Amazon FFSXFSFSFS](#)
- [适用于NetApp ONTAP 的亚马逊 FSx](#)
- [Amazon S3](#)

## 互连网络流量隐私

我们建议将您的源位置和目标位置配置为每个站点所支持的最高安全级别。连接到某个位置时AWS DataSync，使用存储系统使用的最安全版本的数据访问协议。此外，考虑将子网流量限制在已知协议和服务上。

DataSync使用传输层安全性协议 ( TLS ) 1.2，保护位置之间的AWS 账户连接。AWS 区域

## AWS DataSync 中的 Identity and Access Management

AWS 使用安全凭证来识别您的身份并向您授予对 AWS 资源的访问权限。利用 AWS Identity and Access Management (IAM) 的功能，可在不共享您的安全凭证的情况下允许其他用户、服务和应用程序完全使用或受限使用您的 AWS 资源。

默认情况下，IAM 身份 ( 用户、组和角色 ) 没有创建、查看或修改AWS资源的权限。为了允许用户、群组和角色访问AWS DataSync资源并与DataSync控制台和 API 进行交互，我们建议您使用 IAM 策略向他们授予使用他们所需的特定资源和 API 操作的权限。然后，将策略附加到需要访问的 IAM 身份。有关策略的基本要素的概述，请参阅[访问管理AWS DataSync](#)。

### 主题

- [访问管理AWS DataSync](#)
- [AWS适用于 AWS DataSync 的托管策略](#)

- [IAM 客户托管式策略AWS DataSync](#)
- [将服务相关角色用于 AWS DataSync](#)
- [在创建过程中授予标记AWS DataSync资源的权限](#)
- [防止跨服务混淆代理](#)
- [DataSyncAPI 权限：操作和资源](#)

## 的访问管理AWS DataSync

每种AWS资源都归... 所有AWS 账户. 创建或访问资源的权限由权限策略进行管理。账户管理员可以将权限策略附加到AWS Identity and Access Management (IAM) 身份。有些服务 ( 例如 AWS Lambda ) 还支持向资源附加权限策略。

### Note

账户管理员是具有管理员权限的用户AWS 账户。有关更多信息，请参阅 IAM 用户指南中的 [IAM 最佳实践](#)。

### 主题

- [DataSync 资源和操作](#)
- [了解资源所有权](#)
- [管理对资源的访问](#)
- [指定策略元素：操作、效果、资源和主体](#)
- [在策略中指定条件](#)

## DataSync 资源和操作

在中DataSync，主要资源是代理、位置、任务和任务执行。

这些资源具有关联的唯一 Amazon 资源名称 (ARN)，如下表所示。

资源类型	ARN 格式
客服 ARN	arn:aws:datsync: <i>region</i> : <i>account-id</i> :agent/ <i>agent-id</i>



资源类型	ARN 格式
位置 ARN	arn:aws:datsync: <i>region</i> : <i>account-id</i> :location/ <i>location-id</i>
任务 ARN	arn:aws:datsync: <i>region</i> : <i>account-id</i> :task/ <i>task-id</i>
任务执行 ARN	arn:aws:datsync: <i>region</i> : <i>account-id</i> :task/ <i>task-id</i> /executio n/ <i>exec-id</i>

为授予执行特定 API 操作（如创建任务）的权限，需要 DataSync 定义一组您可以在权限策略中指定的操作。一个 API 操作可能需要执行多个操作的权限。有关所有 DataSync API 操作及其适用的资源的列表，请参阅 [DataSync API 权限：操作和资源](#)。

## 了解资源所有权

资源拥有者是创建 AWS 账户该资源的。也就是说，资源拥有者是 AWS 账户委托人实体（例如 IAM 角色）的委托人。以下示例说明了这种行为的工作原理：

- 如果您使用 root 账户凭证创建任务，AWS 账户则您即为该资源的所有者（中 DataSync，资源即任务）。AWS 账户
- 如果您在您的 IAM 角色中创建一个 IAM 角色，AWS 账户并授予该用户执行 CreateTask 操作的权限，则该用户便能创建任务。但是，这些任务资源将归用户所属的用户所有。AWS 账户
- 如果您在您的 IAM 角色中创建 AWS 账户具有创建任务的权限的 IAM 角色，则能够担任该角色的任何人都可以创建任务。该角色所属的任务资源将归该角色所属的所有。AWS 账户

## 管理对资源的访问

权限策略规定谁可以访问哪些内容。下一节介绍创建权限策略时的可用选项。

### Note

本节讨论如何在 DataSync 范围内使用 IAM。这里不提供有关 IAM 服务的详细信息。有关完整的 IAM 文档，请参阅 IAM 用户指南中的 [什么是 IAM？](#)。有关 IAM policy 语法和说明的信息，请参阅《IAM 用户指南》中的 [AWS Identity and Access Management 策略参考](#)。

附加到 IAM 身份的策略称作基于身份的策略（IAM policy），附加到资源的策略称作基于资源的策略。DataSync 只支持基于身份的策略（IAM 策略）。

## 主题

- [基于身份的策略](#)
- [基于资源的策略](#)

### 基于身份的策略

您可以使用 IAM 策略管理DataSync资源访问权限。这些策略可以帮助AWS 账户管理员执行以下操作DataSync：

- 授予创建和管理DataSync资源的权限-创建 IAM 策略，AWS 账户允许您中的 IAM 角色创建和管理DataSync资源，例如代理、地点和任务。
- 向其他角色AWS 账户或中的角色授予权限AWS 服务 — 创建一个 IAM 策略，向其他AWS 账户角色中的 IAM 角色授予权限AWS 服务。例如：
  1. 账户 A 管理员可以创建一个 IAM 角色，然后向该角色附加授予其访问账户 A 中资源的权限策略。
  2. 账户 A 管理员向角色附加信任策略，用于将账户 B 标识为可以担任该角色的委托人。

要授予代入该角色的AWS 服务权限，账户 A 管理员可以在信任策略中指定一个AWS 服务作为委托人。

3. 之后，账户 B 管理员可以委派权限，指定账户 B 中的任何用户担任该角色。这样，账户 B 中使用该角色的任何人都可以创建或访问账户 A 中的资源。

有关使用 IAM 委派权限的更多信息，请参阅 IAM 用户指南中的[访问权限管理](#)。

以下示例策略授予对所有资源List\*执行所有资源的权限。此操作是只读操作，不允许修改资源。

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "AllowAllListActionsOnAllResources",
      "Effect": "Allow",
      "Action": [
        "datasync:List*"
      ],
      "Resource": "*"
    }
  ]
}
```

有关使用基于身份的策略的更多信息DataSync，请参阅[AWS托管策略](#)和[客户托管策略](#)。有关 IAM 身份的更多信息，请参阅 [IAM 用户指南](#)。

## 基于资源的策略

其他服务（如 Amazon S3）支持基于资源的权限策略。例如，您可以将策略附加到 Amazon S3 存储桶以管理对该存储桶的访问权限。但DataSync不支持基于资源的策略。

## 指定策略元素：操作、效果、资源和主体

对于每个 DataSync 资源（请参阅 [DataSyncAPI 权限：操作和资源](#)），该服务都定义了一组 API 操作（请参阅[操作](#)）。为授予这些 API 操作的权限，DataSync 定义了一组您可以在策略中指定的操作。例如，对于 DataSync 资源，定义了以下操作：CreateTask、DeleteTask 和 DescribeTask。执行一个 API 操作可能需要多个操作的权限。

以下是最基本的策略元素：

- Resource（资源）- 在策略中，您可以使用 Amazon Resource Name (ARN) 标识策略应用到的资源。对于 DataSync 资源，您可以在 IAM 策略中使用通配符（\*）。有关更多信息，请参阅[DataSync 资源和操作](#)：
- 操作 – 您可以使用操作关键字标识要允许或拒绝的资源操作。例如，根据指定Effect元素，权datsync:CreateTask限允许或拒绝执行DataSyncCreateTask操作的用户权限。
- Effect（效果）— 您可以指定当用户请求特定操作（可以是**Allow**或）时的效果**Deny**。如果没有显式授予 (Allow) 对资源的访问权限，则隐式拒绝访问。您还可以显式拒绝对资源的访问，这样可确保用户无法访问该资源，即使有其他策略授予了该用户访问权限的情况下也是如此。有关更多信息，请参阅 IAM 用户指南中的[授权](#)。
- 主体 – 在基于身份的策略（IAM policy）中，附加了策略的用户是隐式主体。对于基于资源的策略，您可以指定要接收权限的用户、账户、服务或其他实体（仅适用于基于资源的策略）。DataSync 不支持基于资源的策略。

有关 IAM policy 语法和说明的更多信息，请参阅《IAM 用户指南》中的[AWS Identity and Access Management策略参考](#)。

有关显示所有 DataSync API 操作的表，请参阅 [DataSyncAPI 权限：操作和资源](#)。

## 在策略中指定条件

当您授予权限时，可使用 IAM 策略语言指定一些条件，这些条件规定在授予权限时策略何时生效。例如，您可能希望策略仅在特定日期后应用。有关使用策略语言指定条件的更多信息，请参阅《IAM 用户指南》中的[条件](#)。

要表示条件，您可以使用预定义的条件键。没有特定于 DataSync 的条件键。但有 AWS 范围内的条件密钥，您可以根据需要使用。有关 AWS 范围内的键的完整列表，请参阅《IAM 用户指南》中的[可用键](#)。

## AWS适用于 AWS DataSync 的托管策略

要向用户、组和角色添加权限，与自己编写策略相比，使用 AWS 托管式策略更简单。创建仅为团队提供所需权限的[IAM 客户托管策略](#)需要时间和专业知识。要快速入门，您可以使用我们的 AWS 托管式策略。这些策略涵盖常见使用案例，可在您的 AWS 账户中使用。有关 AWS 托管式策略的更多信息，请参阅 IAM 用户指南中的[AWS 托管式策略](#)。

AWS 服务负责维护和更新 AWS 托管式策略。您无法更改 AWS 托管式策略中的权限。服务偶尔会向 AWS 托管式策略添加额外权限以支持新功能。此类更新会影响附加策略的所有身份（用户、组和角色）。当启动新功能或新操作可用时，服务最有可能会更新 AWS 托管式策略。服务不会从 AWS 托管式策略中删除权限，因此策略更新不会破坏您的现有权限。

此外，AWS 还支持跨多种服务的工作职能的托管式策略。例如，ReadOnlyAccess AWS 托管式策略提供对许多 AWS 服务和资源的只读访问权限。当服务启动新功能时，AWS 会为新操作和资源添加只读权限。有关工作职能策略的列表和说明，请参阅 IAM 用户指南中的[适用于工作职能的 AWS 托管策略](#)。

### AWS 托管策略：AWSDataSyncReadOnlyAccess

您可以将 AWSDataSyncReadOnlyAccess 策略附加得到 IAM 身份。

此策略授予的只读权限DataSync。

```
{
  "Version": "2012-10-17",
```

```

    "Statement": [
      {
        "Effect": "Allow",
        "Action": [
          "datasync:Describe*",
          "datasync:List*",
          "ec2:DescribeSecurityGroups",
          "ec2:DescribeSubnets",
          "elasticfilesystem:DescribeFileSystems",
          "elasticfilesystem:DescribeMountTargets",
          "fsx:DescribeFileSystems",
          "iam:GetRole",
          "iam:ListRoles",
          "logs:DescribeLogGroups",
          "logs:DescribeResourcePolicies",
          "s3:ListAllMyBuckets",
          "s3:ListBucket"
        ],
        "Resource": "*"
      }
    ]
  }
}

```

## AWS托管策略：AWSDataSyncFullAccess

您可以将 `AWSDataSyncFullAccess` 策略附加得到 IAM 身份。

此策略授予访问服务的管理权限DataSync，并且是AWS Management Console访问该服务所必需的。`AWSDataSyncFullAccess`提供对DataSync API 操作和描述相关资源的操作（例如 Amazon S3 存储桶和 Amazon EFS 文件系统）的完全访问权限。该策略还向亚马逊授予权限CloudWatch，包括创建日志组以及创建或更新资源策略。

```

{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "datasync:*",
        "ec2:CreateNetworkInterface",

```

```
"ec2:CreateNetworkInterfacePermission",
"ec2:DeleteNetworkInterface",
"ec2:DescribeNetworkInterfaces",
"ec2:DescribeSecurityGroups",
"ec2:DescribeSubnets",
"ec2:DescribeVpcEndpoints",
"ec2:ModifyNetworkInterfaceAttribute",
"fsx:DescribeFileSystems",
"fsx:DescribeStorageVirtualMachines",
"elasticfilesystem:DescribeAccessPoints",
"elasticfilesystem:DescribeFileSystems",
"elasticfilesystem:DescribeMountTargets",
"iam:GetRole",
"iam:ListRoles",
"logs:CreateLogGroup",
"logs:DescribeLogGroups",
"logs:DescribeResourcePolicies",
"outposts:ListOutposts",
"s3:GetBucketLocation",
"s3:ListAllMyBuckets",
"s3:ListBucket",
"s3-outposts:ListAccessPoints",
"s3-outposts:ListRegionalBuckets"
],
"Resource": "*"
},
{
  "Effect": "Allow",
  "Action": [
    "iam:PassRole"
  ],
  "Resource": "*",
  "Condition": {
    "StringEquals": {
      "iam:PassedToService": [
        "datasync.amazonaws.com"
      ]
    }
  }
}
]
}
```

更改	说明	日期
<a href="#">AWSDataSyncFullAccess</a> – 对现有策略的更新	<p>DataSync向 : 添加了新权限AWSDataSyncFullAccess :</p> <ul style="list-style-type: none"> <li>• ec2:DescribeVpcEndpoints</li> <li>• elasticfilesystem:DescribeAccessPoints</li> <li>• fsx:DescribeStorageVirtualMachines</li> <li>• outposts:ListOutposts</li> <li>• s3:GetBucketLocation</li> <li>• s3-outposts:ListAccessPoints</li> <li>• s3-outposts:ListRegionalBuckets</li> </ul>	2013 3 3 3 3 3
DataSync 已开启跟踪更改	DataSync 为其 AWS 托管式策略开启了跟踪更改。	2021 年 3 月 1 日

## IAM 客户托管式策略AWS DataSync

除托管策略外，您还可以为AWS DataSync API 操作创建自己的基于身份的策略，并将其附加到需要这些权限的AWS Identity and Access Management (IAM) 身份。这些策略称为客户托管型策略，是您在自己的中管理的独立策略AWS 账户。

### Important

在开始之前，我们建议您首先了解管理DataSync资源访问权限的基本概念和选项。有关更多信息，请参阅[访问管理AWS DataSync](#)：

## 自定义策略概览

以下示例是授予使用某些DataSync操作的权限的策略。

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "AllowsSpecifiedActionsOnAllTasks",
      "Effect": "Allow",
      "Action": [
        "datasync:DescribeTask",
        "datasync:ListTasks"
      ],
      "Resource": "arn:aws:datasync:us-east-2:111222333444:task/*"
    }
  ]
}
```

该策略有一条语句（请注意语句中的Action和Resource元素），它执行以下操作：

- 使用亚马逊资源名称 (ARN `arn:aws:datasync:ListTasks`) 授予对某些任务资源执行两项DataSync操作（`datasync:DescribeTask`和）的权限。
- 在任务 ARN 中指定通配符 (\*)，因为允许 IAM 角色对所有任务执行这两个操作。要将操作权限限制为特定任务，请在该语句中指定任务 ID 而不是通配符。

## 自定义策略的示例

以下示例用户策略授予执行各种DataSync操作的权限。如果您使用的是AWS SDK 或AWS Command Line Interface (AWS CLI)，这些策略就会起作用。要在控制台中使用这些策略，还必须使用托管策略AWSDataSyncFullAccess。

示例 1：创建允许DataSync访问您的 Amazon S3 存储桶的信任关系

下面是一个允许 DataSync 担任 IAM 角色的示例信任策略。此角色DataSync允许访问 Amazon S3 存储桶。为防止出现[跨服务混淆的副手问题](#)，我们建议在策略中使用[aws:SourceArn](#)和[aws:SourceAccount](#)全局条件上下文密钥。

```
{
  "Version": "2012-10-17",
  "Statement": [
```



```
{
  "Effect": "Allow",
  "Principal": {
    "Service": "datasync.amazonaws.com"
  },
  "Action": "sts:AssumeRole",
  "Condition": {
    "StringEquals": {
      "aws:SourceAccount": "123456789012"
    },
    "StringLike": {
      "aws:SourceArn": "arn:aws:datasync:us-east-2:123456789012:*"
    }
  }
}
]
```

示例 2：DataSync 允许读取和写入您的 Amazon S3 存储桶

以下示例策略授予 DataSync 对 S3 存储桶读取和写入数据的最低权限。

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Action": [
        "s3:GetBucketLocation",
        "s3:ListBucket",
        "s3:ListBucketMultipartUploads"
      ],
      "Effect": "Allow",
      "Resource": "YourS3BucketArn"
    },
    {
      "Action": [
        "s3:AbortMultipartUpload",
        "s3:DeleteObject",
        "s3:GetObject",
        "s3:ListMultipartUploadParts",
        "s3:GetObjectTagging",
        "s3:PutObjectTagging",
        "s3:PutObject"
      ],
      "Effect": "Allow",
      "Resource": "YourS3BucketArn"
    }
  ]
}
```

```
        "Effect": "Allow",
        "Resource": "YourS3BucketArn/*"
    }
]
}
```

### 示例 3：DataSync允许将日志上传到CloudWatch日志组

DataSync需要权限才能将日志上传到您的亚马逊CloudWatch日志组。您可以使用CloudWatch日志组来监视和调试您的任务。

有关授予此类权限的 IAM 策略的示例，请参阅[允许DataSync将日志上传到CloudWatch日志组](#)。

## 将服务相关角色用于 AWS DataSync

AWS DataSync 使用 AWS Identity and Access Management ( IAM ) [服务相关角色](#)。服务相关角色是一种独特类型的 IAM 角色，它与 IAM 角色直接相关 DataSync。服务相关角色由预定义，DataSync 并包含该服务代表您调用其他AWS服务所需的所有权限。

服务相关角色可让您 DataSync 更轻松地进行设置，因为您不必手动添加必要的权限。DataSync 定义其服务相关角色的权限，除非另外定义，否则 DataSync 只能代入该角色。定义的权限包括信任策略和权限策略，以及不能附加到任何其它 IAM 实体的权限策略。

只有在首先删除相关资源后，才能删除服务相关角色。这将保护您的 DataSync 资源，因为您不会无意中删除对资源的访问权限。

有关支持服务相关角色的其他服务的信息，请参阅[与 IAM 配合使用的 AWS 服务](#)，并查找 Service-linked roles ( 服务相关角色 ) 列中显示为 Yes ( 是 ) 的服务。请选择 Yes 与查看该服务的服务相关角色文档的链接。

### DataSync 的服务相关角色权限

DataSync 使用 DataSync Discovery 使用的名为 AWSServiceRoleForDataSyncDiscovery“服务关联角色”的服务关联角色来实现与其他AWS服务的集成。

AWSServiceRoleForDataSyncDiscovery 服务相关角色信任以下服务代入该角色：

- `discovery-datasync.amazonaws.com`

名为的角色权限策略AWSDataSyncDiscoveryServiceRolePolicy可以 DataSync 对指定资源完成以下操作：

```

{
  "Version": "2012-10-17",
  "Statement": [{
    "Effect": "Allow",
    "Action": [
      "secretsmanager:GetSecretValue"
    ],
    "Resource": [
      "arn*:secretsmanager:*:*:secret:datasync!*"
    ],
    "Condition": {
      "StringEquals": {
        "secretsmanager:ResourceTag/aws:secretsmanager:owningService":
"datasync",
        "aws:ResourceAccount": "${aws:PrincipalAccount}"
      }
    }
  },
  {
    "Effect": "Allow",
    "Action": [
      "logs:CreateLogGroup",
      "logs:CreateLogStream"
    ],
    "Resource": [
      "arn*:logs:*:*:log-group:/aws/datasync*"
    ]
  },
  {
    "Effect": "Allow",
    "Action": [
      "logs:PutLogEvents"
    ],
    "Resource": [
      "arn*:logs:*:*:log-group:/aws/datasync:log-stream:*"
    ]
  }
]
}

```

您必须配置权限以允许 IAM 实体 (例如, 角色) 创建、编辑或删除服务相关角色。有关更多信息, 请参阅 IAM 用户指南中的 [服务相关角色权限](#)。

## 为 DataSync 创建服务相关角色

无需手动创建服务相关角色。当您在 AWS Management Console、或 AWS API 中添加存储系统时，DataSync 会为您创建服务相关角色。AWS CLI

如果删除此服务相关角色，然后需要再次创建，可以使用相同流程在账户中重新创建此角色。添加存储系统时，将再次为您 DataSync 创建服务相关角色。

您也可以使用 IAM 控制台为 DataSync Discovery 使用案例创建服务相关角色。在 AWS CLI 或 AWS API 中，使用 `discovery-datasync.amazonaws.com` 服务名称创建服务相关角色。有关更多信息，请参阅《IAM 用户指南》中的[创建服务相关角色](#)。如果您删除了此服务相关角色，则可以使用此相同过程再次创建角色。

## 为 DataSync 编辑服务相关角色

DataSync 不允许您编辑 `AWSServiceRoleForDataSyncDiscovery` 服务相关角色。这包括该角色的名称，因为可能有不同的实体引用该角色。但是可以使用 IAM 编辑角色说明。有关更多信息，请参阅 IAM 用户指南中的[编辑服务相关角色](#)。

## 删除 DataSync 的服务相关角色

如果不再需要使用某个需要服务相关角色的功能或服务，我们建议您删除该角色。这样您就没有未被主动监控或维护的未使用实体。但是，您必须先清除服务相关角色的资源，然后才能手动删除它。

### Note

如果 DataSync 服务在您试图删除资源时正在使用该角色，则删除操作可能会失败。如果发生这种情况，请等待几分钟后重试。

删除 `AWSServiceRoleForDataSyncDiscovery` 角色使用的 DataSync 资源

1. [移除您在 DataSync Discovery 中使用的本地存储系统](#)。
2. 使用 IAM 删除服务相关角色。

使用 IAM 控制台，即 AWS CLI 或 AWS API 来删除 `AWSServiceRoleForDataSyncDiscovery` 服务相关角色。有关更多信息，请参阅 IAM 用户指南中的[删除服务相关角色](#)。

## DataSync 服务相关角色的受支持区域

DataSync 支持在服务可用的所有区域中使用服务相关角色。有关更多信息，请参阅 [DataSync 区域和终端节点](#)。

## 在创建过程中授予标记AWS DataSync资源的权限

某些资源创建 AWS DataSync API 操作允许您在创建资源时指定标签。您可以使用资源标签来实现基于属性的访问控制 (ABAC)。有关更多信息，请参阅 [什么是适用于的 ABACAWS?](#) 在 IAM 用户指南中。

为使用户能够在创建时为资源添加标签，他们必须具有使用创建该资源的操作（如 `datsync:CreateAgent` 或 `datsync:CreateTask`）的权限。如果在资源创建操作中指定了标签，则用户还必须具有使用 `datsync:TagResource` 操作的显式权限。

仅当用户在资源创建操作中应用了标签时，系统才会评估 `datsync:TagResource` 操作。因此，如果未在此请求中指定任何标签，则拥有创建资源权限（假定没有标记条件）的用户无需具备使用 `datsync:TagResource` 操作的权限。

但是，如果用户不具备使用 `datsync:TagResource` 操作的权限而又试图创建带标签的资源，则请求将失败。

### 示例 IAM policy

使用以下示例 IAM 策略声明向创建DataSync资源的用户授予TagResource权限。

以下语句允许用户在创建DataSync代理时标记代理。

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": "datsync:TagResource",
      "Resource": "arn:aws:datsync:region:account-id:agent/*"
    }
  ]
}
```

以下语句允许用户在创建DataSync位置时标记该位置。

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": "datasync:TagResource",
      "Resource": "arn:aws:datasync:region:account-id:location/*"
    }
  ]
}
```

以下语句允许用户在创建DataSync任务时标记任务。

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": "datasync:TagResource",
      "Resource": "arn:aws:datasync:region:account-id:task/*"
    }
  ]
}
```

## 防止跨服务混淆代理

混淆代理问题是一个安全性问题，即不具有操作执行权限的实体可能会迫使具有更高权限的实体执行该操作。在 AWS 中，跨服务模拟可能会导致混淆代理问题。一个服务（呼叫服务）调用另一项服务（所谓的被调服务）时，可能会发生跨服务模拟。可以操纵调用服务，使用其权限以在其他情况下该服务不应有访问权限的方式对另一个客户的资源进行操作。为防止这种情况，AWS 提供可帮助您保护所有服务的数据的工具，而这些服务中的服务委托人有权访问账户中的资源。

我们建议在资源策略中使用 [aws:SourceArn](#) 和 [aws:SourceAccount](#) 全局条件上下文键，以限制 AWS DataSync 为其他服务提供的资源访问权限。如果同时使用全局条件上下文密钥和包含账户 ID 的 [aws:SourceArn](#) 值，则 [aws:SourceAccount](#) 值和 [aws:SourceArn](#) 值中的账户在同一策略语句中使用时，必须使用相同的账户 ID。如果您只希望将一个资源与跨服务访问相关联，请使用 [aws:SourceArn](#)。[aws:SourceAccount](#) 如果您想将该账户中的任何资源与跨服务使用相关联，请使用。

的值 [aws:SourceArn](#) 必须包括允许担任 IAM 角色 DataSync 的地点 ARN。DataSync

防范混淆代理问题最有效的方法是`aws:SourceArn`使用带有资源完整 ARN。如果您不知道完整 ARN，或者正在指定多个资源，请针对未知部分使用通配符 (\*)。下面是一些如何针对 DataSync 以下操作示例。

- 要将信任策略限制在现有 DataSync 位置，请在策略中包含完整的地点 ARN。DataSync 仅在处理该特定位置时才会担任 IAM 角色。
- 在为创建 Amazon S3 地点时 DataSync，您不知道该地点的 ARN。在这些情况下，使用以下格式作为 `aws:SourceArn` 密钥：`arn:aws:datsync:us-east-2:123456789012:*`。此格式验证分区 (aws)、账户 ID 和区域。

以下完整示例演示如何使用信任策略中的全局条件上下文键 `aws:SourceArn` 和 `aws:SourceAccount` 全局条件上下文键来防范混淆代理策略中的全局条件上下文键和全局条件上下文键来防范混淆代理问题 DataSync

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Principal": {
        "Service": "datsync.amazonaws.com"
      },
      "Action": "sts:AssumeRole",
      "Condition": {
        "StringEquals": {
          "aws:SourceAccount": "123456789012"
        },
        "StringLike": {
          "aws:SourceArn": "arn:aws:datsync:us-east-2:123456789012:*"
        }
      }
    }
  ]
}
```

有关显示如何使用 `aws:SourceArn` 和 `aws:SourceAccount` 全局条件上下文键的更多策略示例 DataSync，请参阅以下主题：

- [创建允许 DataSync 访问您的 Amazon S3 存储桶的信任关系](#)
- [配置 IAM 角色以访问您的 Amazon S3 存储桶](#)

## DataSyncAPI 权限：操作和资源

在创建AWS Identity and Access Management (IAM) 策略时，此页面可以帮助您了解AWS DataSync API 操作、您可授予执行权限的对应操作以及您可授予权限的AWS资源。

通常，以下是向策略添加DataSync权限的方式：

- 在Action元素中指定操作。该值包括前datasync:缀和 API 操作名称。例如，datasync:CreateTask。
- 在Resource元素中指定与操作相关的AWS资源。

您也可以在DataSync策略中使用AWS条件密钥。有关AWS键的完整列表，请参阅《IAM 用户指南》中的[可用键](#)。

有关DataSync资源及其Amazon 资源名称 (ARN) 格式的列表[DataSync 资源和操作](#)。

DataSyncAPI 操作和相应的操作

### [AddStorageSystem](#)

操作：datasync:AddStorageSystem

资源：无

操作：

- kms:Decrypt
- iam:CreateServiceLinkedRole

资源：\*

操作：secretsmanager:CreateSecret

资源：arn:aws:secretsmanager:*region*:*account-id*:secret:datasync!\*

### [CancelTaskExecution](#)

操作：datasync:CancelTaskExecution

资源：arn:aws:datasync:*region*:*account-id*:task/*task-id*/execution/*exec-id*



## [CreateAgent](#)

操作 : `datasync:CreateAgent`

资源 : 无

## [CreateLocationEfs](#)

操作 : `datasync:CreateLocationEfs`

资源 : 无

## [CreateLocationFsxLustre](#)

操作 : `datasync:CreateLocationFsxLustre`

资源 : 无

## [CreateLocationFsxOntap](#)

操作 : `datasync:CreateLocationFsxOntap`

资源 : 无

## [CreateLocationFsxOpenZfs](#)

操作 : `datasync:CreateLocationFsxOpenZfs`

资源 : 无

## [CreateLocationFsxWindows](#)

操作 : `datasync:CreateLocationFsxWindows`

资源 : 无

## [CreateLocationHdfs](#)

操作 : `dataSync:CreateLocationHdfs`

资源 : `arn:aws:datasync:region:account-id:agent/agent-id`

## [CreateLocationNfs](#)

操作 : `datasync:CreateLocationNfs`

资源 : `arn:aws:datasync:region:account-id:agent/agent-id`

## [CreateLocationObjectStorage](#)

操作 : `dataSync:CreateLocationObjectStorage`

资源 : `arn:aws:datasync:region:account-id:agent/agent-id`

## [CreateLocationS3](#)

操作 : `datasync:CreateLocationS3`

资源 : `arn:aws:datasync:region:account-id:agent/agent-id` ( 仅适用于Outposts 基地上的Amazon S3 )

## [CreateLocationSmb](#)

操作 : `datasync:CreateLocationSmb`

资源 : `arn:aws:datasync:region:account-id:agent/agent-id`

## [CreateTask](#)

操作 : `datasync:CreateTask`

资源 :

- `arn:aws:datasync:region:account-id:location/source-location-id`
- `arn:aws:datasync:region:account-id:location/destination-location-id`

## [DeleteAgent](#)

操作 : `datasync>DeleteAgent`

资源 : `arn:aws:datasync:region:account-id:agent/agent-id`

## [DeleteLocation](#)

操作 : `datasync>DeleteLocation`

资源 : `arn:aws:datasync:region:account-id:location/location-id`

## [DeleteTask](#)

操作 : `datasync>DeleteTask`

资源 : `arn:aws:datasync:region:account-id:task/task-id`

## [DescribeAgent](#)

操作 : `datasync:DescribeAgent`

资源 : `arn:aws:datasync:region:account-id:agent/agent-id`

## [DescribeDiscoveryJob](#)

操作 : `datasync:DescribeDiscoveryJob`

资源 : `arn:aws:datasync:region:account-id:system/storage-system-id/job/discovery-job-id`

## [DescribeLocationEfs](#)

操作 : `datasync:DescribeLocationEfs`

资源 : `arn:aws:datasync:region:account-id:location/location-id`

## [DescribeLocationFsxLustre](#)

操作 : `datasync:DescribeLocationFsxLustre`

资源 : `arn:aws:datasync:region:account-id:location/location-id`

## [DescribeLocationFsxOntap](#)

操作 : `datasync:DescribeLocationFsxOntap`

资源 : `arn:aws:datasync:region:account-id:location/location-id`

## [DescribeLocationFsxOpenZfs](#)

操作 : `datasync:DescribeLocationFsxOpenZfs`

资源 : `arn:aws:datasync:region:account-id:location/location-id`

## [DescribeLocationFsxWindows](#)

操作 : `datasync:DescribeLocationFsxWindows`

资源 : `arn:aws:datasync:region:account-id:location/location-id`

## [DescribeLocationHdfs](#)

操作 : `datasync:DescribeLocationHdfs`

资源 : `arn:aws:datsync:region:account-id:location/location-id`

#### [DescribeLocationNfs](#)

操作 : `datsync:DescribeLocationNfs`

资源 : `arn:aws:datsync:region:account-id:location/location-id`

#### [DescribeLocationObjectStorage](#)

操作 : `datsync:DescribeLocationObjectStorage`

资源 : `arn:aws:datsync:region:account-id:location/location-id`

#### [DescribeLocationS3](#)

操作 : `datsync:DescribeLocationS3`

资源 : `arn:aws:datsync:region:account-id:location/location-id`

#### [DescribeLocationSmb](#)

操作 : `datsync:DescribeLocationSmb`

资源 : `arn:aws:datsync:region:account-id:location/location-id`

#### [DescribeStorageSystem](#)

操作 : `datsync:DescribeStorageSystem`

资源 : `arn:aws:datsync:region:account-id:system/storage-system-id`

操作 : `secretsmanager:DescribeSecret`

资源 : `arn:aws:secretsmanager:region:account-id:secret:datsync!*`

#### [DescribeStorageSystemResourceMetrics](#)

操作 : `datsync:DescribeStorageSystemResourceMetrics`

资源 : `arn:aws:datsync:region:account-id:system/storage-system-id/  
job/discovery-job-id`

#### [DescribeStorageSystemResources](#)

操作 : `datsync:DescribeStorageSystemResources`

资源 : `arn:aws:datsync:region:account-id:system/storage-system-id/job/discovery-job-id`

### [DescribeTask](#)

操作 : `datsync:DescribeTask`

资源 : `arn:aws:datsync:region:account-id:task/task-id`

### [DescribeTaskExecution](#)

操作 : `datsync:DescribeTaskExecution`

资源 : `arn:aws:datsync:region:account-id:task/task-id/execution/exec-id`

### [GenerateRecommendations](#)

操作 : `datsync:GenerateRecommendations`

资源 : `arn:aws:datsync:region:account-id:system/storage-system-id/job/discovery-job-id`

### [ListAgents](#)

操作 : `datsync:ListAgents`

资源 : 无

### [ListDiscoveryJobs](#)

操作 : `datsync:ListDiscoveryJobs`

资源 : `arn:aws:datsync:region:account-id:system/storage-system-id`

### [ListLocations](#)

操作 : `datsync:ListLocations`

资源 : 无

### [ListTagsForResource](#)

操作 : `datsync:ListTagsForResource`

资源 :

- `arn:aws:datsync:region:account-id:agent/agent-id`
- `arn:aws:datsync:region:account-id:task/task-id`
- `arn:aws:datsync:region:account-id:location/location-id`

### ListTaskExecutions

操作 : `datsync:ListTaskExecutions`

资源 : `arn:aws:datsync:region:account-id:task/task-id`

### ListTasks

操作 : `datsync:ListTasks`

资源 : 无

### RemoveStorageSystem

操作 : `datsync:RemoveStorageSystem`

资源 : `arn:aws:datsync:region:account-id:system/storage-system-id`

操作 : `secretsmanager>DeleteSecret`

资源 : `arn:aws:secretsmanager:region:account-id:secret:datsync!*`

### StartDiscoveryJob

操作 : `datsync:StartDiscoveryJob`

资源 : `arn:aws:datsync:region:account-id:system/storage-system-id/job/discovery-job-id`

### StopDiscoveryJob

操作 : `datsync:StopDiscoveryJob`

资源 : `arn:aws:datsync:region:account-id:system/storage-system-id/job/discovery-job-id`

### StartTaskExecution

操作 : `datsync:StartTaskExecution`

资源 : `arn:aws:datsync:region:account-id:task/task-id`

## TagResource

操作 : `datasync:TagResource`

资源 :

- `arn:aws:datasync:region:account-id:agent/agent-id`
- `arn:aws:datasync:region:account-id:task/task-id`
- `arn:aws:datasync:region:account-id:location/location-id`

## UntagResource

操作 : `datasync:UntagResource`

资源 :

- `arn:aws:datasync:region:account-id:agent/agent-id`
- `arn:aws:datasync:region:account-id:task/task-id`
- `arn:aws:datasync:region:account-id:location/location-id`

## UpdateAgent

操作 : `datasync:UpdateAgent`

资源 : `arn:aws:datasync:region:account-id:agent/agent-id`

## UpdateDiscoveryJob

操作 : `datasync:UpdateDiscoveryJob`

资源 : `arn:aws:datasync:region:account-id:system/storage-system-id/job/discovery-job-id`

## UpdateLocationHdfs

操作 : `datasync:UpdateLocationHdfs`

资源 :

- `arn:aws:datasync:region:account-id:agent/agent-id`
- `arn:aws:datasync:region:account-id:location/location-id`

## UpdateLocationNfs

操作 : `datasync:UpdateLocationNfs`

资源 : `arn:aws:datsync:region:account-id:location/location-id`

### [UpdateLocationObjectStorage](#)

操作 : `datsync:UpdateLocationObjectStorage`

资源 :

- `arn:aws:datsync:region:account-id:agent/agent-id`
- `arn:aws:datsync:region:account-id:location/location-id`

### [UpdateLocationSmb](#)

操作 : `datsync:UpdateLocationSmb`

资源 :

- `arn:aws:datsync:region:account-id:agent/agent-id`
- `arn:aws:datsync:region:account-id:location/location-id`

### [UpdateStorageSystem](#)

操作 : `datsync:UpdateStorageSystem`

资源 :

- `arn:aws:datsync:region:account-id:agent/agent-id`
- `arn:aws:datsync:region:account-id:system/storage-system-id`

### [UpdateTask](#)

操作 : `datsync:UpdateTask`

资源 : `arn:aws:datsync:region:account-id:task/task-id`

### [UpdateTaskExecution](#)

操作 : `datsync:UpdateTaskExecution`

资源 : `arn:aws:datsync:region:account-id:task/task-id/execution/exec-id`

## 使用记录AWSDataSync API 调用AWS CloudTrail

AWSDataSync与提供用户AWS CloudTrail、角色或服务在中执行的操作记录的AWS服务集成AWSDataSync。CloudTrail将的所有 API 调用AWSDataSync作为事件捕获。捕获的调用包括来自AWSDataSync控制台和代码的 API 操作代码的AWSDataSync API 操作调用。



如果您创建跟踪，则可以使CloudTrail事件持续传送到 Amazon S3 存储桶（包括的事件）AWSDataSync。如果您不配置跟踪记录，则仍可在 CloudTrail 控制台的 Event history（事件历史记录）中查看最新事件。通过使用收集的信息CloudTrail，您可以确定向发出了什么请求AWSDataSync、发出请求的 IP 地址、何人发出的请求、请求的发出时间以及其他详细信息。

要了解有关 CloudTrail 的更多信息，请参阅 [AWS CloudTrail 用户指南](#)。

## 处理中的DataSync信息CloudTrail

在您创建 AWS 账户 账户时，即针对该账户启用了 CloudTrail。当中发生活动时，该活动将记录在 AWSDataSync事件中，并与其他AWS服务CloudTrail事件一同保存在 Event history（事件历史记录）中。您可以在 AWS 账户 中查看、搜索和下载最新事件。有关更多信息，请参阅[使用 CloudTrail 事件历史记录查看事件](#)。

要持续记录您的AWS 账户事件（包括的事件）AWSDataSync，请创建跟踪。通过跟踪CloudTrail，可将日志文件传送至 Amazon S3 存储桶。预设情况下，在控制台中创建跟踪时，此跟踪应用于所有 AWS 区域。此跟踪记录来自 AWS 分区中的所有 AWS 区域的事件，并将日志文件传送至您指定的 Amazon S3 存储桶。此外，您可以配置其他 AWS 服务，进一步分析在 CloudTrail 日志中收集的事件数据并采取行动。有关更多信息，请参阅下列内容：

- [创建跟踪概览](#)
- [CloudTrail 支持的服务和集成](#)
- [为 CloudTrail 配置 Amazon SNS 通知](#)
- [从多个区域接收 CloudTrail 日志文件和从多个账户中接收 CloudTrail 日志文件](#)

所有 DataSync 操作均由 CloudTrail 记录。（有关更多信息，请参阅 [ADataSync](#) |||||）

例如，对 CreateAgent、CreateTask 和 ListLocations 操作的调用会在 CloudTrail 日志文件中生成条目。

每个事件或日志条目都包含有关生成请求的人员信息。身份信息可帮助您确定以下内容：

- 请求是使用根用户凭证发出的。AWS Identity and Access Management
- 请求是使用角色还是联合身份用户的临时安全凭证发出的。
- 请求是否由其它 AWS 服务发出。

有关更多信息，请参阅 [《CloudTrailAWS CloudTrail用户指南》中的 UserIdentity 元素](#)。

## 了解 DataSync 日志文件条目

跟踪是一种配置，可用于将事件作为日志文件传送到您指定的 Amazon S3 桶。CloudTrail 日志文件包含一个或多个日志条目。一个事件表示来自任何源的一个请求，包括有关所请求的操作、操作的日期和时间、请求参数等方面的信息。CloudTrail 日志文件不是公用 API 调用的有序堆栈跟踪，因此它们不会按任何特定顺序显示。

下面的示例显示了一个 CloudTrail 日志条目，该条目说明了 CreateTask 操作。

```
{
  "eventVersion": "1.05",
  "userIdentity": {
    "type": "IAMUser",
    "principalId": "AIDAJ0ERGY7LS5PKXTMX0",
    "arn": "arn:aws:iam::123456789012:user/user1",
    "accountId": "123456789012",
    "accessKeyId": "access key",
    "userName": "user1",
    "sessionContext": {
      "attributes": {
        "mfaAuthenticated": "false",
        "creationDate": "2018-12-13T14:56:46Z"
      }
    },
    "invokedBy": "signin.amazonaws.com"
  },
  "eventTime": "2018-12-13T14:57:02Z",
  "eventSource": "datasync.amazonaws.com",
  "eventName": "CreateTask",
  "awsRegion": "ap-southeast-1",
  "sourceIPAddress": "12.345.123.45",
  "userAgent": "signin.amazonaws.com",
  "requestParameters": {
    "cloudWatchLogGroupArn": "arn:aws:logs:ap-southeast-1:123456789012:log-group:MyLogGroup",
    "name": "MyTask-NTIzMzY1",
    "tags": [],
    "destinationLocationArn": "arn:aws:datasync:ap-southeast-1:123456789012:location/loc-020c33c5d9966f40a",
    "options": {
      "bytesPerSecond": -1,
      "verifyMode": "POINT_IN_TIME_CONSISTENT",
      "uid": "INT_VALUE",
```

```
        "posixPermissions": "PRESERVE",
        "mtime": "PRESERVE",
        "gid": "INT_VALUE",
        "preserveDevices": "NONE",
        "preserveDeletedFiles": "REMOVE",
        "atime": "BEST_EFFORT"
    },
    "sourceLocationArn": "arn:aws:datasync:ap-southeast-1:123456789012:location/
loc-04aaa9c609812135d"
},
"responseElements": {
    "taskArn": "arn:aws:datasync:ap-southeast-1:123456789012:task/
task-00e5db3f3f41f6cd2"
},
"requestID": "5890e03c-fee7-11e8-8b63-0b409054d4dc",
"eventID": "e5f59b6a-05e6-4412-bd56-440d872e90e9",
"eventType": "AwsApiCall",
"recipientAccountId": "123456789012"
}
```

## AWS DataSync 的合规性验证

作为多个 AWS DataSync 合规性计划的一部分，第三方审核员将评估 AWS 的安全性和合规性。其中包括 SOC、PCI、FedRAMP、HIPAA 及其他。

有关特定合规性计划范围内的 AWS 服务的列表，请参阅[合规性计划范围内的 AWS 服务](#)。有关一般信息，请参阅[AWS 合规性计划](#)。

您可以使用 AWS Artifact 下载第三方审计报告。有关更多信息，请参阅[下载 AWS Artifact 中的报告](#)。

您在使用 DataSync 时的合规性责任由您的数据的敏感性、您公司的合规性目标以及适用的法律法规决定。如果您对 DataSync 使用需遵守 HIPAA、PCI 或 FedRAMP 等标准，AWS 提供帮助的资源：

- [安全性与合规性快速入门指南](#) – 这些部署指南讨论了架构注意事项，并提供了在 AWS 上部署基于安全性和合规性的基准环境的步骤。
- [《设计符合 HIPAA 安全性和合规性要求的架构》白皮书](#) – 此白皮书介绍公司如何使用 AWS 创建符合 HIPAA 标准的应用程序。
- [AWS 合规性资源](#) – 此业务手册和指南集合可能适用于您的行业和位置。
- [AWS Config](#) – 此 AWS 服务评估您的资源配置对内部实践、行业指南和法规的遵循情况。

- [AWS Security Hub](#) – 此AWS服务提供了AWS中安全状态的全面视图，可帮助您检查是否符合安全行业标准 and 最佳实践。

## 中的故障恢复能力AWSDataSync

AWS全球基础设施围绕AWS区域和可用区构建。AWS区域提供多个在物理上独立且隔离的可用区，这些可用区通过延迟低、吞吐量高且冗余性高的网络连接在一起。利用可用区，您可以设计和操作在可用区之间无中断地自动实现故障转移的应用程序和数据库。与传统的单个或多个数据中心基础设施相比，可用区具有更高的可用性、容错性和可扩展性。

### Note

如果您要迁移到或迁出的可用区域在运行 DataSync 任务，该任务也将失败。

有关 AWS 区域和可用区的更多信息，请参阅 [AWS 全球基础设施](#)。

## 中的基础设施安全性AWSDataSync

作为托管服务，AWSDataSync 受到AWS全球网络安全程序。有关更多信息，请参阅 [AWS安全、身份和合规性最佳实践](#)。

你用AWS发布的 API 调用通过网络访问 DataSync。客户端必须支持传输层安全性 (TLS) 1.0 或更高版本。建议使用 TLS 1.2 或更高版本。客户端还必须支持具有完全向前保密 (PFS) 的密码套件，例如 Ephemeral Diffie-Hellman (DHE) 或 Elliptic Curve Ephemeral Diffie-Hellman (ECDHE)。大多数现代系统（如 Java 7 及更高版本）都支持这些模式。

此外，必须使用访问密钥 ID 和与 IAM 委托人关联的秘密访问密钥来对请求进行签名。或者，您可以使用 [AWS Security Token Service](#) (AWS STS) 生成临时安全凭证来对请求进行签名。

## AWS DataSync 配额

使用时了解配额和限制AWSDataSync。

### 存储系统、文件和对象限制

下表描述了使用存储系统、文件和对象时存在的限制。DataSync

描述	限制
最大文件路径总长度	4,096 bytes
文件路径组件（文件名、目录或子目录）的最大长度	255 字节
Windows 域名的最大长度	253 个字符
服务器主机名的最大长度	255 个字符
Amazon S3 对象名称的最大长度	1,024 个 UTF-8 字符

### DataSync任务配额

下表描述了特定AWS 账户和中的DataSync任务配额AWS 区域。

资源	配额	你能增加配额吗？
您可以创建的任务的最大数量	100	是
在自管存储或其他云和服务之间传输数据时，每项任务的最大文件、AWS对象或目录数	5000 万	是

**⚠ Important**  
请记住以下有关此配额的内容：

**💡 Tip**  
您可以使用包

资源	配额	你能增加配额吗？
	<ul style="list-style-type: none"><li>• 如果您传输带有前缀的 Amazon S3 对象，则前缀将被视为目录并计入配额。例如，DataSync 将 <code>s3://bucket/foo/bar.txt</code> 视为两个目录（<code>./</code>和<code>./foo/</code>）和一个对象（<code>bar.txt</code>）。</li><li>• 如果您的任务要传输的文件、对象或目录超过 2000 万个，请确保为 DataSync 代理分配至少 64 GB 的 RAM。有关更多信息，请参阅 <a href="#">DataSync 转账的代理额可能需要几天才能使用</a>。</li></ul>	<p>含和排除过滤器创建专注于特定目录的任务，而不是请求增加。有关更多信息，请参阅 <a href="#">筛选传输的数据</a> <a href="#">DataSync</a>。</p>

资源	配额	你能增加配额吗？
在AWS存储服务之间传输数据时每项任务的最大文件、对象或目录数	2500 万	是

**⚠ Important**

如果您传输带有前缀的 Amazon S3 对象，则前缀将被视为目录并计入配额。例如，DataSync 将 `s3://bucket/foo/bar.txt` 视为两个目录（`./`和`./foo/`）和一个对象（`bar.txt`）。

**💡 Tip**

您可以使用包含和排除过滤器创建专注于特定目录的任务，而不是请求增加。有关更多信息，请参阅[筛选传输的数据](#) [DataSync](#)。

资源	配额	你能增加配额吗？
在AWS Snowcone设备DataSync上运行时每项任务的最大文件、对象或目录数	200,000	否
每个任务的最大吞吐量	10 Gbps ( 适用于使用DataSync代理的传输 )	否
	5 Gbps ( 适用于不使用DataSync代理的传输 )	否
任务筛选器中可以包含的最大字符数	102,400 个字符	否
<div style="border: 1px solid #0070C0; border-radius: 10px; padding: 10px; background-color: #E1F5FE;"> <p> <b>Note</b></p> <p>如果您使用的是主DataSync机，则此限制包括您的包含和排除模式中组合的所有字符。</p> </div>		
任务执行历史记录保留的天数	30	否

## DataSync发现配额

下表描述了特定AWS 账户和中的 DataSync Discovery 配额AWS 区域。

资源	配额	你能增加配额吗？
可以与 DataSync Discovery 一起使用的最大存储系统数	10	否
DataSync代理一次可以访问的最大存储系统数	4	否



## 请求提高限额

您可以请求增加某些配DataSync额。增加限额可能需要几天才能生效。

### 请求提高配额

1. 打开 CratS [AWSupport](#) t ( 创建案例 ) 页面，按需登录，然后选择 Crate ( 创建案例 )。
2. 在“创建案例”中，选择“提高服务限额”。
3. 对于 Limit Type (限制类型)，请选择 DataSync。
4. 对于区域，选择您的AWS 区域，对于“限制”，选择要增加的配额。
5. 填写案例描述，然后选择您的首选联系方式。

如果您需要增加其他配额，请单独填写申请。



验证您的安全组是否已设置为允许代理连接到 VPC 终端节点并且您已允许所需的端口。有关所需端口的信息，请参阅[AWS DataSync网络要求](#)。

此外，请检查您的防火墙和路由器设置，并确保它们允许与AWS服务端点通信。有关信息，请参阅[AWS DataSync网络要求](#)。

## 我无法使用 VPC 终端节点激活代理

如果您在激活使用 VPC 终端节点创建的代理时遇到问题，请针对您的 VPC 终端节点elastic network interface 打开支持渠道。有关支持渠道的信息，请参阅[通过以下方式向您的代理人寻求帮助AWS Support](#)。

## 我不知道我的经纪人怎么了。有人能帮我吗？

如果您遇到无法解决的DataSync代理问题，请参阅[通过以下方式向您的代理人寻求帮助AWS Support](#)。

## 对AWSDataSync传输问题进行故障排除

以下主题描述了AWSDataSync地点和任务的常见问题以及如何解决这些问题。

## DataSync如何配置为使用特定的 NFS 或 SMB 版本来挂载我的文件共享？

对于支持网络文件系统 (NFS) 或服务器消息块 (SMB) 的位置，DataSync默认情况下为您选择协议版本。您也可以使用DataSync控制台或 API 自己指定版本。

要采取的操作 (DataSync控制台)

创建 NFS 或 SMB 位置时，配置DataSync要使用的协议版本。有关更多信息[配置来自 NFS 文件服务器的AWS DataSync传输](#)，请参[配置来自 SMB 文件服务器的AWS DataSync传输](#)阅。

要采取的行动 (DataSyncAPI)

创建或更新 NFS 或 SMB 位置时，请指定Version参数。例如，请参阅 [CreateLocationNfs](#) 或 [CreateLocationSmb](#)。

以下示例AWS CLI命令创建了一个使用 NFS 版本 4.0 进行DataSync装载的 NFS 位置。

```
$ aws datasync create-location-nfs --server-hostname your-server-address \  
  --on-prem-config AgentArns=your-agent-arns \  
  --subdirectory nfs-export-path \  
  --mount-options Version="NFS4_0"
```

以下示例AWS CLI命令创建使用 SMB 版本 3DataSync 装载的 SMB 位置。

```
$ aws datasync create-location-smb --server-hostname your-server-address \  
  --on-prem-config AgentArns=your-agent-arns \  
  --subdirectory smb-export-path \  
  --mount-options Version="SMB3"
```

**错误：** SyncOption值无效。选项： TransferMode ， PreserveDeletedFiles ，  
**值：** 全部 ， 删除。

当您创建或编辑DataSync任务时，选择“传输所有数据”选项并取消选择“保留已删除的文件”选项时，就会出现此错误。当您传输所有数据时，DataSync不会扫描您的目标位置，也不知道要删除什么。

## 我的任务一直因EniNotFound错误而失败

如果您在虚拟私有云 ( VPC ) 中删除任务的网络接口，则会发生问题。如果您的任务已安排或已排队，则如果任务缺少[传输数据所需的网络接口](#)，则该任务将失败。

### 要采取的操作

问题

- 手动重启任务。执行此操作时，DataSync将创建运行任务所需的任何缺失的网络接口。
- 如果您需要清理 VPC 中的资源，请确保不要删除与仍在使用的DataSync任务相关的网络接口。

要查看分配任务的网络接口，请执行以下操作之一：

- 使用该[DescribeTask](#)操作。您可以在SourceNetworkInterfaceArns和DestinationNetworkInterfaceArns响应元素中查看网络接口。
- 在 Amazon EC2 控制台中，搜索您的任务 ID ( 例如task-f012345678abcdef0 ) 以查找其网络接口。
- 考虑不要自动运行任务。这可能包括禁用任务排队或调度 ( 通过DataSync或自定义自动化 ) 。

## 我的任务因DataSync当前不支持服务器端 NFSv4 ID 映射错误而失败

如果您的传输中涉及的文件系统使用 NFS 版本 4 ID 映射 ( DataSync不支持此功能 ) ，则可能会发生此错误。



## 我的任务因输入/输出错误而失败

如果您的存储系统未通过DataSync代理发出的 I/O 请求，您可能会收到输入/输出错误消息。常见原因包括服务器磁盘故障、防火墙配置更改或网络路由器故障。

如果错误涉及 NFS 服务器或 Hadoop 分布式文件系统 (HDFS) 集群，请使用以下步骤解决错误。

### 要采取的行动 (NFS)

首先检查 NFS 服务器的日志和指标，以确定问题是否始于 NFS 服务器。如果是，请解决该问题。

接下来，检查网络配置是否尚未更改。要检查 NFS 服务器的配置是否正确以及是否DataSync可以访问它，请执行以下操作：

1. 在与代理相同的网络子网上设置另一个 NFS 客户端。
2. 在客户端上装载您的共享。
3. 验证客户端可以成功地读取和写入共享。

### 要采取的操作 (HDFS)

确保您的 HDFS 集群允许代理与集群NameNode和DataNode端口通信。在大多数集群中，您可以在以下配置文件中找到集群使用的端口号。

1. 要找到端NameNode口，请查看fs.default或fs.default.name属性下的core-site.xml文件（取决于 Hadoop 发行版）。
2. 要找到端DataNode口，请查看该dfs.datanode.address属性下的hdfs-site.xml文件。

## 我的任务执行处于启动状态，但似乎什么也没发生

当无法指示指定的源代理开始任务时，您的任务执行DataSync可能会停留在启动状态。出现该问题通常是因为，代理已关闭电源或断开网络连接。

### 要采取的操作

确保代理已连接，并且状态为 ONLINE (联机)。如果状态为 OFFLINE (脱机)，则代理未连接。有关如何测试网络连接的信息，请参阅[测试您的代理与的连接AWS](#)。

接下来，确保代理的电源已开启。否则，请打开同步代理电源。

如果代理已开机且任务仍处于启动状态，则代理与之间的网络连接问题DataSync是最有可能的问题。检查网络和防火墙设置，以确保代理可以连接到 DataSync。

如果您执行了这些操作但问题仍未解决，请通过以下方式打开支持频道AWS Support。有关如何打开支持渠道的信息，请参阅[通过以下方式向您的代理人寻求帮助AWS Support](#)。

## 我的任务执行处于准备状态很长时间了

DataSync处于 Preparing (正在准备) 状态的时间取决于源和目标文件系统中的文件数以及这些文件系统的性能。在任务启动时，DataSync 会执行递归目录列表操作以查找源和目标文件系统中的所有文件和文件元数据。这些列表用于查找差异和确定要复制的内容。此过程通常需要几分钟到几小时不等。有关更多信息，请参阅[开始你的AWS DataSync任务](#)：

### 要采取的操作

你不必做任何事情。继续等待“准备中”状态更改为“正在传输”。如果状态仍未更改，请联系AWS Support。

## 我的任务因权限被拒绝错误而失败

如果配置 NFS 服务器并启用 root\_squash 或 all\_squash，并且您的文件没有所有读取访问权限，则可能会显示“permissions denied (权限被拒绝)”错误消息。

### 要采取的操作

要修复此问题，您可以使用 no\_root\_squash 配置 NFS 导出。或者，您可以确保要传输的所有文件的权限允许所有用户具有读取权限。通过执行上述任一操作，将允许代理读取这些文件。要使代理能够访问目录，您还必须启用全部执行访问权限。

要确保可以挂载目录，首先连接到与代理具有相同网络配置的任何计算机。然后运行以下 CLI 命令。

```
mount -t nfs -o nfsvers=<your-nfs-server-version> <your-nfs-server-name>:<nfs-export-path-youspecified> <new-test-folder-on-your-computer>
```

如果您执行了这些操作但问题仍未解决，请联系AWS Support。

## 验证任务需要多长时间？

DataSync 花在 VERIFYING (正在验证) 状态上的时间取决于许多因素。其中包括源和目标文件系统中的文件数、所有文件的总大小以及这些文件系统的性能。默认情况下，将在选项设置中启用



Verification mode (验证模式)。DataSync执行的验证包括对所有文件内容进行 SHA256 校验和以及对所有文件元数据的精确比较。

## 要采取的操作

你不必做任何事情。继续等待“正在验证”状态完成。如果状态仍未更改，请联系AWS Support。

## 转移到另一个 S3 存储桶时，我的任务失败AWS 账户

与同一资源之间的DataSync传输不同AWS 账户，将数据复制到不同资源的 S3 存储桶AWS 账户需要一些额外的步骤。

- 如果您的DataSync任务因与 S3 存储桶权限相关的错误而失败：创建任务时，请确保AWS Management Console使用您在目标 S3 存储桶策略中指定的相同 IAM 角色登录。（注意：这不是让写入 S3 存储桶的 IAM 角色的 IAM 角色。DataSync
- 如果您还将数据复制到另一个存储桶中AWS 区域并出现 S3 终端节点连接错误：在目标 S3 存储桶所在的区域中创建DataSync任务。

有关使用 Amazon S3 进行跨账户转账的完整说明，请参阅以下教程：

- [将数据从本地存储传输到不同的 Amazon S3AWS 账户](#)
- [使用不同的设备将数据从 Amazon S3 传输到 Amazon S3 传输数据AWS 账户](#)

## 从 Google 云存储分区传输时我的任务失败

由于使用 Amazon S3 API与DataSync 与 Google Cloud Storage 通信，因此存在一个限制，如果您尝试复制对象标签，可能会导致DataSync任务失败。以下与该问题相关的消息出现在您的CloudWatch日志中：

[警告] 无法读取文件/*your-bucket* /*your-* object的元数据：S3 获取对象标记失败：在不标记的情况下继续操作

为防止出现这种情况，请在配置任务设置时取消选择“复制对象标签”选项。

## 我的任务开始和结束时间与日志不符

您在DataSync控制台中看到的任务执行的开始和结束时间可能因您在与转移相关的其他地方看到的时间戳而有所不同。这是因为控制台没有考虑任务执行在启动或排队状态下所花费的时间。



例如，您的亚马逊CloudWatch日志可能表明您的任务执行结束时间晚于DataSync控制台中显示的时间。您可能会注意到以下领域存在类似的差异：

- 传输中涉及的文件系统或对象存储系统的日志
- DataSync写入的 Amazon S3 对象的上次修改日期
- 来自DataSync代理的网络流量
- 亚马逊EventBridge活动

## Error: SyncTaskDeletedByUser

在自动执行某些DataSync工作流程时，您可能会意外地看到此错误。例如，也许你有一个脚本在任务执行完成或进入[队列](#)之前删除你的任务。

要修复此问题，请重新配置您的自动化，以使这些类型的操作不会重叠。

## Error: NoMem

您尝试传输的数据集可能太大DataSync。如果您看到此错误，请联系 [AWS Support 中心](#)。

## Error: FsS3UnableToConnectToEndpoint

DataSync无法连接到您的 [Amazon S3 位置](#)。这可能意味着无法访问该位置的 S3 存储桶或该位置的配置不正确。

在解决问题之前，请执行以下操作：

- 检查是否DataSync可以[访问您的 S3 存储桶](#)。
- 使用DataSync控制台或 [DescribeLocationS3](#) 操作确保位置配置正确。

## Error: FsS3HeadBucketFailed

DataSync无法访问您要传输到或传出的 S3 存储桶。使用 Amazon S3 [HeadBucket](#)操作检查是否DataSync有权访问存储桶。

## 我的 Amazon S3 存储成本高于预期AWS DataSync

如果您的 Amazon S3 存储成本高于您在AWS DataSync传输后想象的费用，则可能是由于以下一个或多个原因造成的：

- 向 S3 存储桶传输或从 S3 存储桶传输时，您会产生与 S3 API 请求相关的费用DataSync。
- DataSync使用 Amazon S3 分段上传功能将对象上传到 S3 存储桶。这种方法可能会导致上传未成功完成时产生意想不到的存储费用。
- 可能对您的 S3 存储桶启用了对象版本控制。对象版本控制导致 Amazon S3 存储具有相同名称的对象多个副本。

## 要采取的操作

在这些情况下，您可以执行以下步骤：

- 确保您了解如何DataSync使用 S3 请求以及它们可能如何影响您的存储成本。有关更多信息，请参阅[使用时评估 S3 请求成本 DataSync](#)：
- 如果问题与分段上传有关，请为您的 S3 存储桶配置分段上传策略，以清理未完成的分段上传以降低存储成本。有关更多信息，请参阅AWS博客文章 [S3 生命周期管理更新-Support 分段上传和删除标记](#)。
- 如果问题与对象版本控制有关，请在 S3 存储桶上禁用对象版本控制。
- 如果您需要更多帮助，请联系[AWS Support中心](#)。

# AWS DataSync 教程

这些教程将引导您完成一些现实世界中的场景AWS DataSync。

## 主题

- [教程：将数据从本地存储传输到不同的 Amazon S3AWS 账户](#)
- [教程：将数据从 Amazon S3 传输到 Amazon S3AWS 账户](#)

## 教程：将数据从本地存储传输到不同的 Amazon S3AWS 账户

AWS DataSync与本地存储一起使用时，通常将数据复制到与您的DataSync代理同AWS 账户属的AWS存储服务。但是，在某些情况下，您可能需要将数据传输到其他账户关联的 Amazon S3 存储桶。

### Important

只有当 Amazon S3 是其中一个DataSync位置时，使用本教程中的方法复制数据才有效。AWS 账户

## 概览

在本教程中，您将学习AWS Identity and Access Management (IAM) 和AWS Command Line Interface (AWS CLI) 如何帮助您创建DataSync任务，将数据从本地存储传输到不同的 S3 存储桶AWS 账户。

以下是这种场景的样子：

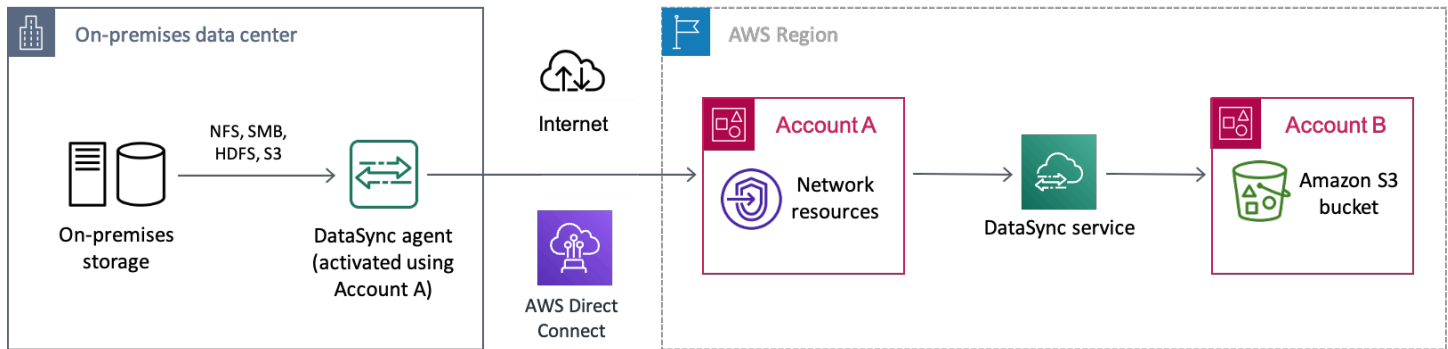
- 账户 A：用于管理网络资源的账户。AWS 账户您激活DataSync代理时使用的终端节点也属于此帐户。

### Note

本教程中的步骤适用于激活代理时使用的[任何类型的终端节点](#)。

- 账户 B：您希望将数据复制到其中的 S3 存储桶。AWS 账户

下图阐明了此方案。



## 先决条件

在开始 IAM 工作以促进跨账户转移之前，如果您还没有这样做，请执行以下操作：

1. [配置您的网络](#)，以便您的本地存储系统可以连接AWS。
2. 使用账户 A [部署和激活DataSync代理](#)。
3. 使用账户 A [为要从中复制数据的本地存储系统创建DataSync源位置](#)。
4. AWS CLI使用账户 A [进行设置](#)。AWS CLI您需要在账户 B 中创建 S3 存储桶的DataSync目标位置。

## 步骤 1：为账户 A DataSync 中创建 IAM 角色

您需要一个DataSync允许写入到账户 B 中的 S3 存储桶的 IAM 角色

当您为存储段创建位置时，DataSync可以自动创建并代入具有访问该存储桶的正确权限的角色。由于您跨账户转移，因此您必须手动创建角色。

有关更多信息，请参阅 IAM 用户指南中的 [AWS 服务（控制台）创建角色](#)。

### 创建 IAM 角色

以可信实体DataSync身份创建角色。

#### 创建 IAM 角色

1. 使用账户AWS Management Console A 登录
2. 通过以下网址打开 IAM 控制台：<https://console.aws.amazon.com/iam/>。
3. 在左侧导航窗格的“访问管理”下，选择“角色”，然后选择“创建角色”。
4. 在“选择可信实体”页面上，对于可信实体类型，选择AWS 服务。

5. 对于用例，DataSync在下拉列表中选择并选择DataSync。选择 Next ( 下一步 )。
6. 在 Add permissions ( 添加权限 ) 页面上，选择 Next ( 下一步 )。
7. 为您的角色命名，然后选择创建角色。

## 将自定义策略附加到 IAM 角色

IAM 角色需要一个DataSync允许写入账户 B 中的 S3 存储桶的策略。

### 将自定义策略附加到 IAM 角色

1. 在 IAM 控制台的角色页面上，搜索您刚刚创建的角色并选择其名称。
2. 在角色的详细信息页面上，选择 Permission ( 权限 ) 选项卡。选择添加权限，然后选择创建内联策略。
3. 选择 JSON 选项卡并执行以下操作：
  - a. 将下面的 JSON 粘贴到策略编辑器中：

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Action": [
        "s3:GetBucketLocation",
        "s3:ListBucket",
        "s3:ListBucketMultipartUploads"
      ],
      "Effect": "Allow",
      "Resource": "arn:aws:s3:::account-b-bucket"
    },
    {
      "Action": [
        "s3:AbortMultipartUpload",
        "s3:DeleteObject",
        "s3:GetObject",
        "s3:ListMultipartUploadParts",
        "s3:PutObject",
        "s3:GetObjectTagging",
        "s3:PutObjectTagging"
      ],
      "Effect": "Allow",
      "Resource": "arn:aws:s3:::account-b-bucket/*"
    }
  ]
}
```

```
    }  
  ]  
}
```

- b. `account-b-bucket` 替换为账户 B 中 S3 存储桶的名称
4. 选择 Review policy ( 查看策略 )。
5. 为您的策略命名，然后选择创建策略。

## 步骤 2：在账户 B 中禁用 S3 存储桶的 ACL

您复制到 S3 存储桶的所有数据都必须属于账户 B。为确保账户 B 是数据的所有者，请禁用存储段的访问控制列表 (ACL)。有关更多信息，请参阅《Amazon S3 用户指南》中的[控制存储桶的对象所有权利禁用 ACL](#)。

### 禁用 S3 存储桶的 ACL

1. 在中 AWS Management Console，切换到账户 B。
2. 通过以下网址打开 Simple Storage Service ( Amazon S3 ) 控制台：<https://console.aws.amazon.com/s3/>。
3. 在左侧导航窗格中，选择桶。
4. 在存储段列表中，选择要将数据传输到的 S3 存储桶。
5. 在存储桶的详细信息页面上，选择权限选项卡。
6. 在 Object Ownership ( 对象所有权 ) 下方，请选择 Edit ( 编辑 )。
7. 如果尚未选择 ACL，请选择 ACL 已禁用 ( 推荐 ) 选项。
8. 选择保存更改。

## 步骤 3：更新账户 B 中的 S3 存储桶策略

在账户 B 中，修改 S3 存储桶策略以允许访问您为账户 A DataSync 中创建的 IAM 角色

更新后的政策 ( 在以下说明中提供给您 ) 包括两项主要内容：

- 第一个委托人指定了您在步骤 1 中创建的账户 A 中的 IAM 角色。此角色允许写入 DataSync 到账户 B 中的 S3 存储桶
- 第二个委托人指定了账户 A 中的 IAM 角色，您使用该角色访问控制台或 AWS CLI。在步骤 4 中，您将在创建 S3 存储桶的目标位置时使用此角色。

## 更新 S3 存储桶策略

1. 当您仍在 S3 控制台中并使用账户 B 时，选择要将数据复制到的 S3 存储桶。
2. 在存储桶的详细信息页面上，选择权限选项卡。
3. 在存储桶策略下，选择编辑，然后执行以下操作来修改您的 S3 存储桶策略：
  - a. 更新编辑器中的内容以包括以下政策声明：

```
{
  "Version": "2008-10-17",
  "Statement": [
    {
      "Sid": "DataSyncCreateS3LocationAndTaskAccess",
      "Effect": "Allow",
      "Principal": {
        "AWS": "arn:aws:iam::account-a-id:role/name-of-datasync-role"
      },
      "Action": [
        "s3:GetBucketLocation",
        "s3:ListBucket",
        "s3:ListBucketMultipartUploads",
        "s3:AbortMultipartUpload",
        "s3>DeleteObject",
        "s3:GetObject",
        "s3:ListMultipartUploadParts",
        "s3:PutObject",
        "s3:GetObjectTagging",
        "s3:PutObjectTagging"
      ],
      "Resource": [
        "arn:aws:s3::account-b-bucket",
        "arn:aws:s3::account-b-bucket/*"
      ]
    },
    {
      "Sid": "DataSyncCreateS3Location",
      "Effect": "Allow",
      "Principal": {
        "AWS": "arn:aws:iam::account-a-id:role/name-of-your-role"
      },
      "Action": "s3:ListBucket",
      "Resource": "arn:aws:s3::account-b-bucket"
    }
  ]
}
```

```
]
}
```

- b. *account-a-id* 替换为账户 A 的 AWS 账户号码。
  - c. *name-of-datasync-role* 替换为您在账户 A DataSync 中创建的 IAM 角色 ( 返回步骤 1 ) 。
  - d. *account-b-bucket* 替换为账户 B 中 S3 存储桶的名称
  - e. *name-of-your-role* 替换为您用于访问控制台的 IAM 角色或 AWS CLI 账户 A。
4. 选择保存更改。

## 步骤 4 : 为 S3 存储桶创建 DataSync 目标位置

为 S3 存储桶创建位置后，即可运行 DataSync 任务。但是，该 DataSync 主机不支持在不同的账户中创建地点。必须先使用创建位置，AWS CLI 然后才能运行任务。

使用 CLI 创建 DataSync 地点

1. 打开 终端。
2. 确保您的 CLI 配置文件配置为使用账户 A。
3. 复制以下命令：

```
aws datasync create-location-s3 \  
  --s3-bucket-arn arn:aws:s3:::account-b-bucket \  
  --s3-config '{"BucketAccessRoleArn":"arn:aws:iam::account-a-id:role/name-of-  
datasync-role"}'
```

4. *account-b-bucket* 替换为账户 B 中 S3 存储桶的名称
5. *account-a-id* 替换为账户 A 的 AWS 账户号码。
6. *name-of-datasync-role* 替换为您在账户 A DataSync 中创建的 IAM 角色 ( 返回步骤 1 ) 。
7. 运行 命令。

如果该命令返回与以下 DataSync 位置相似的位置 ARN，则表示您成功创建了位置：

```
{  
  "LocationArn": "arn:aws:datasync:us-east-2:123456789012:location/loc-  
  abcdef01234567890"  
}
```



- 在中切换回账户 A 的 AWS Management Console。
- 通过 <https://console.aws.amazon.com/datasync/> 打开 DataSync 主机。
- 在左侧导航窗格中，选择位置。

您可以看到刚刚使用 CLI 创建的账户 B 中的 S3 存储桶的位置。

## 步骤 5：创建并启动 DataSync 任务

在移动数据之前，让我们回顾一下你到目前为止所做的事情：

- 在账户 A 中部署并激活了您的 DataSync 代理，以便代理可以从您的自管理存储系统读取数据并与其通信 AWS。
- 在账户 A 中创建了一个 IAM 角色，以便 DataSync 可以将数据写入账户 B 中的 S3 存储桶。
- 在账户 B 中配置您的 S3 存储桶以确保您的 DataSync 任务正常运行。
- 在账户 A 中创建了您的 DataSync 来源和目的地地点。

### 创建和启动 DataSync 任务

- 在账户 A 中仍使用 DataSync 控制台时，展开左侧导航窗格中的“数据传输”，然后选择“任务和创建任务”。

#### Note

您必须使用在步骤 3 中的 S3 存储桶策略中为账户 A 指定的 IAM 角色登录控制台。

- 
- 
- 在“配置源位置”页面上，选择“选择现有位置”。选择要从中复制数据的源位置（您的本地存储），然后选择下一步。
- 在配置目标位置页面上，选择选择现有位置。选择要将数据复制到的目标位置（账户 B 中的 S3 存储桶），然后选择下一步。
- 在配置设置页面上，为任务命名。根据需要配置其他设置，例如指定 Amazon CloudWatch 日志组。选择下一步。
- 在“查看”页面上，查看您的设置并选择“创建任务”。
- 在任务的详细信息页面上，选择 Start（开始），然后选择以下选项之一：
  - 要在不修改的情况下运行任务，请选择“以默认值启动”。

- 要在运行任务之前对其进行修改，请选择“使用替代选项开始”。

任务完成后，您将在 S3 存储桶中看到来自本地存储的数据。现在，您可以从账户 B 访问存储桶数据。

## 相关资源

有关您在本教程中执行的操作的更多信息，请参阅以下主题：

- [为AWS 服务 \(控制台\) 创建角色](#)
- [修改角色信任策略 \(控制台\)](#)
- [使用 Amazon S3 控制台添加存储桶策略](#)
- [使用创建 S3 位置AWS CLI](#)

## 教程：将数据从 Amazon S3 传输到 Amazon S3AWS 账户

使用AWS DataSync，您可以在属于不同的 Amazon S3 存储桶之间移动数据AWS 账户。

### Important

AWS 账户使用本教程中的方法复制数据仅适用于 Amazon S3。

## 概览

在本教程中，您将学习AWS Identity and Access Management (IAM) 和AWS Command Line Interface (AWS CLI) 如何帮助您创建DataSync任务，将数据从 Amazon S3 传输到另一个存储桶中的另一个 S3 存储桶AWS 账户。

### Tip

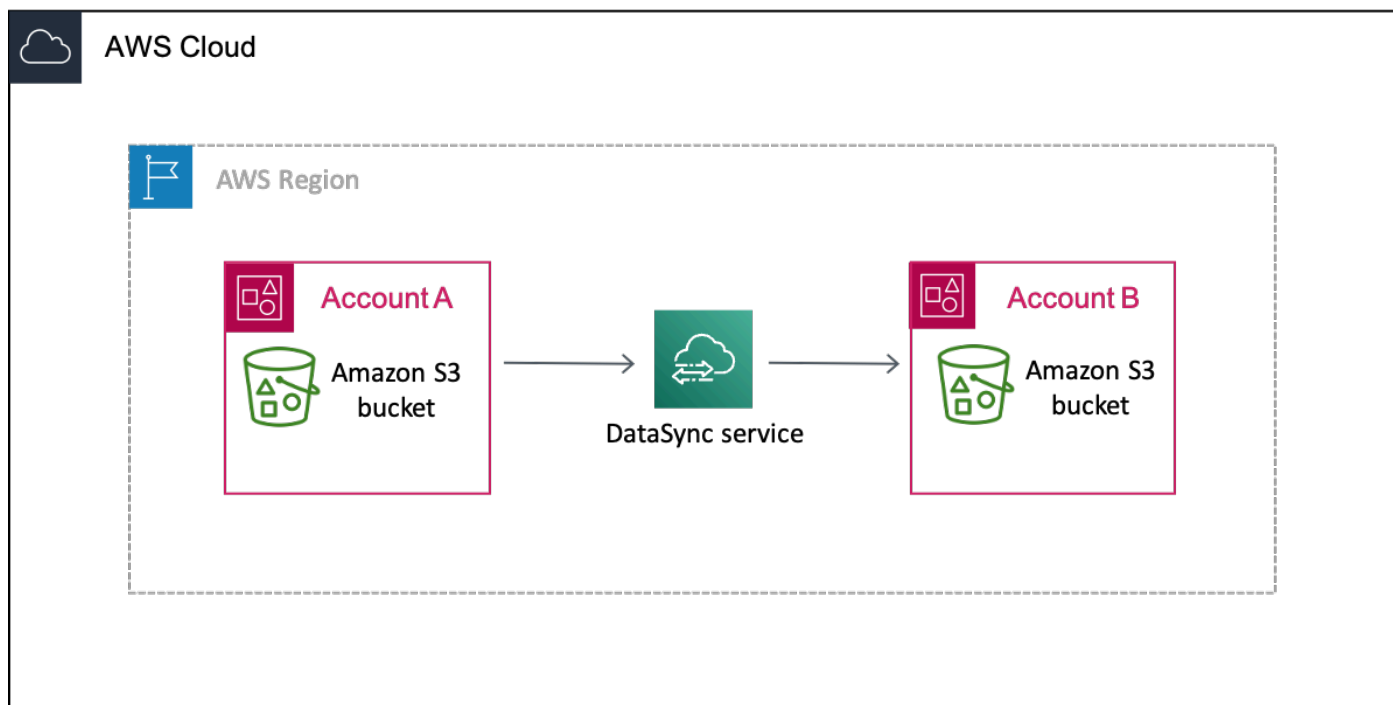
如果您的 S3 存储桶也位于不同的存储桶中，请按照本教程进行操作AWS 区域。除了一些额外的步骤外，该过程基本相同。但是请记住，[默认情况下，这DataSync不支持对区域](#)进行此类传输。

以下是这种场景的样子：

- 账户 A：您用于管理要复制数据的 S3 存储桶。AWS 账户
- 账户 B：您用于管理要将数据复制到其中的 S3 存储桶。AWS 账户

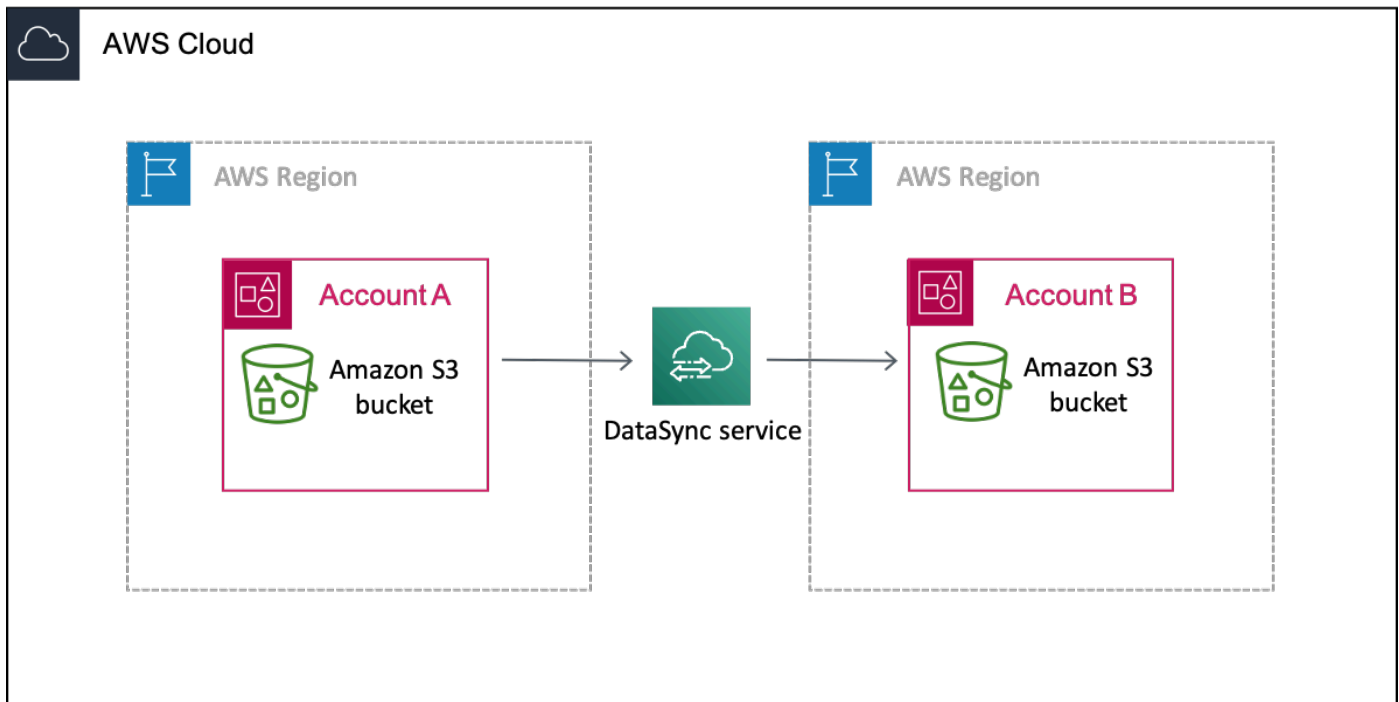
## Transfers across accounts

下图说明了将数据从 S3 存储桶复制到另一个 S3 存储桶的场景AWS 账户。



## Transfers across accounts and Regions

下图说明了将数据从 S3 存储桶复制到位于不同AWS 账户区域的另一个 S3 存储桶的场景。



## 先决条件

在开始 IAM 工作以促进跨账户转移之前，如果您还没有这样做，请执行以下操作：

1. 确定要复制的对象数量。使用 [Amazon S3 存储镜头](#)来计算您的存储桶中有多少对象。

### Tip

在 S3 存储段之间传输时，每个任务DataSync不能复制超过 2,500 万个对象。如果您的存储桶包含超过 2,500 万个对象，我们推荐几个选项：

- 使用包含不超过 2,500 万个@@ [对象的前缀来组织](#)对象。然后，您可以为每个前缀创建单独的DataSync任务。
- [筛选传输的数据](#)DataSync。

2. 使用账户 A 为要从中复制数据的 S3 存储桶创建DataSync源位置。
3. AWS CLI使用账户 A [进行设置](#)。AWS CLI您需要在账户 B 中创建 S3 存储桶的DataSync目标位置。

## 步骤 1：DataSync在账户 A 中创建 IAM 角色

您需要一个 IAM 角色，以DataSync允许写入账户 B 中的 S3 存储桶

当您为存储段创建位置时，DataSync可以自动创建并代入具有访问该存储桶的正确权限的角色。由于您要跨账户转移，因此必须手动创建角色。

有关更多信息，请参阅 IAM 用户指南中的 [AWS 服务（控制台）创建角色](#)。

## 创建 IAM 角色

以可信实体DataSync身份创建角色。

### 创建 IAM 角色

1. 使用账户AWS Management Console A 登录
2. 通过以下网址打开 IAM 控制台：<https://console.aws.amazon.com/iam/>。
3. 在左侧导航窗格的“访问管理”下，选择“角色”，然后选择“创建角色”。
4. 在“选择可信实体”页面上，对于可信实体类型，选择AWS 服务。
5. 对于用例，DataSync在下拉列表中选择并选择DataSync。选择 Next（下一步）。
6. 在 Add permissions（添加权限）页面上，选择 Next（下一步）。
7. 为您的角色命名，然后选择创建角色。

## 将自定义策略附加到 IAM 角色

IAM 角色需要一个DataSync允许写入账户 B 中的 S3 存储桶的策略。

### 将自定义策略附加到 IAM 角色

1. 在 IAM 控制台的角色页面上，搜索您刚刚创建的角色并选择其名称。
2. 在角色的详细信息页面上，选择 Permissions（权限）选项卡。选择添加权限，然后选择创建内联策略。
3. 选择 JSON 选项卡并执行以下操作：
  - a. 将下面的 JSON 粘贴到策略编辑器中：

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Action": [
```

```

        "s3:GetBucketLocation",
        "s3:ListBucket",
        "s3:ListBucketMultipartUploads"
    ],
    "Effect": "Allow",
    "Resource": "arn:aws:s3:::account-b-bucket"
},
{
    "Action": [
        "s3:AbortMultipartUpload",
        "s3:DeleteObject",
        "s3:GetObject",
        "s3:ListMultipartUploadParts",
        "s3:PutObject",
        "s3:GetObjectTagging",
        "s3:PutObjectTagging"
    ],
    "Effect": "Allow",
    "Resource": "arn:aws:s3:::account-b-bucket/*"
}
]
}

```

b. *account-b-bucket* 替换为账户 B 中 S3 存储桶的名称

4. 选择 Review policy ( 查看策略 )。
5. 为您的策略命名，然后选择创建策略。

## 步骤 2：在账户 B 中禁用 S3 存储桶的 ACL

您复制到 S3 存储桶的所有数据都必须属于账户 B。为确保账户 B 是数据的所有者，请禁用存储段的访问控制列表 (ACL)。有关更多信息，请参阅《Amazon S3 用户指南》中的 [控制存储桶的对象所有权利和禁用 ACL](#)。

### 禁用 S3 存储桶的 ACL

1. 在中 AWS Management Console，切换到账户 B。
2. 通过以下网址打开 Simple Storage Service ( Amazon S3 ) 控制台：<https://console.aws.amazon.com/s3/>。
3. 在左侧导航窗格中，选择桶。
4. 在存储段列表中，选择要将数据传输到的 S3 存储桶。

5. 在存储桶的详细信息页面上，选择权限选项卡。
6. 在 Object Ownership (对象所有权) 下方，请选择 Edit (编辑)。
7. 如果尚未选择 ACL，请选择 ACL 已禁用 (推荐) 选项。
8. 选择保存更改。

## 步骤 3：更新账户 B 中的 S3 存储桶策略

在账户 B 中，修改 S3 存储桶策略以允许访问您为账户 A DataSync 中创建的 IAM 角色

更新后的政策 (在以下说明中提供给您) 包括两项主要内容：

- 第一个委托人指定了您在步骤 1 中创建的账户 A 中的 IAM 角色。此角色 DataSync 允许写入账户 B 中的 S3 存储桶
- 第二个委托人指定了账户 A 中的 IAM 角色，您使用该角色访问控制台或 AWS CLI。在步骤 4 中，您将在创建 S3 存储桶的目标位置时使用此角色。

### 更新 S3 存储桶策略

1. 当您仍在 S3 控制台中并使用账户 B 时，选择要将数据复制到的 S3 存储桶。
2. 在存储桶的详细信息页面上，选择权限选项卡。
3. 在存储桶策略下，选择编辑，然后执行以下操作来修改您的 S3 存储桶策略：
  - a. 更新编辑器中的内容以包括以下政策声明：

```
{
  "Version": "2008-10-17",
  "Statement": [
    {
      "Sid": "DataSyncCreateS3LocationAndTaskAccess",
      "Effect": "Allow",
      "Principal": {
        "AWS": "arn:aws:iam::account-a-id:role/name-of-datasync-role"
      },
      "Action": [
        "s3:GetBucketLocation",
        "s3:ListBucket",
        "s3:ListBucketMultipartUploads",
        "s3:AbortMultipartUpload",
        "s3:DeleteObject",

```

```

        "s3:GetObject",
        "s3:ListMultipartUploadParts",
        "s3:PutObject",
        "s3:GetObjectTagging",
        "s3:PutObjectTagging"
    ],
    "Resource": [
        "arn:aws:s3:::account-b-bucket",
        "arn:aws:s3:::account-b-bucket/*"
    ]
},
{
    "Sid": "DataSyncCreateS3Location",
    "Effect": "Allow",
    "Principal": {
        "AWS": "arn:aws:iam::account-a-id:role/name-of-your-role"
    },
    "Action": "s3:ListBucket",
    "Resource": "arn:aws:s3:::account-b-bucket"
}
]
}

```

- b. *account-a-id* 替换为账户 A 的 AWS 账户号码。
  - c. *name-of-datasync-role* 替换为您在账户 A DataSync 中创建的 IAM 角色 ( 返回步骤 1 )。
  - d. *account-b-bucket* 替换为账户 B 中 S3 存储桶的名称
  - e. *name-of-your-role* 替换为您用于访问控制台的 IAM 角色或 AWS CLI 账户 A。
4. 选择保存更改。

## 步骤 4 : 为 S3 存储桶创建 DataSync 目标位置

为 S3 存储桶创建位置后，即可运行 DataSync 任务。但是，该 DataSync 主机不支持在不同的账户中创建地点。必须先使用创建位置，AWS CLI 然后才能运行任务。

使用 CLI 创建 DataSync 地点

1. 打开 终端。
2. 确保您的 CLI 配置文件配置为使用账户 A。
3. 复制以下命令：



```
aws datasync create-location-s3 \  
  --s3-bucket-arn arn:aws:s3:::account-b-bucket \  
  --s3-config '{"BucketAccessRoleArn":"arn:aws:iam::account-a-id:role/name-of-  
datasync-role"}'
```

4. *account-b-bucket* 替换为账户 B 中 S3 存储桶的名称
5. *account-a-id* 替换为账户 A 的 AWS 账户号码。
6. *name-of-datasync-role* 替换为您在账户 A DataSync 中创建的 IAM 角色（返回步骤 1）。
7. 如果账户 B 中的存储桶与账户 A 中的存储桶位于不同的区域，请在命令末尾添加 `--region` 选项以指定账户 B 存储桶所在的区域。例如，`--region us-west-1`。
8. 运行命令。

如果该命令返回与以下 DataSync 位置相似的位置 ARN，则表示您成功创建了位置：

```
{  
  "LocationArn": "arn:aws:datasync:us-east-2:123456789012:location/loc-  
  abcdef01234567890"  
}
```

9. 在中切换回账户 A AWS Management Console。
10. 通过 <https://console.aws.amazon.com/datasync/> 打开 DataSync 主机。
11. 在左侧导航窗格中，选择位置。

您可以看到刚刚使用 CLI 创建的账户 B 中的 S3 存储桶的位置。

## 步骤 5：创建并启动 DataSync 任务

在移动数据之前，让我们回顾一下你到目前为止所做的事情：

- 在账户 A 中创建了一个 IAM 角色，以便 DataSync 可以将数据写入账户 B 中的 S3 存储桶。
- 在账户 B 中配置您的 S3 存储桶以确保您的 DataSync 任务正常运行。
- 在账户 A 中创建了您的 DataSync 来源和目的地地点。

### 创建和启动 DataSync 任务

1. 在账户 A 中仍使用 DataSync 控制台时，展开左侧导航窗格中的“数据传输”，然后选择“任务和创建任务”。

**Note**

您必须使用在步骤 3 中的 S3 存储桶策略中为账户 A 指定的 IAM 角色登录控制台。

2. 如果账户 B 中的存储段与账户 A 中的存储段位于不同的区域，请在导航窗格中选择账户 B 存储段的区域。

您必须从目标位置的区域（在本例中为账户 B 存储桶）启动 DataSync 任务，以避免连接错误。

3. 在“配置源位置”页面上，选择“选择现有位置”。
4. 要跨区域转移，请选择 Account A 存储桶所在的区域。
5. 选择要从中复制数据的源位置（账户 A 中的 S3 存储桶），然后选择下一步。
6. 在配置目标位置页面上，选择选择现有位置。选择要将数据复制到的目标位置（账户 B 中的 S3 存储桶），然后选择下一步。
7. 在配置设置页面上，为任务命名。根据需要配置其他设置，例如指定 AmazonCloudWatch 日志组。选择下一步。
8. 在“查看”页面上，查看您的设置并选择“创建任务”。
9. 在任务的详细信息页面上，选择 Start（开始），然后选择以下选项之一：
  - 要在不修改的情况下运行任务，请选择“以默认值启动”。
  - 要在运行任务之前对其进行修改，请选择“使用替代选项开始”。

任务完成后，检查账户 B 中的 S3 存储桶。您应该会看到账户 A 存储桶中的数据。

## 相关资源

有关您在本教程中执行的操作的更多信息，请参阅以下主题：

- [为 AWS 服务（控制台）创建角色](#)
- [修改角色信任策略（控制台）](#)
- [使用 Amazon S3 控制台添加存储桶策略](#)
- [使用创建 S3 位置 AWS CLI](#)

## AWS DataSync 的其他资源

在本节中，您可以找到 AWS DataSync 的更多信息和资源。

### 主题

- [从自我管理的存储阵列传输数据](#)
- [其他AWS DataSync 用例](#)
- [的开源组件AWS DataSync](#)

## 从自我管理的存储阵列传输数据

您可能希望将数据从自管企业存储阵列传输到 Amazon EFS。在这种情况下，在将文件从网络文件系统 (NFS) 或服务器消息块 (SMB) 文件共享传输到 Amazon EFS 时，其他应用程序可能会修改源文件系统中的文件。

要确保 DataSync 成功执行传输并进行完整一致性验证，我们建议源位置指向一个只读快照。该设置可确保在传输文件时无法修改源位置中的文件，并确保验证正常工作。

有关如何在企业存储阵列中拍摄快照的信息，请参阅以下内容之一：

- EMC VNX：[如何创建 VNX 快照并将其附加到服务器](#)
- NetApp：[快照管理](#)
- HPE 3PAR：[创建虚拟卷快照](#)
- HDS：[Hitachi 写入时复制快照用户指南](#)

## 其他AWS DataSync 用例

在本节中，您可以找到有关大多数用户AWS DataSync 不常见的用例的信息。

### 主题

- [朝相反的方向传输文件](#)
- [使用多个任务写入同一 Amazon S3 存储桶](#)
- [允许 DataSync 访问受限的 Amazon S3 存储桶](#)

## 朝相反的方向传输文件

向相反方向传输数据允许在工作流程中活动应用程序在不同位置之间移动。AWS DataSync 不支持多个活动应用程序同时写入两个位置的工作流程。使用以下过程中的步骤进行配置，以 DataSync 向相反方向传输数据。

### 配置为 DataSync 向相反方向传输数据

1. 创建一个位置并将其命名为位置 A。
2. 创建第二个位置并将其命名为位置 B。
3. 创建一个任务，将其命名为任务 A-B，然后将位置 A 配置为源位置，将位置 B 配置为目标位置。
4. 创建第二个任务，将其命名为任务 B-A，然后将位置 B 配置为源位置，将位置 A 配置为目标位置。
5. 要使用来自位置 A 的数据更新位置 B，请运行任务 A-B。

要使用来自位置 B 的数据更新位置 A，请运行任务 B-A。

不要同时运行这两个任务。DataSync 可以定期朝相反的方向传输文件。但是，它不支持多个活动应用程序同时写入位置 A 和位置 B 的工作流。

## 使用多个任务写入同一 Amazon S3 存储桶

在某些使用案例中，您可能希望将不同的任务写入同一 Amazon S3 存储桶。在这种情况下，您可在 S3 存储桶中为每个任务创建不同的文件夹。此方法可防止任务之间的文件名冲突，同时意味着您可以为每个文件夹设置不同的权限。

例如，您可能有三个任务：task1、task2 和 task3 写入到名为 MyBucket 的 S3 存储桶。

您可在存储桶中创建三个文件夹：

```
s3://MyBucket/task1
```

```
s3://MyBucket/task2
```

```
s3://MyBucket/task3
```

对于每个任务，您可选择对应该任务的 MyBucket 中的文件夹作为目标，并为三个文件夹的每一个设置不同的权限。

## 允许 DataSync 访问受限的 Amazon S3 存储桶

在某些情况下，您可能希望限制对 Amazon S3 存储桶的访问权限。您可以编辑 S3 存储桶策略，以便在运行任务时仍 DataSync 可以访问存储桶。

### 允许 DataSync 访问受限的 S3 存储桶

1. 复制以下示例策略。

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Deny",
      "Principal": "*",
      "Action": "s3:*",
      "Resource": [
        "arn:aws:s3:::bucket-name",
        "arn:aws:s3:::bucket-name/*"
      ],
      "Condition": {
        "StringNotLike": {
          "aws:userid": [
            "datasync-role-id:*",
            "your-role-id"
          ]
        }
      }
    }
  ]
}
```

2. 在示例策略中，替换以下值：

- 存储@@ ###：您限制访问的 S3 存储桶的名称。
- *datasync-role-id*：DataSync 用于访问 S3 存储桶的 IAM 角色的 ID。运行以下 AWS CLI 命令以获取 IAM 角色 ID：

```
aws iam get-role --role-name datasync-iam-role-name
```

在输出中，查找 RoleId 值：

```
"RoleId": "ANPAJ2UCCR6DPCEXAMPLE"
```

- *your-role-id* : 您用来为 S3 存储桶创建 DataSync 位置的 IAM 角色的 ID。运行以下命令以获取 IAM 角色 ID :

```
aws iam get-role --role-name your-iam-role-name
```

在输出中，查找UserId值：

```
"RoleId": "AIDACKCEVSQ6C2EXAMPLE"
```

3. 将此策略添加到 S3 存储桶策略中。有关更多信息，请参阅 Amazon S3 用户指南的如何[编辑存储桶策略](#)。

更新 S3 存储桶策略后，您必须在策略中为需要访问 S3 存储桶的用户添加其他 IAM 角色或用户。

## 的开源组件AWS DataSync

要查看使用的开源组件AWS DataSync，请下载以下链接：

- [datasync-open-source-components.zip](#)

# 将 AWS Command Line Interface 与 AWS DataSync 结合使用

在本节中，您可以找到对 AWS DataSync 使用 AWS Command Line Interface (AWS CLI) 命令的示例。您可以使用这些命令创建代理，创建源和目标位置以及运行任务。

在开始之前，我们建议您阅读[AWS DataSync 工作原理](#)以了解中使用的组件和术语 DataSync 以及该服务的工作原理。我们还建议您阅读[IAM 客户托管策略 AWS DataSync](#)以了解 DataSync 所需的 AWS Identity and Access Management (IAM) 权限。

在使用 AWS CLI 命令之前，请安装 AWS CLI。有关如何安装的信息 AWS CLI，请参阅 [《AWS Command Line Interface 用户指南》AWS Command Line Interface 中的安装](#)。安装后 AWS CLI，您可以使用 `help` 命令查看 DataSync 操作以及与之相关的参数。

要查看可用操作，请输入以下命令。

```
aws datasync help
```

要查看与特定操作关联的参数，请输入以下命令。

```
aws datasync operation help
```

有关“的更多信息 AWS CLI，请参阅 [什么是 AWS Command Line Interface ?](#)

主题

- [使用创建 AWS DataSync 代理 AWS CLI](#)
- [使用创建 AWS DataSync 地点 AWS CLI](#)
- [使用创建 AWS DataSync 任务 AWS CLI](#)
- [使用... 启动 AWS DataSync 任务 AWS CLI](#)
- [筛选 AWS DataSync 资源](#)

有关支持的终端节点 AWS 区域和终端节点的信息，请参阅 [AWS DataSync 终端节点和配额](#)。

有关 DataSync Amazon 资源名称 (ARN) 值的信息，请参阅 [Amazon 资源名称](#)。

## 使用创建AWS DataSync代理AWS CLI

要访问您的自管理存储，请先部署并激活AWS DataSync代理。激活过程可将您的代理与您关联AWS账户。在同一AWS存储服务内的存储服务之间传输时不需要代理AWS账户。要设置两者之间的数据传输AWS服务，请参见[使用创建AWS DataSync地点AWS CLI](#)。

DataSync代理可以通过公共服务终端节点、联邦信息处理标准 (FIPS) 终端节点和 Amazon VPC 终端节点传输数据。有关更多信息，请参阅[创建代AWS DataSync理](#)：

### Note

当您将代理配置为使用 Amazon VPC 终端节点时，代理和 DataSync 服务之间传输的数据不会通过公共互联网，也不需要公有 IP 地址。有关此配置的 end-to-end 说明，请参阅[将AWS DataSync代理与 VPC 终端节点一起使用](#)。

创建代理以从网络文件系统 (NFS)、服务器消息块 (SMB)、Hadoop 分布式文件系统 (HDFS) 或自管理的对象存储源位置读取

1. 从AWS DataSync控制台下载当前 DataSync .ova映像或启动基于 DataSync Amazon EC2 的当前亚马逊机器映像 (AMI)。有关如何获取 .ova 镜像或 Amazon EC2 AMI 的信息，请参阅[创建代AWS DataSync理](#)。有关硬件要求和推荐的 Amazon EC2 实例类型的信息，请参阅[虚拟机要求](#)。

### Important

如果您在 Amazon EC2 上部署代理，请部署代理，使其不需要在可用区之间进行网络流量（以避免为此类流量收费）。

- 要访问您的 Amazon EFS 或 Amazon FSx for Windows 文件系统，请在具有文件系统挂载目标的可用区中部署代理。
- 对于自管理文件系统，请在文件系统所在的可用区部署代理。

要了解有关所有数据传输价格的更多信息AWS 区域，请参阅 [Amazon EC2 按需定价](#)。

2. 确保满足代理的网络连接要求。有关网络要求的信息，请参阅[AWS DataSync网络要求](#)。
3. 在虚拟机管理程序中部署 .ova 映像，打开虚拟机管理程序的电源，并记下代理的 IP 地址。确保您可以到达代理的端口 80。可以使用以下命令进行检查。



```
nc -vz agent-ip-address 80
```

### Note

.ova默认凭据是登录名**admin**、密码**password**。您可以在虚拟机 (VM) 本地控制台上更改密码。您无需登录虚拟机即可使用基本 DataSync 功能。登录主要用于故障排除、网络特定设置等。

您可以使用虚拟机的虚拟机管理程序客户端登录代理虚拟机本地控制台。有关如何使用虚拟机本地控制台的信息，请参阅[使用AWS DataSync代理的本地控制台](#)。

4. 将 HTTP/1.1 GET 请求发送到代理以获取激活密钥。您可以使用标准 Unix 工具执行此操作：

- 要使用公共服务端点激活代理，请使用以下命令。

```
curl "http://agent-ip-address/?gatewayType=SYNC&activationRegion=aws-region&no_redirect"
```

- 要使用虚拟私有云 (VPC) 端点激活代理，请使用 VPC 终端节点的 IP 地址。使用以下命令。

```
curl "http://agent-ip-address/?gatewayType=SYNC&activationRegion=aws-region&privateLinkEndpoint=IP address of VPC endpoint&endpointType=PRIVATE_LINK&no_redirect"
```

要找到正确的 IP 地址，请打开 Amazon VPC 控制台 <https://console.aws.amazon.com/vpc/>，然后从左侧的导航窗格中选择终端节点。选择 DataSync 终端节点，然后检查子网列表以查找与您为 VPC 终端节点设置选择的子网相对应的私有 IP 地址。

有关 VPC 终端节点配置的更多信息，请参阅[将您的DataSync代理配置为使用 VPC 终端节点](#)中的步骤 5。

- 要使用联邦信息处理标准 (FIPS) 终端节点激活代理，请指定 `endpointType=FIPS`。此外，在美国AWS 区域境内，该`activationRegion`值必须设置为 `a`。要激活 FIPS 终端节点，请使用以下命令。

```
curl "http://agent-IP-address/?gatewayType=SYNC&activationRegion=US-based-aws-region&endpointType=FIPS&no_redirect"
```

此命令返回一个如下所示的激活密钥。

F0EFT-7FPPR-GG7MC-3I9R3-27DOH

5. 在具有激活密钥后，执行下列操作之一：

- 要使用公有终端节点或 FIPS 终端节点激活代理，请使用以下命令。

```
aws datasync create-agent \
  --agent-name agent-name-you-specify \
  --activation-key obtained-activation-key
```

- 要使用 VPC 终端节点激活代理，请使用以下命令。

```
aws datasync create-agent \
  --agent-name agent-name-you-specify \
  --vpc-endpoint-id vpc-endpoint-id \
  --subnet-arns subnet-arns \
  --security-group-arns security-group-arns \
  --activation-key obtained-activation-key
```

在此命令中，使用以下参数：

- *vpc-endpoint-id*— 代理连接到的AWS端点。要查找终端节点 ID，请打开 Amazon VPC 控制台 <https://console.aws.amazon.com/vpc/>，然后从左侧的导航窗格中选择终端节点。复制终端节点的 DataSync 端点 ID 值。有关 VPC 终端节点配置的更多信息，请参阅[将您的 DataSync 代理配置为使用 VPC 终端节点](#)中的步骤 5。

- *security-group-arn*— 用于任务终端节点的安全组的 Amazon 资源名称 (ARN)。

这是您在[将您的 DataSync 代理配置为使用 VPC 终端节点](#)的步骤 3 中创建的安全组。

- *subnet-arns*— 创建代理任务端点的子网的 ARN。

这是您在[将您的 DataSync 代理配置为使用 VPC 终端节点](#)的步骤 1 中选择的子网。

这些命令返回刚激活的代理的 ARN。该 ARN 如下所示。

```
{
  "AgentArn": "arn:aws:datasync:us-east-1:111222333444:agent/
agent-0b0addbeef44baca3"
}
```

**Note**

选择一个服务终端节点后，以后不能进行更改。

激活代理后，它将关闭端口 80，无法再访问该端口。如果激活代理后无法连接到代理，请使用以下命令验证激活是否成功：

```
aws datasync list-agents
```

**Note**

确保在整个过程中使用相同的AWS证书。不要在使用不同AWS证书进行身份验证的多个终端之间切换。

## 使用创建AWS DataSync地点AWS CLI

每AWS DataSync 项任务都由传输中的两个地点组成。源位置 定义要从中读取数据的存储系统或服务。目标位置 定义要向其中写入数据的存储系统或服务。

使用AWS Command Line Interface (AWS CLI) 可以为以下存储系统和服务创建位置：

- 网络文件系统 (NFS)
- 服务器消息块 (SMB)
- Hadoop Distributed File System ( HDFS )
- 自我管理的对象存储源位置
- Amazon Elastic File System (Amazon EFS)
- Amazon FSx for Windows File Server
- Amazon FSx for Lustre
- Amazon FSx for OpenZFS
- Amazon FSx for NetApp ONTAP
- Amazon Simple Storage Service (Amazon S3)

有关更多信息，请参阅[我可以在哪里传输我的数据AWS DataSync？](#)：

## 创建 NFS 位置

NFS 位置定义 NFS 服务器上可以读取或写入的文件系统。您也可以使用`创建 NFS 位置AWS Management Console。有关更多信息，请参阅[配置来自 NFS 文件服务器的AWS DataSync传输](#)：

### Note

如果您在设备上使用 NFS 位置，请参阅[使用以下AWS DataSync方式配置传输AWS Snowcone](#)，了解有关向该AWS Snowcone设备传输数据或从该设备传输数据的更多信息。

### 使用 CLI 创建 NFS 位置

- 使用以下命令创建 NFS 源位置。

```
$ aws datasync create-location-nfs \  
  --server-hostname nfs-server-address \  
  --on-prem-config AgentArns=datasync-agent-arns \  
  --subdirectory nfs-export-path
```

对于上述命令，以下情况适用：

- 您为--subdirectory参数提供的路径 (*nfs-export-path*) 必须是 NFS 服务器导出的路径，或者是子目录。网络中的其他 NFS 客户端必须能够安装此路径。要查看 NFS 服务器导出的所有路径，请从可访问服务器的 NFS 客户端运行命令 `showmount -e nfs-server-address`。您可以指定显示在结果中的任何目录以及该目录的任何子目录。
- 要传输指定文件夹中的所有数据，DataSync 需要有权读取所有数据。要 DataSync 授予权限，您可以执行以下两项操作之一。您可以使用 `no_root_squash` 配置 NFS 导出。或者，对于您希望 DataSync 访问的所有文件，您可以确保权限允许所有用户读取。通过执行上述任一操作，将允许代理读取这些文件。要使代理能够访问目录，您还必须额外为所有用户授予执行访问权限。
- 确保无需 Kerberos 身份验证即可访问 NFS 导出路径。

DataSync 自动选择用于从 NFS 位置读取的 NFS 版本。要指定 NFS 版本，请在[NfsMountOptions](#) API 操作中使用可选 `Version` 参数。

此命令返回 NFS 位置的 Amazon 资源名称 (ARN)，与以下所示的 ARN 类似。

```
{ "LocationArn": "arn:aws:datsync:us-east-1:111222333444:location/loc-0f01451b140b2af49" }
```

要确保可以挂载目录，您可以连接到与代理具有相同网络配置的任何计算机，并运行以下命令。

```
mount -t nfs -o nfsvers=<nfs-server-version <nfs-server-address:<nfs-export-path <test-folder
```

以下是该命令的示例。

```
mount -t nfs -o nfsvers=3 198.51.100.123:/path_for_sync_to_read_from /temp_folder_to_test_mount_on_local_machine
```

## 创建 SMB 位置

SMB 位置定义 SMB 服务器上可以读取或写入的文件系统。您还可以使用控制台创建 SMB 位置。有关更多信息，请参阅[配置来自 SMB 文件服务器的AWS DataSync传输](#)：

使用 CLI 创建 SMB 位置

- 使用以下命令创建 SMB 源位置。

```
aws datsync create-location-smb \
  --server-hostname smb-server-address \
  --user user-who-can-mount-share \
  --domain windows-domain-of-smb-server \
  --password user-password \
  --agent-arns datsync-agent-arns \
  --subdirectory smb-export-path
```

您为--subdirectory参数提供的路径应该是 SMB 服务器导出的路径。*smb-export-path*使用正斜杠指定路径；例如/path/to/folder。网络中的其他 SMB 客户端应该能够访问此路径。

DataSync 自动选择用于从 SMB 位置读取的 SMB 版本。要指定 SMB 版本，请在[SmbMountOptions](#) API 操作中使用可选 Version 参数。

此命令返回 SMB 位置的 Amazon 资源名称 (ARN)，与以下所示的 ARN 类似。

```
{
```

```
"LocationArn": "arn:aws:datsync:us-east-1:111222333444:location/loc-0f01451b140b2af49"
}
```

## 创建 HDFS 位置

HDFS 位置定义了 Hadoop 集群上可读取或可写入的文件系统。您也可以使用创建 HDFS 位置AWS Management Console。有关更多信息，请参阅[配置来自 HDFS 的AWS DataSync传输](#)：

### 使用创建 HDFS 位置AWS CLI

- 使用以下命令创建 HDFS 位置。在以下示例中，将每个 *user input placeholder* 替换为您自己的信息。

```
aws datsync create-location-hdfs --name-nodes [{"Hostname":"host1", "Port": 8020}] \
  \
  --authentication-type "SIMPLE|KERBEROS" \
  --agent-arns [arn:aws:datsync:us-east-1:123456789012:agent/agent-01234567890example] \
  --subdirectory "/path/to/my/data"
```

create-location-hdfs命令中必需参数如下：

- name-nodes— 指定 Hadoop 集群 NameNode 中的主机名或 IP 地址以及 NameNode 正在监听的 TCP 端口。
- authentication-type— 连接到 Hadoop 集群时要使用的身份验证类型。指定 SIMPLE 或 KERBEROS。

如果您使用SIMPLE身份验证，请使用--simple-user参数指定用户的用户名。如果您使用KERBEROS身份验证，请使用--kerberos-principal--kerberos-keytab、和--kerberos-krb5-conf参数。有关更多信息，请参阅[create-location-hdfs](#)。

- agent-arns— 用于 HDFS 位置的 DataSync 代理的 ARN。

上述命令返回位置 ARN，类似于以下内容：

```
{
  "arn:aws:datsync:us-east-1:123456789012:location/loc-01234567890example"
}
```

## 创建对象存储位置

位置是代表本地或其他云提供商（例如 Google Cloud Storage 存储分区）托管的对象存储系统的 DataSync 终端节点。

有关对象存储位置的更多信息，包括兼容性要求，请参阅[配置来自对象存储系统的AWS DataSync传输](#)。

使用创建对象存储位置AWS CLI

1. 复制以下create-location-object-storage命令：

```
aws datasync create-location-object-storage \  
  --server-hostname object-storage-server.example.com \  
  --bucket-name your-bucket \  
  --agent-arns arn:aws:datasync:us-east-1:123456789012:agent/  
agent-01234567890deadfb
```

2. 在命令中指定以下必需参数：

- --server-hostname— 指定对象存储服务器的域名或 IP 地址。
- --bucket-name— 指定您要传输到或传输的对象存储服务器上存储分区的名称。
- --agent-arns— 指定要连接到对象存储服务器的 DataSync 代理。

3. （可选）将以下任何参数添加到命令中：

- --server-port— 指定对象存储服务器接受入站网络流量的端口（例如，端口443）。
- --server-protocol— 指定对象存储服务器通信时使用的协议（HTTP或HTTPS）。
- --access-key— 如果需要凭证来向对象存储服务器进行身份验证，则指定访问密钥（例如，用户名）。
- --secret-key— 如果需要凭证来向对象存储服务器进行身份验证，则指定私有密钥（例如，密码）。
- --server-certificate— 指定 DataSync 代理在连接到您的对象存储服务器时将信任的私有或自签名证书。

自定义证书文件必须具有.pem扩展名（例如file:///home/user/mycert.pem）。

- --subdirectory— 指定对象存储服务器的对象前缀。

DataSync 仅复制具有此前缀的对象。

- `--tags`— 指定表示要添加到位置资源的标签的键值对。

标签可帮助您管理、筛选和搜索资源。我们建议为您的位置创建一个名称标签。

#### 4. 运行 `create-location-object-storage` 命令。

您会收到一条显示您刚刚创建的位置 ARN 的响应。

```
{
  "LocationArn": "arn:aws:datsync:us-east-1:123456789012:location/
loc-01234567890abcdef"
}
```

## 创建 Amazon EFS 位置

位置是可以读取或可写入的 Amazon EFS 文件系统的终端节点。您也可以使用控制台创建此类位置。有关更多信息，请参阅[使用亚马逊 EFS 配置AWS DataSync传输](#)：

使用创建 Amazon EFS 位置AWS CLI

1. 如果您没有 Amazon EFS 文件系统，请创建一个。有关如何创建 EFS 文件系统的信息，请参阅《Amazon Elastic File System 用户指南》中的[Amazon Elastic File System 入门](#)。
2. 标识至少有该文件系统的挂载目标的子网。您可以使用`describe-mount-targets`命令查看所有挂载目标和与 EFS 文件系统关联的子网。

```
aws efs describe-mount-targets \
  --region aws-region \
  --file-system-id file-system-id
```

### Note

您指定的是目标 S3 存储桶或 EFS 文件系统所在的 AWS 区域

此命令返回与以下信息类似的目标相关信息。

```
{
  "MountTargets": [
    {
```



```

        "OwnerId": "111222333444",
        "MountTargetId": "fsmt-22334a10",
        "FileSystemId": "fs-123456ab",
        "SubnetId": "subnet-f12a0e34",
        "LifecycleState": "available",
        "IpAddress": "11.222.0.123",
        "NetworkInterfaceId": "eni-1234a044"
    }
]
}

```

3. 指定可以访问挂载目标的 Amazon EC2 安全组。您可以运行以下命令来找到挂载目标的安全组。

```

aws efs describe-mount-target-security-groups \
  --region aws-region \
  --mount-target-id mount-target-id

```

您提供的安全组必须能够与指定子网中的挂载目标的安全组进行通信。

挂载目标的安全组 M 与您提供在此阶段使用的安全组 S 之间的关系如下所示：DataSync

- 安全组 M (与挂载目标关联) 必须允许从安全组 S 对 NFS 端口 (2049) 上的 TCP 协议进行入站访问。

您可以通过其 IP 地址 (CIDR 范围) 或其安全组启用入站连接。

- 您提供的访问 Amazon EFS 的安全组 S 应具有一个规则，该规则允许与 NFS 端口进行出站连接。DataSync 它在文件系统的其中一个挂载目标上启用出站连接。

您可以按 IP 地址 (CIDR 范围) 或安全组启用出站连接。

有关安全组和挂载目标的信息，请参阅 [Amazon Elastic File System 用户指南中的 Amazon EC2 实例和挂载目标的安全组](#)。

4. 创建位置。要创建 EFS 位置，您需要您的 Amazon EC2 子网、Amazon EC2 安全组和文件系统的 ARN。由于 DataSync API 接受完全限定的 ARN，您可以构建这些 ARN。有关如何为不同的服务构建 ARN 的信息，请参阅中的 [亚马逊资源名称 \(ARN\) AWS 一般参考](#)。

使用以下命令创建 EFS 位置。

```

aws datasync create-location-efs \
  --subdirectory /path/to/your/subdirectory \

```

```
--efs-filesystem-arn 'arn:aws:elasticfilesystem:region:account-id:file-
system/filesystem-id' \
--ec2-config SecurityGroupArns='arn:aws:ec2:region:account-id:security-
group/security-group-id',SubnetArn='arn:aws:ec2:region:account-id:subnet/subnet-id'
```

### Note

您指定的是目标 S3 存储桶或 EFS 文件系统所在的 AWS 区域。

这些命令返回一个如下所示的位置 ARN。

```
{
  "LocationArn": "arn:aws:datsync:us-west-2:111222333444:location/
loc-07db7abfc326c50fb"
}
```

## 创建 Amazon FSx for Windows File Server 位置

位置是 FSx for Windows File Server 的端点，您可以对其进行读取或写入。

您还可以使用控制台创建 FSx for Windows File Server 位置。有关更多信息，请参阅[使用 Amazon FSx for Windows File Server AWS DataSync](#)：

使用创建 FSx for Windows File Server 位置 AWS CLI

- 使用以下命令创建 Amazon FSx 位置。

```
aws datsync create-location-fsx-windows \
--fsx-filesystem-arn arn:aws:fsx:region:account-id:file-system/filesystem-id \
--security-group-arns arn:aws:ec2:region:account-id:security-group/group-id \
--user smb-user --password password
```

在 create-location-fsx-windows 命令中，指定以下项：

- fsx-filesystem-arn— 要读取或写入的文件系统的完全限定的 Amazon Resource Name (ARN)。

DataSync API 接受完全合格的 ARN，您可以构造这些 ARN。有关如何为不同的服务构建 ARN 的信息，请参阅中的[亚马逊资源名称 \(ARN\) AWS 一般参考](#)。

- `security-group-arns`— 可应用于文件系统首选子网网络接口的 Amazon EC2 安全组的 ARN。
- `AWS 区域`— 您指定的区域是您的目标 Amazon FSx 文件系统所在的区域。

上一命令返回一个如下所示的位置 ARN。

```
{
  "LocationArn": "arn:aws:datsync:us-west-2:111222333444:location/
loc-07db7abfc326c50fb"
}
```

## 创建 Amazon FSx for Lustre 位置

位置是您可以读取或写入 FSx for Lustre 文件系统的端点。

您还可以使用控制台创建 FSx for Lustre 位置。有关更多信息，请参阅[使用 Amazon FSx for Lustre 配置 AWS DataSync 传](#)：

要为 Lustre 位置创建 FSx，请使用 AWS CLI

- 使用以下命令为 Lustre 位置创建 FSx。

```
aws datsync create-location-fsx-lustre \
  --fsx-filesystem-arn arn:aws:fsx:region:account-id:file-system:filesystem-id \
  --security-group-arns arn:aws:ec2:region:account-id:security-group/group-id
```

`create-location-fsx-lustre` 命令中必需参数。

- `fsx-filesystem-arn`— 要读取或写入的文件系统的完全限定的 Amazon Resource Name (ARN)。
- `security-group-arns`— 适用于文件系统首选子网网络接口的 Amazon EC2 安全组的 ARN。

上述命令会返回类似以下内容的位置 ARN。

```
{
  "LocationArn": "arn:aws:datsync:us-west-2:111222333444:location/
loc-07sb7abfc326c50fb"
}
```

```
}
```

## 创建 Amazon FSx for OpenZFS 位置

位置是 DataSync 可以访问以进行传输的 FSx for OpenZFS 文件系统的端点。你也可以在[控制台中为 OpenZFS 位置创建 FSx](#)。

使用创建 FSx for OpenZFS 位置AWS CLI

### 1. 复制以下命令：

```
$ aws datasync create-location-fsx-opensfs \  
  --fsx-filesystem-arn arn:aws:fsx:region:account-id:file-system/filesystem-id \  
  --security-group-arns arn:aws:ec2:region:account-id:security-group/group-id \  
  --protocol NFS={}
```

### 2. 在命令中指定以下必需选项：

- 对于 `fsx-filesystem-arn`，请指定位置文件系统的完全合格的 Amazon Resource Name (ARN)。这包括您的文件系统所在的位置AWS 账户、您的和文件系统 ID。AWS 区域
- 对于 `security-group-arns`，指定 Amazon EC2 安全组的 ARN，该安全组提供对 OpenZFS 文件系统首选子网的 FSx [网络接口](#) 的访问权限。这包括您的 Amazon EC2 实例所在的位置 AWS 账户、您的和安全组 ID。AWS 区域

有关安全组的更多信息，请参阅《Amazon FSx for [OpenZFS 用户指南](#)》[VPC Amazon FSx for OpenZFS 用户指南](#)中的 Amazon FS x

- 为 `protocol`，指定 DataSync 用于访问文件系统的协议。（DataSync 目前仅支持 NFS。）

### 3. 运行命令。您会收到一条回复，其中显示您刚刚创建的位置。

```
{  
  "LocationArn": "arn:aws:datasync:us-west-2:123456789012:location/loc-  
  abcdef01234567890"  
}
```

## 创建 Amazon FSx for NetApp or ONTAP 位置

位置是 DataSync 可以访问以进行传输的 FSx for ONTAP 文件系统的终端节点。您还可以在[控制台中为 ONTAP 位置创建 FSx](#)。

## 使用创建 FSx for ONTAP 位置AWS CLI

### 1. 复制以下命令：

```
$ aws datasync create-location-fsx-ontap \  
  --storage-virtual-machine-arn arn:aws:fsx:region:account-id:storage-virtual-  
  machine/fs-file-system-id \  
  --security-group-arns arn:aws:ec2:region:account-id:security-group/group-id \  
  --protocol data-transfer-protocol={}
```

### 2. 在命令中指定以下必需选项：

- 对于 `storage-virtual-machine-arn`，请指定要将数据复制到或从中复制数据的文件系统  
中的存储虚拟机 ( SVM ) 的完全限定的 Amazon Resource Name (ARN)。

此 ARN 包括您的文件系统所在的位置AWS 账户、您的文件系统和 SVM ID。AWS 区域

- 对于 `security-group-arns`，请指定 Amazon EC2 安全组的 ARN，这些安全组提供对文件  
系统首选子网[网络接口](#)的访问权限。

这包括您的 Amazon EC2 实例所在的位置AWS 账户、您的和您的安全组 ID。AWS 区域您最多  
可以指定五个安全组 ARN。

有关安全组的更多信息，请参阅《[Amazon FSx for NetApp ONTAP 用户指南](#)》中的 [Amazon VPC 文件系统访问控制](#)。

- 对于 `protocol`，请配置用于访问文件系统的 DataSync SVM 访问文件系统的 SVM 的协议来访问  
文件系统的 SVM。
  - 对于 NFS，您可以使用默认配置：

```
--protocol NFS={}
```

- 对于 SMB，必须指定可以访问 SVM 的用户名和密码：

```
--protocol SMB={User=smb-user, Password=smb-password}
```

### 3. 运行 命令。

您会收到一条显示刚刚创建位置的回复。

```
{  
  "LocationArn": "arn:aws:datasync:us-west-2:123456789012:location/loc-  
  abcdef01234567890"
```

```
}
```

## 创建Amazon S3 位置

Amazon S3 位置需要一个可以读取或写入的 S3 存储桶。要创建 S3 存储桶，请参阅 Amazon S3 用户指南中的[创建存储桶](#)。

DataSync 要访问 S3 存储桶，DataSync 需要具有所需权限的AWS Identity and Access Management (IAM) 角色。按照以下程序，您可以使用创建 IAM 角色、所需的 IAM 策略和 S3 位置AWS CLI。

DataSync 要担任 IAM 角色，必须在您的账户和区域中激活AWS Security Token Service (AWS STS)。有关临时安全凭证的更多信息，请参阅《IAM 用户指南》中的[IAM 临时安全凭证](#)。

您还可以使用控制台创建 S3 位置。有关更多信息，请参阅[使用 Amazon S3 配置AWS DataSync传输](#)：

### 使用 CLI 创建 S3 位置

1. 创建 IAM 信任策略，DataSync 允许代入访问 S3 存储桶所需的 IAM 角色。

以下是信任策略的示例。

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Principal": {
        "Service": "datasync.amazonaws.com"
      },
      "Action": "sts:AssumeRole"
    }
  ]
}
```

2. 为该 IAM 策略创建一个临时文件，如以下示例所示。

```
$ ROLE_FILE=$(mktemp -t sync.iam.role.filename.json)
$ IAM_ROLE_NAME='YourBucketAccessRole'

$ cat<<EOF> ${ROLE_FILE}
```

```
{
  "Version": "2012-10-17",
  "Statement": [{
    "Effect": "Allow",
    "Principal": {
      "Service": "datasync.amazonaws.com"
    },
    "Action": "sts:AssumeRole"
  }]
}
```

### 3. 创建 IAM 角色并向其附加此 IAM policy。

以下命令创建一个 IAM 角色并向其附加该策略。

```
$ aws iam create-role --role-name ${IAM_ROLE_NAME} --assume-role-policy-document
file://${ROLE_FILE}
{
  "Role": {
    "Path": "/",
    "RoleName": "YourBucketAccessRole",
    "RoleId": "role-id",
    "Arn": "arn:aws:iam::account-id:role/YourBucketAccessRole",
    "CreateDate": "2018-07-27T02:49:23.117Z",
    "AssumeRolePolicyDocument": {
      "Version": "2012-10-17",
      "Statement": [{
        "Effect": "Allow",
        "Principal": {
          "Service": "datasync.amazonaws.com"
        },
        "Action": "sts:AssumeRole"
      }]
    }
  }
}
```

### 4. 允许您创建的 IAM 角色写入 S3 存储桶。

向 IAM 角色附加一个具有访问您的 S3 存储桶的足够权限的 IAM 策略。以下示例显示读取和写入 S3 存储桶所需的最低权限AWS 区域。 DataSync

```
{
```

```

"Version": "2012-10-17",
"Statement": [
  {
    "Action": [
      "s3:GetBucketLocation",
      "s3:ListBucket",
      "s3:ListBucketMultipartUploads"
    ],
    "Effect": "Allow",
    "Resource": "YourS3BucketArn"
  },
  {
    "Action": [
      "s3:AbortMultipartUpload",
      "s3:DeleteObject",
      "s3:GetObject",
      "s3:ListMultipartUploadParts",
      "s3:PutObjectTagging",
      "s3:GetObjectTagging",
      "s3:PutObject"
    ],
    "Effect": "Allow",
    "Resource": "YourS3BucketArn/*"
  }
]
}

```

要将该策略附加到您的 IAM 角色，请运行以下命令。

```

$ aws iam attach-role-policy \
  --role-name role-name \
  --policy-arn 'arn:aws:iam::aws:policy/YourPolicyName'

```

对于启用的 Amazon S3 存储桶 AWS Outposts，请使用以下策略。

```

{
  "Version": "2012-10-17",
  "Statement": [{
    "Action": [
      "s3-outposts:ListBucket",
      "s3-outposts:ListBucketMultipartUploads"
    ],

```



```

    "Effect": "Allow",
    "Resource": [
        "s3outpostsBucketArn",
        "s3outpostsAccessPointArn"
    ],
    "Condition": {
        "StringLike": {
            "s3-outposts:DataAccessPointArn": "s3outpostsAccessPointArn"
        }
    }
},
{
    "Action": [
        "s3-outposts:AbortMultipartUpload",
        "s3-outposts>DeleteObject",
        "s3-outposts:GetObject",
        "s3-outposts:ListMultipartUploadParts",
        "s3-outposts:PutObjectTagging",
        "s3-outposts:GetObjectTagging",
        "s3-outposts:PutObject"
    ],
    "Effect": "Allow",
    "Resource": [
        "s3outpostsBucketArn/*",
        "s3outpostsAccessPointArn"
    ],
    "Condition": {
        "StringLike": {
            "s3-outposts:DataAccessPointArn": "s3outpostsAccessPointArn"
        }
    }
},
{
    "Effect": "Allow",
    "Action": [
        "s3-outposts:GetAccessPoint"
    ],
    "Resource": "s3outpostsAccessPointArn"
}
]
}

```

## 5. 创建 S3 位置。

使用以下命令创建您的 Amazon S3 位置。

```
$ aws datasync create-location-s3 \  
  --s3-bucket-arn 'arn:aws:s3:::DOC-EXAMPLE-BUCKET' \  
  --s3-storage-class 'your-S3-storage-class' \  
  --s3-config 'BucketAccessRoleArn=arn:aws:iam::account-id:role/role-allowing-DS-  
operations' \  
  --subdirectory /your-folder
```

这些命令返回一个如下所示的位置 ARN。

```
{  
  "LocationArn": "arn:aws:datasync:us-east-1:111222333444:location/  
loc-0b3017fc4ba4a2d8d"  
}
```

位置类型信息在中编码LocationUri。在此示例中，中的s3://前缀LocationUri显示了位置  
的类型。

如果您的 Amazon S3 存储桶位于AWS前哨站上，则必须在前哨上部署 Amazon EC2 代理。代  
理必须位于允许访问命令中指定的接入点的虚拟私有云 (VPC) 中。还必须在 Outpost 的父区域  
激活代理，并且能够在存储桶的AWS Outposts终端节点上路由到 Amazon S3。有关在上启动  
DataSync 代理的更多信息AWS Outposts，请参阅[在上部署您的代理 AWS Outposts](#)。

使用以下命令在前哨创建 Amazon S3 位置。

```
aws datasync create-location-s3 \  
  --s3-bucket-arn access-point-arn \  
  --s3-config BucketAccessRoleArn=arn:aws:iam::account-id:role/role-allowing-DS-  
operations
```

#### Note

- 对对象数据或元数据的更改等同于删除一个对象并创建一个新对象来替换它。在以下情况  
下，这会导致额外费用：
  - 使用对象版本控制时-对对象数据或元数据的更改会创建对象的新版本。

- 当使用可能因覆盖、删除或检索对象而产生额外费用的存储类时，对对象数据或元数据的更改会产生此类费用。有关更多信息，请参阅[Amazon S3 传输的存储类注意事项](#)：
- 当您使用对象版本控制时，单个 DataSync 任务执行可能会创建 Amazon S3 对象的多个版本。
- 除了授予 DataSync 权限的 IAM 策略外，我们还建议为您的 S3 存储桶创建分段上传存储桶策略。这样做可帮助您控制存储成本。有关更多信息，请参阅博客帖子 [S3 生命周期管理更新-支持分段上传和删除标记](#)的支持。

## 使用创建AWS DataSync任务AWS CLI

创建源位置和目标位置后，即可创建AWS DataSync任务。

### Important

如果您计划将数据传输到或传出 Amazon S3 地点，请在开始之前查看[会 DataSync 如何影响您的 S3 请求费用](#)和[DataSync 定价页面](#)。

### 使用创建任务AWS CLI

1. 使用以下命令创建 Amazon Log CloudWatch s 日志组。

```
aws logs create-log-group \  
  --log-group-name your-log-group
```

2. 将 IAM 资源策略附加到您的日志组。有关如何附加策略的说明，请参阅 [允许DataSync将日志上传到CloudWatch日志组](#)。
3. 使用以下命令创建任务。

```
aws datasync create-task \  
  --source-location-arn 'arn:aws:datasync:region:account-id:location/location-id' \  
  \  
  --destination-location-arn 'arn:aws:datasync:region:account-id:location/location-id' \  
  --cloud-watch-log-group-arn 'arn:aws:logs:region:account-id:log-group:log-group' \  
  --name task-name
```

此命令返回任务的 Amazon 资源名称 (ARN)，如下所示。

```
{
  "TaskArn": "arn:aws:datsync:us-east-1:111222333444:task/
task-08de6e6697796f026"
}
```

创建AWS 服务在不同区域之间传输数据的任务时，必须在不同的区域中指定另一个位置（例如，要在us-east-1和之间传输数据us-east-2），请在其中一个区域 DataSync 中使用并使用以下命令创建任务。

```
aws datsync create-task \
  --source-location-arn 'arn:aws:datsync:us-east-1:account-id:location/location-id \
  --destination-location-arn 'arn:aws:datsync:us-east-2:account-id:location/location-id \
  --cloud-watch-log-group-arn 'arn:aws:logs:region:account-id' \
  --name task-name \
  --options
VerifyMode=NONE,OverwriteMode=NEVER,Atime=BEST_EFFORT,Mtime=PRESERVE,Uid=INT_VALUE,Gid=INT
```

将使用默认配置选项创建任务。如果您要在任务创建期间配置不同的选项，请将 `--options` 参数添加到 `create-task` 命令。以下示例介绍如何指定不同的选项。有关这些选项的描述，请参阅 [the section called “Options”](#)。

```
aws datsync create-task \
  --source-location-arn 'arn:aws:datsync:region:account-id:location/location-id' \
  --destination-location-arn 'arn:aws:datsync:region:account-id:location/location-id' \
  --cloud-watch-log-group-arn 'arn:aws:logs:region:account-id:log-group:log-group' \
  --name task-name \
  --options
VerifyMode=NONE,OverwriteMode=NEVER,Atime=BEST_EFFORT,Mtime=PRESERVE,Uid=INT_VALUE,Gid=INT
```

创建任务时，您可以将任务配置为包括或排除特定的文件、文件夹和对象。有关更多信息，请参阅 [筛选传输的数据AWS DataSync](#)：您还可以安排希望任务的运行时间。有关更多信息，请参阅 [安排AWS DataSync任务](#)：

## 使用... 启动AWS DataSync任务AWS CLI

您可以使用开始传输数据AWS CLI。AWS DataSync

要启动 DataSync 任务，您只需要指定要运行的任务的 Amazon Resource Name (ARN)。以下为示例 `start-task-execution` 命令：

```
aws datasync start-task-execution \  
  --task-arn 'arn:aws:datasync:region:account-id:task/task-id'
```

您可以修改特定任务执行的任务设置，如以下示例所示。有关这些选项的描述，请参阅 [the section called “Options”](#)。您还可以配置特定的任务运行以专注于要传输的特定文件、文件夹和对象上使用的特定文件、文件夹和对象进行传输。有关更多信息，请参阅 [筛选传输的数据AWS DataSync](#)：

以下示例使用一些不同于任务默认设置的设置启动任务：

```
aws datasync start-task-execution \  
  --override-options VerifyMode=NONE,OverwriteMode=NEVER,PosixPermissions=NONE
```

命令会返回类似于以下示例的 ARN，用于执行任务：

```
{  
  "TaskExecutionArn": "arn:aws:datasync:us-east-1:209870788375:task/  
task-08de6e6697796f026/execution/exec-04ce9d516d69bd52f"  
}
```

### Note

每个代理每次可以运行单个任务。

## 筛选AWS DataSync资源

您可以使用中的和 `ListTasks` API 操作筛选您的AWS DataSync 位置 `ListLocations`和任务AWS CLI。例如，检索您最近的任务列表。

## 用于筛选的参数

您可以使用 API 过滤器来缩小 `ListTasks` 和返回的资源列表的范围 `ListLocations`。例如，要检索您的所有 Amazon S3 位置，您可以将筛选器名称 `LocationTypeS3` 和 `ListLocations` 一起使用 `OperatorEquals`。

要筛选 API 结果，必须指定筛选器名称、运算符和值。

- **Name**— 正在使用的过滤器的名称。每个 API 调用都支持可用于它的过滤器列表（`LocationType` 例如，`forListLocations`）。
- **Values**— 要筛选的值。例如，您可能希望仅显示 Amazon S3 位置。
- **Operator**— 用于比较筛选器值的运算符（例如，`Equals` 或 `Contains`）。

下表列出了可用的运算符。

操作符	密钥类型
<code>Equals</code>	字符串，数字
<code>NotEquals</code>	字符串，数字
<code>LessThan</code>	数字
<code>LessThanOrEqual</code>	数字
<code>GreaterThan</code>	数字
<code>GreaterThanOrEqual</code>	数字
<code>In</code>	字符串
<code>Contains</code>	字符串
<code>NotContains</code>	字符串
<code>BeginsWith</code>	字符串

## 按位置筛选

ListLocations支持以下过滤器名称：

- LocationType— 按位置类型筛选：
  - SMB
  - NFS
  - HDFS
  - OBJECT\_STORAGE
  - S3
  - OUTPOST\_S3
  - FSX\_WINDOWS
  - FSX\_LUSTRE
  - FSX\_OPENZFS\_NFS
  - FSX\_ONTAP\_NFS
  - FSX\_ONTAP\_SMB
- LocationUri— 根据DescribeLocation\* API 调用返回的分配给该位置的统一资源标识符 (URI) 进行筛选（例如，对于 Amazon S3 位置）。s3://*bucket-name/your-prefix*
- CreationTime— 根据创建位置的时间进行筛选。输入格式采用yyyy-MM-dd:mm:ss协调世界时 (UTC)。

以下AWS CLI示例列出了所有类型为 Amazon S3 的位置，这些位置 URI 以字符串开头"s3://DOC-EXAMPLE-BUCKET"，创建于 2019-12-15 17:15:20 UTC 或之后。

```
aws datasync list-locations \  
  --filters [{Name=LocationType, Values=["S3"], Operator=Equals},  
  {Name=LocationUri, Values=["s3://DOC-EXAMPLE-BUCKET"], Operator=BeginsWith},  
  {Name=CreationTime, Values=["2019-12-15 17:15:20"], Operator=GreaterThanOrEqual}]
```

该命令会返回类似以下内容的输出。

```
{  
  "Locations": [  
    {  
      "LocationArn": "arn:aws:datasync:us-east-1:111122223333:location/  
loc-3333333333abcdef0",
```

```
        "LocationUri": "s3://DOC-EXAMPLE-BUCKET-examples/"
    },
    {
        "LocationArn": "arn:aws:datasync:us-east-1:123456789012:location/
loc-987654321abcdef0",
        "LocationUri": "s3://DOC-EXAMPLE-BUCKET-examples-2/"
    }
]
}
```

## 按任务筛选

ListTasks支持以下过滤器名称。

- LocationId— 根据亚马逊资源名称 (ARN) 值对源位置和目标位置进行筛选。
- CreationTime— 根据任务创建时间进行筛选。输入格式采用yyyy-MM-dd:mm:ss UTC 表示。

以下AWS CLI示例显示了筛选时的语法LocationId。

```
aws datasync list-tasks \
  --filters Name=LocationId,Values=arn:aws:datasync:us-east-1:your-account-id:location/your-location-id,Operator=Contains
```

此命令的输出与以下内容类似。

```
{
  "Tasks": [
    {
      "TaskArn": "arn:aws:datasync:us-east-1:your-account-id:task/your-task-id",
      "Status": "AVAILABLE",
      "Name": "DOC-EXAMPLE-BUCKET"
    }
  ]
}
```



# AWS DataSync API

除了AWS Management Console和之外AWS CLI，您还可以使用 AWS DataSync API 对 [AWSSDK](#) 进行配置和管理DataSync。

## 主题

- [操作](#)
- [数据类型](#)
- [常见错误](#)
- [常见参数](#)

## 操作

支持以下操作：

- [AddStorageSystem](#)
- [CancelTaskExecution](#)
- [CreateAgent](#)
- [CreateLocationEfs](#)
- [CreateLocationFsxLustre](#)
- [CreateLocationFsxOntap](#)
- [CreateLocationFsxOpenZfs](#)
- [CreateLocationFsxWindows](#)
- [CreateLocationHdfs](#)
- [CreateLocationNfs](#)
- [CreateLocationObjectStorage](#)
- [CreateLocationS3](#)
- [CreateLocationSmb](#)
- [CreateTask](#)
- [DeleteAgent](#)
- [DeleteLocation](#)
- [DeleteTask](#)

- [DescribeAgent](#)
- [DescribeDiscoveryJob](#)
- [DescribeLocationEfs](#)
- [DescribeLocationFsxLustre](#)
- [DescribeLocationFsxOntap](#)
- [DescribeLocationFsxOpenZfs](#)
- [DescribeLocationFsxWindows](#)
- [DescribeLocationHdfs](#)
- [DescribeLocationNfs](#)
- [DescribeLocationObjectStorage](#)
- [DescribeLocationS3](#)
- [DescribeLocationSmb](#)
- [DescribeStorageSystem](#)
- [DescribeStorageSystemResourceMetrics](#)
- [DescribeStorageSystemResources](#)
- [DescribeTask](#)
- [DescribeTaskExecution](#)
- [GenerateRecommendations](#)
- [ListAgents](#)
- [ListDiscoveryJobs](#)
- [ListLocations](#)
- [ListStorageSystems](#)
- [ListTagsForResource](#)
- [ListTaskExecutions](#)
- [ListTasks](#)
- [RemoveStorageSystem](#)
- [StartDiscoveryJob](#)
- [StartTaskExecution](#)
- [StopDiscoveryJob](#)
- [TagResource](#)

- [UntagResource](#)
- [UpdateAgent](#)
- [UpdateDiscoveryJob](#)
- [UpdateLocationHdfs](#)
- [UpdateLocationNfs](#)
- [UpdateLocationObjectStorage](#)
- [UpdateLocationSmb](#)
- [UpdateStorageSystem](#)
- [UpdateTask](#)
- [UpdateTaskExecution](#)

## AddStorageSystem

为您希望 DataSync Discovery 收集有关信息的本地存储系统创建AWS资源。

### 请求语法

```
{
  "AgentArns": [ "string" ],
  "ClientToken": "string",
  "CloudWatchLogGroupArn": "string",
  "Credentials": {
    "Password": "string",
    "Username": "string"
  },
  "Name": "string",
  "ServerConfiguration": {
    "ServerHostname": "string",
    "ServerPort": number
  },
  "SystemType": "string",
  "Tags": [
    {
      "Key": "string",
      "Value": "string"
    }
  ]
}
```

### 请求参数

有关所有操作的通用参数的信息，请参阅[通用参数](#)。

请求接受采用 JSON 格式的以下数据。

#### [AgentArns](#)

指定连接本地存储系统的管理界面并从中读取数据的数据Sync代理的 Amazon 资源名称 (ARN)。

类型：字符串数组

数组成员：固定数量为 1 项。

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\-\0-9]+:[0-9]{12}:agent/agent-[0-9a-z]{17}$`

必需：是

### ClientToken

指定客户端令牌以确保使用此 API 操作的请求是等同的。如果您没有指定客户代币，则自动 DataSync 生成一个客户代币。

类型：字符串

模式：`[a-f0-9]{8}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{12}`

必需：是

### CloudWatchLogGroupArn

指定 Amazon CloudWatch 日志组的 ARN，用于监控和记录发现任务事件。

类型：字符串

长度约束：最大长度为 562。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):logs:[a-z\-\0-9]+:[0-9]{12}:log-group:([^\:*]*)(:\*)?$$`

必需：否

### Credentials

指定用于访问本地存储系统的管理界面的用户名和密码。

类型：[Credentials](#) 对象

必需：是

### Name

为您的本地存储系统指定一个熟悉的名称。

类型：字符串

长度限制：最小长度为 1。长度上限为 256。

模式：`^[\\p{L}\\p{M}\\p{N}\\s+=. _:@\\/- ]+$`

必需：否

### [ServerConfiguration](#)

指定连接本地存储系统的管理接口所需的服务器名称和网络端口。

类型：[DiscoveryServerConfiguration](#) 对象

必需：是

### [SystemType](#)

指定您希望 DataSync Discovery 收集的本地存储系统的相关信息类型。

#### Note

DataSyncDiscovery 目前支持运行 ONTAP 9.7 或更高版本的 NetApp 结构连接存储 (FAS) 和全闪存 FAS (AFF) 系统。

类型：字符串

有效值：NetAppONTAP

必需：是

### [Tags](#)

指定帮助您分类、筛选和搜索 AWS 资源的标签。我们建议至少为本地存储系统创建一个名称标签。

类型：[TagListEntry](#) 对象数组

数组成员：最少 0 项。最多 50 项。

必需：否

### 响应语法

```
{
  "StorageSystemArn": "string"
}
```

```
}
```

## 响应元素

如果此操作成功，则该服务将会发送回 HTTP 200 响应。

服务以 JSON 格式返回的以下数据。

### StorageSystemArn

您可以与 DataSync Discovery 一起使用的本地存储系统的 ARN。

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\n-0-9]+:[0-9]{12}:system/storage-system-[a-f0-9]{8}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{12}$`

## 错误

有关所有操作常见错误的信息，请参阅[常见错误](#)。

### InternalException

当AWS DataSync服务中发生错误时，会引发此异常。

HTTP 状态代码：500

### InvalidRequestException

当客户端提交格式错误的请求时，会引发此异常。

HTTP 状态代码：400

## 示例

### 示例请求

以下示例将本地存储系统添加到 DataSync Discovery。

```
{
  "ServerConfiguration": {
    "ServerHostname": "172.16.0.0",
    "ServerPort": 443
  },
  "SystemType": "NetAppONTAP",
  "AgentArns": [
    "arn:aws:datasync:us-east-1:111222333444:agent/agent-012345abcde012345"
  ],
  "CloudWatchLogGroupArn": "arn:aws:logs:us-east-1:111222333444:log-group:/aws/
datasync/discovery:*",
  "Tags": [
    {
      "Key": "Migration Plan",
      "Value": "1"
    }
  ],
  "Name": "MyOnPremStorage",
  "Credentials": {
    "Username": "admin",
    "Password": "1234"
  }
}
```

## 示例响应

响应会返回您刚刚添加到 DataSync Discovery 的本地存储系统的 ARN。

```
{
  "StorageSystemArn": "arn:aws:datasync:us-east-1:111222333444:system/storage-system-
abcdef01234567890"
}
```

## 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [AWS 命令行界面](#)
- [适用于 .NET 的 AWS SDK](#)
- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)



- [适用于 Java V2 的 AWS SDK](#)
- [AWSSDK 用于 JavaScript](#)
- [适用于 PHP V3 的 AWS SDK](#)
- [适用于 Python 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)

## CancelTaskExecution

停止正在执行的AWS DataSync任务。某些文件的传输突然中断。传输到目标的文件内容可能不完整或与源文件不一致。

但是，如果您使用相同的任务开始执行新任务并允许其完成，则目标上的文件内容将完整且一致。这适用于中断任务执行的其他意外故障。在所有这些情况下，当您开始执行下一个任务时，DataSync成功完成传输。

### 请求语法

```
{
  "TaskExecutionArn": "string"
}
```

### 请求参数

有关所有操作的通用参数的信息，请参阅[常用参数](#)。

请求接受采用 JSON 格式的以下数据。

#### TaskExecutionArn

要停止执行的 Amazon 资源名称 ( ARN ) 。

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\-\0-9]*:[0-9]{12}:task/task-[0-9a-f]{17}/execution/exec-[0-9a-f]{17}$`

必需：是

### 响应元素

如果此操作成功，则该服务会发送回带有空 HTTP 正文的 HTTP 200 响应。

### 错误

有关所有操作常见错误的信息，请参阅[常见错误](#)。

## InternalException

当AWS DataSync服务中发生错误时，会抛出此异常。

HTTP 状态代码：500

## InvalidRequestException

当客户端提交格式错误的请求时，会引发此异常。

HTTP 状态代码：400

## 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [AWS 命令行界面](#)
- [适用于 .NET 的 AWS SDK](#)
- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [AWS适用于JavaScript](#)
- [适用于 PHP V3 的 AWS SDK](#)
- [适用于 Python 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)

## CreateAgent

激活您在存储环境中部署的AWS DataSync代理。激活过程可将代理与您的 Agent 关联AWS 账户。

如果您尚未部署代理，请参阅以下主题以了解更多信息：

- [代理要求](#)
- [创建代理](#)

### Note

如果您在AWS存储服务之间传输，则不需要代DataSync理。

### 请求语法

```
{
  "ActivationKey": "string",
  "AgentName": "string",
  "SecurityGroupArns": [ "string" ],
  "SubnetArns": [ "string" ],
  "Tags": [
    {
      "Key": "string",
      "Value": "string"
    }
  ],
  "VpcEndpointId": "string"
}
```

### 请求参数

有关所有操作的通用参数的信息，请参阅[通用参数](#)。

请求接受采用 JSON 格式的以下数据。

#### [ActivationKey](#)

指定您的DataSync代理的激活密钥。如果您没有激活密钥，请参阅[激活您的代理](#)。

类型：字符串

长度约束：最大长度为 29。

模式：`[A-Z0-9]{5}(-[A-Z0-9]{5}){4}`

必需：是

### AgentName

指定您的 Agent 的名称。您可以在 DataSync 控制台中看到这个名字。

类型：字符串

长度限制：最小长度为 1。长度上限为 256。

模式：`^[a-zA-Z0-9\s+=. _:@/-]+$`

必需：否

### SecurityGroupArns

指定安全组的 Amazon 资源名称 (ARN)，该安全组在[使用虚拟私有云 \(VPC\) 终端节点](#)时保护您的任务的[网络接口](#)。

类型：字符串数组

数组成员：固定数量为 1 项。

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):ec2:[a-z\-\0-9]*:[0-9]{12}:security-group/sg-[a-f0-9]+$`

必需：否

### SubnetArns

指定使用 VPC 终端节点时要在其中运行 DataSync 任务的子网的 ARN。这是为您的传输 DataSync 创建和管理[网络接口](#)的子网。

类型：字符串数组

数组成员：固定数量为 1 项。

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):ec2:[a-z\-\0-9]*:[0-9]{12}:subnet/.*$`

必需：否

## Tags

指定帮助您分类、筛选和搜索 AWS 资源的标签。我们建议为您的代理创建至少一个标签。

类型：[TagListEntry](#) 对象数组

数组成员：最少 0 项。最多 50 项。

必需：否

## VpcEndpointId

指定您希望代理连接到的 VPC 端点 ID。例如，VPC 终端节点的样子 `vpce-01234d5aff67890e1`。

### Important

您使用的 VPC 终端节点必须包含 DataSync 服务名称（例如 `com.amazonaws.us-east-2.datasync`）。

类型：字符串

模式：`^vpce-[0-9a-f]{17}$`

必需：否

## 响应语法

```
{
  "AgentArn": "string"
}
```

## 响应元素

如果此操作成功，则该服务将会发送回 HTTP 200 响应。

服务以 JSON 格式返回的以下数据。

## AgentArn

您刚刚激活的代理的 ARN。使用[ListAgents](#)操作返回AWS 账户和中的代理列表AWS 区域。

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\\-0-9]+:[0-9]{12}:agent/agent-[0-9a-z]{17}$`

## 错误

有关所有操作常见错误的信息，请参阅[常见错误](#)。

### InternalException

当AWS DataSync服务中发生错误时，会引发此异常。

HTTP 状态代码：500

### InvalidRequestException

当客户端提交格式错误的请求时，会引发此异常。

HTTP 状态代码：400

## 示例

### 示例请求

以下示例激活DataSync代理并将其与您的AWS 账户相关联。

```
{
  "ActivationKey": "AAAAA-1AAAA-BB1CC-33333-EEEEEE",
  "AgentName": "MyAgent",
  "Tags": [{
    "Key": "Job",
    "Value": "TransferJob-1"
  }]
}
```

```
}
```

## 示例响应

响应返回已激活代理的 ARN。

```
{
  "AgentArn": "arn:aws:datasync:us-east-2:111222333444:agent/agent-0b0addbeef44baca3"
}
```

## 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [AWS 命令行界面](#)
- [适用于 .NET 的 AWS SDK](#)
- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [AWSSDK 用于 JavaScript](#)
- [适用于 PHP V3 的 AWS SDK](#)
- [适用于 Python 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)



## CreateLocationEfs

为AWS DataSync可以访问的 Amazon EFS 文件系统创建终端节点。有关更多信息，请参阅[Amazon EFS 创建位置](#)。

### 请求语法

```
{
  "AccessPointArn": "string",
  "Ec2Config": {
    "SecurityGroupArns": [ "string" ],
    "SubnetArn": "string"
  },
  "EfsFileSystemArn": "string",
  "FileSystemAccessRoleArn": "string",
  "InTransitEncryption": "string",
  "Subdirectory": "string",
  "Tags": [
    {
      "Key": "string",
      "Value": "string"
    }
  ]
}
```

### 请求参数

有关所有操作的通用参数的信息，请参阅[通用参数](#)。

请求接受采用 JSON 格式的以下数据。

#### [AccessPointArn](#)

指定用于访问 Amazon EFS 文件系统的访问点的 DataSync Amazon 资源名称 (ARN)。

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):elasticfilesystem:[a-z\-\0-9]+:[0-9]{12}:access-point/fsap-[0-9a-f]{8,40}$`

必需：否

## [Ec2Config](#)

指定用于访问 Amazon EFS 文件系统的子网和安全组DataSync。

类型：[Ec2Config](#) 对象

必需：是

## [EfsFilesystemArn](#)

指定 Amazon EFS 文件系统的 ARN。

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):elasticfilesystem:[a-z\-\0-9]*:[0-9]{12}:file-system/fs-.*$`

必需：是

## [FileSystemAccessRoleArn](#)

指定当AWS Identity and Access Management ( IAM ) 角色在挂载 Amazon EFS 文件系统时 DataSync担任的 ( IAM ) 角色。

类型：字符串

长度限制：最大长度为 2048。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):iam::[0-9]{12}:role/.*$`

必需：否

## [InTransitEncryption](#)

指定当 TLS 将数据复制DataSync到 Amazon EFS 文件系统或从文件系统复制数据时，您是否希望使用传输层安全性协议 ( TLS ) 1.2 加密。

如果您使用 `AccessPointArn` 指定访问点或使用 `FileSystemAccessRoleArn` 指定 IAM 角色，则必须将此参数设置为 `TLS1_2`。

类型：字符串

有效值：NONE | TLS1\_2

必需：否

### Subdirectory

指定 Amazon EFS 文件系统的挂载路径。这是DataSync读取或写入数据的位置（具体视源位置还是目标位置而定）。默认情况下，DataSync使用根目录，但您也可以包含子目录。

#### Note

必须使用正斜杠指定值（例如，/path/to/folder）。

类型：字符串

长度限制：最大长度为 4096。

模式：`^[a-zA-Z0-9_-\+\.\/\(\)\p{Zs}]*$`

必需：否

### Tags

指定表示要添加到资源的标签的键值对。该值可为空字符串。该值可帮助您管理、筛选和搜索资源。我们建议您为自己的位置创建名称标签。

类型：[TagListEntry](#) 对象数组

数组成员：最少 0 项。最多 50 项。

必需：否

### 响应语法

```
{
  "LocationArn": "string"
}
```

### 响应元素

如果此操作成功，则该服务将会发送回 HTTP 200 响应。

服务以 JSON 格式返回的以下数据。

### LocationArn

您创建的 Amazon EFS 文件系统位置的 Amazon 资源名称 (ARN)。

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\\-0-9]+:[0-9]{12}:location/loc-[0-9a-z]{17}$`

### 错误

有关所有操作常见错误的信息，请参阅[常见错误](#)。

#### InternalException

当 AWS DataSync 服务中发生错误时，会引发此异常。

HTTP 状态代码：500

#### InvalidRequestException

当客户端提交格式错误的请求时，会引发此异常。

HTTP 状态代码：400

### 示例

#### 示例请求

以下示例为 Amazon EFS 文件系统创建位置。

```
{
  "Ec2Config": {
    "SubnetArn": "arn:aws:ec2:us-east-2:11122233344:subnet/subnet-1234567890abcdef1",
    "SecurityGroupArns": [
      "arn:aws:ec2:us-east-2:11122233344:security-group/sg-1234567890abcdef2"
    ]
  }
}
```

```

    ]
  },
  "EfsFileSystemArn": "arn:aws:elasticfilesystem:us-east-2:111222333444:file-system/
fs-021345abcdef6789",
  "Subdirectory": "/mount/path",
  "Tags": [{
    "Key": "Name",
    "Value": "ElasticFileSystem-1"
  }]
}

```

示例请求：为受限的 Amazon EFS 文件系统创建位置

以下示例使用 Amazon EFS 为具有限制访问权限的 Amazon EFS 文件系统创建一个位置。在这种情况下，您可能需要在请求中指定 `AccessPointArn`、`FileSystemAccessRoleArn` 和 `InTransitEncryption` 的值。

```

{
  "AccessPointArn": "arn:aws:elasticfilesystem:us-east-2:111222333444:access-point/
fsap-1234567890abcdef0",
  "Ec2Config": {
    "SubnetArn": "arn:aws:ec2:us-east-2:111222333444:subnet/
subnet-1234567890abcdef1",
    "SecurityGroupArns": [
      "arn:aws:ec2:us-east-2:111222333444:security-group/sg-1234567890abcdef2"
    ]
  },
  "FileSystemAccessRoleArn": "arn:aws:iam::111222333444:role/
AwsDataSyncFullAccessNew",
  "InTransitEncryption": "TLS1_2",
  "LocationArn": "arn:aws:datasync:us-east-2:111222333444:location/loc-
abcdef01234567890",
  "LocationUri": "efs://us-east-2.fs-021345abcdef6789/",
  "Subdirectory": "/mount/path",
  "Tags": [{
    "Key": "Name",
    "Value": "ElasticFileSystem-1"
  }]
}

```

示例响应

响应返回 Amazon EFS 文件系统的位置 ARN。

```
{
  "LocationArn": "arn:aws:datasync:us-east-2:111222333444:location/
loc-12abcdef012345678"
}
```

## 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [AWS 命令行界面](#)
- [适用于 .NET 的 AWS SDK](#)
- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [AWSSDK 用于 JavaScript](#)
- [适用于 PHP V3 的 AWS SDK](#)
- [适用于 Python 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)

## CreateLocationFsxLustre

为 Amazon FSx for Lustre 文件系统创建终端节点

### 请求语法

```
{
  "FsxFilesystemArn": "string",
  "SecurityGroupArns": [ "string" ],
  "Subdirectory": "string",
  "Tags": [
    {
      "Key": "string",
      "Value": "string"
    }
  ]
}
```

### 请求参数

有关所有操作的通用参数的信息，请参阅[常用参数](#)。

请求接受采用 JSON 格式的以下数据。

#### [FsxFilesystemArn](#)

FSx for Lustre 文件系统的 Amazon 资源名称 (ARN)。

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):fsx:[a-z\-\0-9]*:[0-9]{12}:file-system/fs-.*$`

必需：是

#### [SecurityGroupArns](#)

用于配置 FSx for Lustre 文件系统的安全组的 Amazon 资源名称 (ARN)。

类型：字符串数组

数组成员：最少 1 项。最多 5 项。

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):ec2:[a-z\-\0-9]*:[0-9]{12}:security-group/sg-[a-f0-9]+$`

必需：是

### Subdirectory

位置路径中的子目录。FSx for Lustre 文件系统中用于从 FSx for Lustre 源位置读取数据或将数据写入 FSx for Lustre 目标的子目录。

类型：字符串

度度度度度度度度度度度度度度度度度度

模式：`^[a-zA-Z0-9_\-\+\.\(\)\$\p{Zs}]+$`

必需：否

### Tags

表示要添加到资源的标签的键值对。该值可为空字符串。该值可帮助您管理、筛选和搜索资源。我们建议您为自己的位置创建名称标签。

类型：[TagListEntry](#) 对象数组

数组成员：最少 0 项。最多 50 项。

必需：否

### 响应语法

```
{
  "LocationArn": "string"
}
```

### 响应元素

如果此操作成功，则该服务将会发送回 HTTP 200 响应。



服务以 JSON 格式返回的以下数据。

### LocationArn

创建的 FSx for Lustre 文件系统位置的 Amazon 资源名称 (ARN)。

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\\-0-9]+:[0-9]{12}:location/loc-[0-9a-z]{17}$`

### 错误

有关所有操作常见错误的信息，请参阅[常见错误](#)。

#### InternalException

当AWS DataSync服务中发生错误时，会抛出此异常。

HTTP 状态代码：500

#### InvalidRequestException

当客户端提交格式错误的请求时，会引发此异常。

HTTP 状态代码：400

### 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [AWS 命令行界面](#)
- [适用于 .NET 的 AWS SDK](#)
- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [AWS适用于JavaScript](#)
- [适用于 PHP V3 的 AWS SDK](#)

- [适用于 Python 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)

## CreateLocationFsxOntap

为 Amazon FSx for NetApp ONTAP 文件系统创建终端节点 AWS DataSync，用于创建终端节点。有关更多信息，请参阅 [《为 FSx for ONTAP》创建位置](#)。

### 请求语法

```
{
  "Protocol": {
    "NFS": {
      "MountOptions": {
        "Version": "string"
      }
    },
    "SMB": {
      "Domain": "string",
      "MountOptions": {
        "Version": "string"
      },
      "Password": "string",
      "User": "string"
    }
  },
  "SecurityGroupArns": [ "string" ],
  "StorageVirtualMachineArn": "string",
  "Subdirectory": "string",
  "Tags": [
    {
      "Key": "string",
      "Value": "string"
    }
  ]
}
```

### 请求参数

有关所有操作的通用参数的信息，请参阅 [常用参数](#)。

请求接受采用 JSON 格式的以下数据。

#### Protocol

指定 AWS DataSync 用于访问您的 Amazon FSx 文件系统的数据传输协议。

类型：[FsxProtocol](#) 对象

必需：是

### [SecurityGroupArns](#)

指定提供对文件系统首选子网访问权限的 Amazon EC2 安全组。

安全组必须允许以下端口上的出站流量（具体您使用的协议而定）：

- 网络文件系统（NFS）：TCP 端口 111、635 和 2049
- 服务器消息块（SMB）：TCP 端口 445

文件系统的安全组还必须允许同一端口上的入站流量。

类型：字符串数组

数组成员：最少 1 项。最多 5 项。

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):ec2:[a-z\-\0-9]*:[0-9]{12}:security-group/sg-[a-f0-9]+$`

必需：是

### [StorageVirtualMachineArn](#)

指定要将数据复制到或从中复制数据的文件系统存储虚拟机（SVM）的 ARN。

类型：字符串

长度约束：最大长度为 162。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):fsx:[a-z\-\0-9]*:[0-9]{12}:storage-virtual-machine/fs-[0-9a-f]+/svm-[0-9a-f]{17,}$`

必需：是

### [Subdirectory](#)

指定您将在其中复制数据的指向 SVM 中的文件共享路径。

您可以指定连接路径（也称为挂载点）、qtree 路径（用于 NFS 文件共享）或共享名称（用于 SMB 文件共享）。例如，您的挂载路径可能是 /vol1、/vol1/tree1 或 /share1。

**Note**

不要在 SVM 的根卷中指定连接路径。有关更多信息，请参阅《Amazon FSx for ONTAP 用户指南》中的管理 FSx fNetApp or ONTAP [存储虚拟机](#)。

类型：字符串

长度限制：最大长度为 255。

模式：`^[^\u0000\u0085\u2028\u2029\r\n]{1,255}$`

必需：否

## Tags

指定帮助您分类、筛选和搜索 AWS 资源的标签。我们建议至少为您的位置创建一个名称标签。

类型：[TagListEntry](#) 对象数组

数组成员：最少 0 项。最多 50 项。

必需：否

## 响应语法

```
{
  "LocationArn": "string"
}
```

## 响应元素

如果此操作成功，则该服务将会发送回 HTTP 200 响应。

服务以 JSON 格式返回的以下数据。

### [LocationArn](#)

指定您创建的 FSx for ONTAP 文件系统位置的 ARN。

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\\-0-9]+:[0-9]{12}:location/loc-[0-9a-z]{17}$`

## 错误

有关所有操作常见错误的信息，请参阅[常见错误](#)。

### InternalException

当AWS DataSync服务中发生错误时，会抛出此异常。

HTTP 状态代码：500

### InvalidRequestException

当客户端提交格式错误的请求时，会引发此异常。

HTTP 状态代码：400

## 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [AWS 命令行界面](#)
- [适用于 .NET 的 AWS SDK](#)
- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [AWS适用于JavaScript](#)
- [适用于 PHP V3 的 AWS SDK](#)
- [适用于 Python 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)

## CreateLocationFsxOpenZfs

为 Amazon FSx for OpenZFS 文件系统创建终端节点。AWS DataSync有关更多信息，请参阅为 [OpenZFS 的 FSx 创建位置](#)。

### Note

该CreateLocationFsxOpenZfs操作SMB不支持与相关的请求参数。

### 请求语法

```
{
  "FsxFilesystemArn": "string",
  "Protocol": {
    "NFS": {
      "MountOptions": {
        "Version": "string"
      }
    },
    "SMB": {
      "Domain": "string",
      "MountOptions": {
        "Version": "string"
      },
      "Password": "string",
      "User": "string"
    }
  },
  "SecurityGroupArns": [ "string" ],
  "Subdirectory": "string",
  "Tags": [
    {
      "Key": "string",
      "Value": "string"
    }
  ]
}
```

### 请求参数

有关所有操作的通用参数的信息，请参阅[常用参数](#)。

请求接受采用 JSON 格式的以下数据。

### FsxFilesystemArn

FSx for OpenZFS 文件系统的 Amazon 资源名称 ( ARN ) 。

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):fsx:[a-z\-\0-9]*:[0-9]{12}:file-system/fs-.*$`

必需：是

### Protocol

AWS DataSync 用于访问文件系统的协议类型。

类型：[FsxProtocol](#) 对象

必需：是

### SecurityGroupArns

用于配置 FSx for OpenZFS 文件系统的安全组的 ARN。

类型：字符串数组

数组成员：最少 1 项。最多 5 项。

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):ec2:[a-z\-\0-9]*:[0-9]{12}:security-group/sg-[a-f0-9]+$`

必需：是

### Subdirectory

位置路径中必须以 /fsx 开头的子目录。DataSync使用该子目录来读取或写入数据 ( 具体视文件系统是源位置还是目标位置而定 ) 。

类型：字符串

长度限制：最大长度为 4096。



模式：`^[^\u0000\u0085\u2028\u2029\r\n]{1,4096}$`

必需：否

## Tags

表示要添加到资源的标签的键值对。该值可为空字符串。该值可帮助您管理、筛选和搜索资源。我们建议您为自己的位置创建名称标签。

类型：[TagListEntry](#) 对象数组

数组成员：最少 0 项。最多 50 项。

必需：否

## 响应语法

```
{
  "LocationArn": "string"
}
```

## 响应元素

如果此操作成功，则该服务将会发送回 HTTP 200 响应。

服务以 JSON 格式返回的以下数据。

## [LocationArn](#)

您创建的 FSx for OpenZFS 文件系统位置的 ARN。

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\n-0-9]+:[0-9]{12}:location/loc-[0-9a-z]{17}$`

## 错误

有关所有操作常见错误的信息，请参阅[常见错误](#)。

## InternalException

当AWS DataSync服务中发生错误时，会抛出此异常。

HTTP 状态代码：500

## InvalidRequestException

当客户端提交格式错误的请求时，会引发此异常。

HTTP 状态代码：400

## 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [AWS 命令行界面](#)
- [适用于 .NET 的 AWS SDK](#)
- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [AWS适用于JavaScript](#)
- [适用于 PHP V3 的 AWS SDK](#)
- [适用于 Python 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)

## CreateLocationFsxWindows

为 Amazon FSx for Windows File Server 文件系统创建终端节点。

### 请求语法

```
{
  "Domain": "string",
  "FsxFilesystemArn": "string",
  "Password": "string",
  "SecurityGroupArns": [ "string" ],
  "Subdirectory": "string",
  "Tags": [
    {
      "Key": "string",
      "Value": "string"
    }
  ],
  "User": "string"
}
```

### 请求参数

有关所有操作的通用参数的信息，请参阅[通用参数](#)。

请求接受采用 JSON 格式的以下数据。

#### Domain

指定 FSx for Windows File Server 所属的 Windows 域的名称。

类型：字符串

长度限制：最大长度为 253。

模式：`^[A-Za-z0-9](\\.|-+)?[A-Za-z0-9]{0,252}$`

必需：否

#### FsxFilesystemArn

指定 FSx for Windows File Server 文件系统的 Amazon 资源名称 ( ARN )。

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):fsx:[a-z\-\0-9]*:[0-9]{12}:file-system/fs-.*$`

必需：是

## Password

指定有权访问文件系统中文件和文件夹的用户的密码。

类型：字符串

长度限制：最大长度为 104。

模式：`^\.{0,104}$`

必需：是

## SecurityGroupArns

指定用于访问您的文件系统首选子网的 ARN。

### Note

如果您选择不允许从内部进行连接的安全组，执行以下操作之一：

- 将安全组配置为允许其内部通信。
- 选择可以与装载目标的安全组通信的其他安全组。

类型：字符串数组

数组成员：最少 1 项。最多 5 项。

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):ec2:[a-z\-\0-9]*:[0-9]{12}:security-group/sg-[a-f0-9]+$`

必需：是

## Subdirectory

使用正斜杠指定文件系统的挂载路径。这是DataSync读取或写入数据的位置（具体视源位置还是目标位置而定）。

类型：字符串

长度限制：最大长度为 4096。

模式：`^[a-zA-Z0-9_\-\.\/\(\)\$\p{Zs}]+$`

必需：否

## Tags

指定帮助您分类、筛选和搜索 AWS 资源的标签。我们建议至少为您的位置创建一个名称标签。

类型：[TagListEntry](#) 对象数组

数组成员：最少 0 项。最多 50 项。

必需：否

## User

指定有权访问您的文件系统中文件、文件夹和元数据的用户。

有关选择具有足够权限的用户的信息，请参阅[所需权限](#)。

类型：字符串

长度限制：最大长度为 104。

模式：`^[^\x5B\x5D\\\/:;|=, +*?]{1,104}$`

必需：是

## 响应语法

```
{
  "LocationArn": "string"
}
```

## 响应元素

如果此操作成功，则该服务将会发送回 HTTP 200 响应。

服务以 JSON 格式返回的以下数据。

### LocationArn

您创建的 FSx for Windows File Server 文件系统位置的 ARN。

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\\-0-9]+:[0-9]{12}:location/loc-[0-9a-z]{17}$`

## 错误

有关所有操作常见错误的信息，请参阅[常见错误](#)。

### InternalException

当AWS DataSync服务中发生错误时，会引发此异常。

HTTP 状态代码：500

### InvalidRequestException

当客户端提交格式错误的请求时，会引发此异常。

HTTP 状态代码：400

## 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [AWS 命令行界面](#)
- [适用于 .NET 的 AWS SDK](#)
- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)

- [适用于 Java V2 的 AWS SDK](#)
- [AWSSDK 用于 JavaScript](#)
- [适用于 PHP V3 的 AWS SDK](#)
- [适用于 Python 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)

## CreateLocationHdfs

为 Hadoop Distributed File System ( HDFS ) 创建终端

### 请求语法

```
{
  "AgentArns": [ "string" ],
  "AuthenticationType": "string",
  "BlockSize": number,
  "KerberosKeytab": blob,
  "KerberosKrb5Conf": blob,
  "KerberosPrincipal": "string",
  "KmsKeyProviderUri": "string",
  "NameNodes": [
    {
      "Hostname": "string",
      "Port": number
    }
  ],
  "QopConfiguration": {
    "DataTransferProtection": "string",
    "RpcProtection": "string"
  },
  "ReplicationFactor": number,
  "SimpleUser": "string",
  "Subdirectory": "string",
  "Tags": [
    {
      "Key": "string",
      "Value": "string"
    }
  ]
}
```

### 请求参数

有关所有操作的通用参数的信息，请参阅[通用参数](#)。

请求接受采用 JSON 格式的以下数据。



## [AgentArns](#)

用于连接到 HDFS 集群的代理的 Amazon 资源名称 ( ARN )。

类型：字符串数组

数组成员：最少 1 项。最多 4 项。

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\\-0-9]+:[0-9]{12}:agent/agent-[0-9a-z]{17}$`

必需：是

## [AuthenticationType](#)

用于确定用户身份的身份验证类型。

类型：字符串

有效值：SIMPLE | KERBEROS

必需：是

## [BlockSize](#)

要写入 HDFS 集群的数据块大小。块大小必须是 512 字节的倍数。默认块大小为 128 兆字节 ( MiB )。

类型：整数

有效值：10485476 值为 10485 最大值：10737473741824 值

必需：否

## [KerberosKeytab](#)

Kerberos 密钥表 ( keytab )，其中包含定义的 Kerberos 主体和加密密钥之间的映射。您可以通过提供文件的地址从文件加载密钥表。如果您使用的 AWS CLI，它将执行 base64 编码。否则，

### Note

如果为 `AuthenticationType` 指定了 `KERBEROS`，则此参数为必填项。

类型：Base64 编码的二进制数据对象

长度限制：最大长度长度长度长度长度长度长度长度

必需：否

### KerberosKrb5Conf

包含 Kerberos 配置信息的 `krb5.conf` 文件。您可以通过提供 `krb5.conf` 文件的地址来加载文件。如果您使用的 AWS CLI，它将执行 base64 编码。否则，

#### Note

如果为 `AuthenticationType` 指定了 `KERBEROS`，则此参数为必填项。

类型：Base64 编码的二进制数据对象

长度限制：最大长度长度长度长度长度长度为 13106

必需：否

### KerberosPrincipal

有权访问 HDFS 集群上的文件和文件夹的 Kerberos 主体。

#### Note

如果为 `AuthenticationType` 指定了 `KERBEROS`，则此参数为必填项。

类型：字符串

长度限制：最小长度为 1。长度上限为 256。

模式：`^\.+`

必需：否

### KmsKeyProviderUri

HDFS 集群的密钥管理服务器 ( KMS ) 的 URI。

类型：字符串

长度限制：最小长度为 1。最大长度为 255。

模式：`^kms:\//http[s]?@((([a-zA-Z0-9\-\_]*[a-zA-Z0-9])\.)*)*([A-Za-z0-9\-\_]*[A-Za-z0-9])(;((([a-zA-Z0-9\-\_]*[a-zA-Z0-9])\.)*)*([A-Za-z0-9\-\_]*[A-Za-z0-9])))*:[0-9]{1,5}\//kms$`

必需：否

## NameNodes

管理 HDFS 命名空间的 FS 命名空间。NameNodeNameNode 执行打开、关闭和重命名文件和目录等操作。NameNode 包含将数据块映射到的信息 DataNodes。您只能使用一个 NameNode。

类型：[HdfsNameNode](#) 对象数组

数组成员：最少 1 项。

必需：是

## QopConfiguration

保护质量 ( QOP ) 配置指定在 Hadoop Distributed File System ( HDFS ) 集群上配置的远程程序调用 ( RPC ) 和数据传输保护设置。如果未指定 QopConfiguration，则 RpcProtection 和 DataTransferProtection 默认值为 PRIVACY。如果设置 RpcProtection 或 DataTransferProtection，则其他参数采用相同的值。

类型：[QopConfiguration](#) 对象

必需：否

## ReplicationFactor

写入 HDFS 集群时要将数据复制到的数量。DataNodes 默认情况下，数据将复制到三项 DataNodes。

类型：整数

有效范围：最小值为 1。最大值值值值值值值值值

必需：否

## SimpleUser

用于在主机操作系统上标识客户端的用户名。

**Note**

如果为 `AuthenticationType` 指定了 `SIMPLE`，则此参数为必填项。

类型：字符串

长度限制：最小长度为 1。长度上限为 256。

模式：`^[_.A-Za-z0-9][-_.A-Za-z0-9]*$`

必需：否

### Subdirectory

HDFS 集群中的子目录。此子目录用于从 HDFS 集群读取数据或向其写入数据。如果未指定子目录，则默认为 `/`。

类型：字符串

长度长度长度为 4096 长度为 4096。

模式：`^[a-zA-Z0-9_\-\.\/\(\)\$\p{Zs}]+$`

必需：否

### Tags

表示要添加到位置的标签的键值对。该值可为空字符串。建议您使用标签来命名资源。

类型：[TagListEntry](#) 对象数组

数组成员：最少 0 项。最多 50 项。

必需：否

### 响应语法

```
{
  "LocationArn": "string"
}
```

## 响应元素

如果此操作成功，则该服务将会发送回 HTTP 200 响应。

服务以 JSON 格式返回的以下数据。

### LocationArn

创建的源 HDFS 集群的 ARN 集群的 ARN。

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\\-0-9]+:[0-9]{12}:location/loc-[0-9a-z]{17}$`

## 错误

有关所有操作常见错误的信息，请参阅[常见错误](#)。

### InternalException

当AWS DataSync服务中发生错误时，会引发此异常。

HTTP 状态代码：500

### InvalidRequestException

当客户端提交格式错误的请求时，会引发此异常。

HTTP 状态代码：400

## 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [AWS 命令行界面](#)
- [适用于 .NET 的 AWS SDK](#)
- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)

- [适用于 Java V2 的 AWS SDK](#)
- [AWS用于 SDK JavaScript](#)
- [适用于 PHP V3 的 AWS SDK](#)
- [适用于 Python 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)

## CreateLocationNfs

定义网络文件系统 (NFS) 服务器上可读取或可写入的文件系统。

### 请求语法

```
{
  "MountOptions": {
    "Version": "string"
  },
  "OnPremConfig": {
    "AgentArns": [ "string" ]
  },
  "ServerHostname": "string",
  "Subdirectory": "string",
  "Tags": [
    {
      "Key": "string",
      "Value": "string"
    }
  ]
}
```

### 请求参数

有关所有操作的通用参数的信息，请参阅[通用参数](#)。

请求接受采用 JSON 格式的以下数据。

#### [MountOptions](#)

可用于挂载 NFS 共享的DataSync NFS 挂载选项。

类型：[NfsMountOptions](#) 对象

必需：否

#### [OnPremConfig](#)

包含用于连接到 NFS 服务器的代理的 Amazon 资源名称 (ARN) 列表。

如果您要将数据复制到设备或从AWS Snowcone设备复制数据，请参阅[NFS Server on \( 上的 NFS 服务器 \)](#)，了解AWS Snowcone更多信息。

类型：[OnPremConfig](#) 对象

必需：是

### [ServerHostname](#)

NFS 服务器的名称。该值是 NFS 服务器的 IP 地址或域名服务 (DNS) 名称。在本地安装的代理使用此主机名在网络中挂载 NFS 服务器。

如果您要将数据复制到设备或从AWS Snowcone设备复制数据，请参阅 [NFS Server on \( 上的 NFS 服务器 \)](#)，了解AWS Snowcone更多信息。

#### Note

此名称必须符合 DNS 标准，或者必须是 IP 版本 4 (IPv4) 地址。

类型：字符串

长度限制：最大长度为 255。

模式： $^([a-zA-Z0-9\-\_]*[a-zA-Z0-9])\.[A-Za-z0-9\-\_]*[A-Za-z0-9])\$$

必需：是

### [Subdirectory](#)

NFS 文件系统中的子目录，用于从 NFS 源位置读取数据或将数据写入 NFS 目标。NFS 路径应该是由 NFS 服务器导出的路径，或者是该路径的一个子目录。该路径应使其能由网络中的其他 NFS 客户端挂载。

要查看 NFS 服务器导出的所有路径，请从可访问服务器的 NFS 客户端运行“`showmount -e nfs-server-name`”。您可以指定显示在结果中的任何目录以及该目录的任何子目录。确保无需 Kerberos 身份验证即可访问 NFS 导出。

要传输指定文件夹中的所有数据，DataSync 需要有权读取所有数据。为确保这一点，请为 NFS 导出配置 `no_root_squash`，或者确保需要的所有文件的权限DataSync允许所有用户进行读取访问。通过执行上述任一操作，将允许代理读取这些文件。要使代理能够访问目录，您还必须启用全部执行访问权限。

如果您要将数据复制到设备或从AWS Snowcone设备复制数据，请参阅 [NFS Server on \( 上的 NFS 服务器 \)](#)，了解AWS Snowcone更多信息。



有关 NFS 导出配置的信息，请参阅 [18.7。Red Hat Enterprise Linux 文档中的 /etc/exports 配置文件](#)。

类型：字符串

长度约束：最大长度为 4096。

模式：`^[a-zA-Z0-9_\-\.\/\(\)\p{Zs}]+$`

必需：是

## Tags

表示要添加到位置的标签的键值对。该值可为空字符串。建议您使用标签来命名资源。

类型：[TagListEntry](#) 对象数组

数组成员：最少 0 项。最多 50 项。

必需：否

## 响应语法

```
{
  "LocationArn": "string"
}
```

## 响应元素

如果此操作成功，则该服务将会发送回 HTTP 200 响应。

服务以 JSON 格式返回的以下数据。

### [LocationArn](#)

创建的 NFS 文件系统位置的 Amazon 资源名称 (ARN)。

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\-\0-9]+:[0-9]{12}:location/loc-[0-9a-z]{17}$`

## 错误

有关所有操作常见错误的信息，请参阅[常见错误](#)。

### InternalException

当AWS DataSync服务中发生错误时，会抛出此异常。

HTTP 状态代码：500

### InvalidRequestException

当客户端提交格式错误的请求时，会引发此异常。

HTTP 状态代码：400

## 示例

### 示例

以下示例使用指定的 NFS 版本作为装载选项为 NFS 文件系统创建端点。

### 示例请求

```
{
  "MountOptions": {
    "Version": : "NFS4_0"
  },
  "OnPremConfig": {
    "AgentArn": [ "arn:aws:datasync:us-east-2:111222333444:agent/agent-0b0addbeef44b3nfs" ]
  },
  "ServerHostname": "MyServer@amazon.com",
  "Subdirectory": "/MyFolder",
  "Tags": [
    {
      "Key": "Name",
      "Value": "ElasticFileSystem-1"
    }
  ]
}
```

## 示例

响应将返回 NFS 位置的 Amazon 资源名称 (ARN)。

## 示例响应

```
{
  "LocationArn": "arn:aws:datsync:us-east-2:111222333444:location/
loc-07db7abfc326c50aa"
}
```

## 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [AWS 命令行界面](#)
- [适用于 .NET 的 AWS SDK](#)
- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [AWS适用于 SDKJavaScript](#)
- [适用于 PHP V3 的 AWS SDK](#)
- [适用于 Python 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)

## CreateLocationObjectStorage

为AWS DataSync可以访问进行传输的对象存储系统创建端点。有关更多信息，请参阅[创建对象存储位置](#)。

### 请求语法

```
{
  "AccessKey": "string",
  "AgentArns": [ "string" ],
  "BucketName": "string",
  "SecretKey": "string",
  "ServerCertificate": blob,
  "ServerHostname": "string",
  "ServerPort": number,
  "ServerProtocol": "string",
  "Subdirectory": "string",
  "Tags": [
    {
      "Key": "string",
      "Value": "string"
    }
  ]
}
```

### 请求参数

有关所有操作的通用参数的信息，请参阅[通用参数](#)。

请求接受采用 JSON 格式的以下数据。

#### [AccessKey](#)

如果需要凭证来向对象存储服务器进行身份验证，则指定访问密钥（例如，用户名）。

类型：字符串

长度限制：最小长度为 1。最大长度为 200。

模式： $^{\wedge}.\+ \$$

必需：否

## AgentArns

指定可以安全连接到您所在位置的 Amazon 资源名称 ( ARN )。DataSync

类型：字符串数组

数组成员：最少 1 项。最多 4 项。

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\-\0-9]+:[0-9]{12}:agent/agent-[0-9a-z]{17}$`

必需：是

## BucketName

指定传输中涉及的对象桶的名称。

类型：字符串

长度限制：最小长度为 3。长度上限为 63。

模式：`^[a-zA-Z0-9_\-\+\.\^(\)\$\p{Zs}]+$`

必需：是

## SecretKey

如果需要凭证来向对象存储服务器进行身份验证，则指定私有密钥（例如，密码）。

类型：字符串

长度限制：最小长度为 1。最大长度为 200。

模式：`^.+`

必需：否

## ServerCertificate

指定一个包含用于签署对象存储服务器证书的证书的文件（例如，`file:///home/user/.ssh/storage_sys_certificate.pem`）。您指定的文件必须包括以下参数：

- 签名证书颁发机构 (CA) 的证书 ( CA )。
- 任何中间证书

- base64 编码
- 一个 .pem 扩展

该文件最多可以是 32768 字节（在 base64 编码之前）。

要使用此参数，请配置 `ServerProtocol` 为 HTTPS。

类型：Base64 编码的二进制数据对象

长度约束：最大长度为 32768。

必需：否

### ServerHostname

指定对象存储服务器的域名或 IP 地址。DataSync 代理使用此主机名在网络中挂载对象存储服务。

类型：字符串

长度限制：最大长度为 255。

模式：`^(([a-zA-Z0-9\-.]*[a-zA-Z0-9])\.)*([A-Za-z0-9\-.]*[A-Za-z0-9])$`

必需：是

### ServerPort

指定对象存储服务器接受入站网络流量的端口（例如，端口 443）。

类型：整数

有效范围：最小值为 1。最大值为 65536。

必需：否

### ServerProtocol

指定对象存储服务器通信时使用的协议。

类型：字符串

有效值：HTTPS | HTTP

必需：否

## Subdirectory

指定对象存储服务器的对象前缀。如果这是源位置，则DataSync仅复制具有此前缀的对象。如果这是目标位置，则DataSync写入具有此前缀的所有对象。

类型：字符串

长度约束：最大长度为 4096。

模式：`^[a-zA-Z0-9_\-\.\/\(\)\p{Zs}]*$`

必需：否

## Tags

指定表示要添加到资源的标签的键值对。标签可帮助您管理、筛选和搜索资源。我们建议为您的位置创建一个名称标签。

类型：[TagListEntry](#) 对象数组

数组成员：最少 0 项。最多 50 项。

必需：否

## 响应语法

```
{
  "LocationArn": "string"
}
```

## 响应元素

如果此操作成功，则该服务将会发送回 HTTP 200 响应。

服务以 JSON 格式返回的以下数据。

## LocationArn

指定您创建的对象存储系统位置的 ARN。

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\\-0-9]+:[0-9]{12}:location/loc-[0-9a-z]{17}$`

## 错误

有关所有操作常见错误的信息，请参阅[常见错误](#)。

### InternalException

当AWS DataSync服务中发生错误时，会引发此异常。

HTTP 状态代码：500

### InvalidRequestException

当客户端提交格式错误的请求时，会引发此异常。

HTTP 状态代码：400

## 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [AWS 命令行界面](#)
- [适用于 .NET 的 AWS SDK](#)
- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [AWSSDK 用于 JavaScript](#)
- [适用于 PHP V3 的 AWS SDK](#)
- [适用于 Python 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)



## CreateLocationS3

位置是 Amazon S3 桶的 Amazon 终端节点。AWS DataSync 可以将该位置用作复制数据的源或目的地。

### Important

创建位置之前，请确保阅读以下部分：

- [Amazon S3 位置的存储类注意事项](#)
- [使用时评估 S3 请求成本 DataSync](#)

有关更多信息，请参阅[创建 Amazon S3 位置](#)。

### 请求语法

```
{
  "AgentArns": [ "string" ],
  "S3BucketArn": "string",
  "S3Config": {
    "BucketAccessRoleArn": "string"
  },
  "S3StorageClass": "string",
  "Subdirectory": "string",
  "Tags": [
    {
      "Key": "string",
      "Value": "string"
    }
  ]
}
```

### 请求参数

有关所有操作的通用参数的信息，请参阅[通用参数](#)。

请求接受采用 JSON 格式的以下数据。



## S3StorageClass

在将此位置用作任务目标时，您希望在其中存储文件的 Amazon S3 存储类。对于中的桶AWS 区域，默认存储类为“标准”。对于开启的 Amazon AWS Outposts 存储桶，默认存储类为 AWS S3 Outposts s。

有关 S3 存储类的更多信息，请参阅 [Amazon S3 存储类](#)。有些存储类别的行为会影响您的 S3 存储成本。有关详细信息，请参阅在 [中的使用 S3 存储类时的注意事项DataSync](#)。

类型：字符串

有效值：STANDARD | STANDARD\_IA | ONEZONE\_IA | INTELLIGENT\_TIERING | GLACIER | DEEP\_ARCHIVE | OUTPOSTS | GLACIER\_INSTANT\_RETRIEVAL

必需：否

## Subdirectory

Amazon S3 存储桶中的子目录。Amazon S3 中的此子目录用于从 S3 源位置读取数据或将数据写入 S3 目标。

类型：字符串

长度4096。

模式：`^[a-zA-Z0-9_\-\+\.\/\(\)\p{Zs}]*$`

必需：否

## Tags

表示要添加到位置的标签的键值对。该值可为空字符串。建议您使用标签来命名资源。

类型：[TagListEntry](#) 对象数组

数组成员：最少 0 项。最多 50 项。

必需：否

## 响应语法

```
{
  "LocationArn": "string"
}
```

```
}
```

## 响应元素

如果此操作成功，则该服务将会发送回 HTTP 200 响应。

服务以 JSON 格式返回的以下数据。

### LocationArn

创建的 Amazon S3 桶的 Amazon 资源名称 ( ARN ) 。

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\\-0-9]+:[0-9]{12}:location/loc-[0-9a-z]{17}$`

## 错误

有关所有操作常见错误的信息，请参阅[常见错误](#)。

### InternalException

当AWS DataSync服务中发生错误时，会引发此异常。

HTTP 状态代码：500

### InvalidRequestException

当客户端提交格式错误的请求时，会引发此异常。

HTTP 状态代码：400

## 示例

第 1 步 允许承担写入存储桶的 IAM 角色。

以下示例显示了最简单的策略，该策略授予AWS DataSync访问目标 Amazon S3 存储桶所需的权限，然后是该create-location-s3-iam-role策略所关联的 IAM 角色。

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Principal": {
        "Service": "datasync.amazonaws.com"
      },
      "Action": "sts:AssumeRole"
    }
  ]
}
```

```
"Role": {
  "Path": "/",
  "RoleName": "MyBucketAccessRole",
  "RoleId": "role-id",
  "Arn": "arn:aws:iam::account-id:role/MyBucketAccessRole",
  "CreateDate": "2018-07-27T02:49:23.117Z",
  "AssumeRolePolicyDocument": {
    "Version": "2012-10-17",
    "Statement": [
      {
        "Effect": "Allow",
        "Principal": {
          "Service": "datasync.amazonaws.com"
        },
        "Action": "sts:AssumeRole"
      }
    ]
  }
}
```

第 2 步 允许创建的 IAM 角色写入存储桶。

将具有足够权限访问存储桶的策略附加到该角色。此类策略的一个示例是 `AWSDataSyncFullAccess` 托管策略。

有关更多信息，请参阅 IAM 控制台 [AWSDataSyncFullAccess](#) 中的。

您无需创建该策略。它由管理AWS，因此您只需在attach-role-policy命令中指定其 ARN 即可。

```
IAM_POLICY_ARN='arn:aws:iam::aws:policy/AWSDataSyncFullAccess'
```

### 第 3 步 为 Amazon S3 存储桶创建终端节点

以下示例为 Amazon S3 存储桶创建终端节点。

创建 S3 终端节点时，类似于以下第二个示例的响应会返回新 Amazon S3 位置的亚马逊资源名称 (ARN)。

#### 示例请求

```
{
  "S3BucketArn": "arn:aws:s3:::MyBucket",
  "S3Config": {
    "BucketAccessRoleArn": "arn:aws:iam::111222333444:role/MyBucketAccessRole",
  },
  "S3StorageClass": "STANDARD",
  "Subdirectory": "/MyFolder",
  "Tags": [
    {
      "Key": "Name",
      "Value": "s3Bucket-1"
    }
  ]
}
```

#### 示例响应

```
{
  "LocationArn": "arn:aws:datasync:us-east-2:111222333444:location/loc-07db7abfc326c50s3"
}
```

### 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [AWS 命令行界面](#)

- [适用于 .NET 的 AWS SDK](#)
- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [AWSSDK 用于 JavaScript](#)
- [适用于 PHP V3 的 AWS SDK](#)
- [适用于 Python 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)

## CreateLocationSmb

为可以访问传输的服务器消息块 (SMB) 文件服务器创建AWS DataSync可以访问的服务器消息块 (SMB) 文件服务器的端点。有关更多信息，请参阅[创建 SMB 项](#)。

### 请求语法

```
{
  "AgentArns": [ "string" ],
  "Domain": "string",
  "MountOptions": {
    "Version": "string"
  },
  "Password": "string",
  "ServerHostname": "string",
  "Subdirectory": "string",
  "Tags": [
    {
      "Key": "string",
      "Value": "string"
    }
  ],
  "User": "string"
}
```

### 请求参数

有关所有操作的通用参数的信息，请参阅[通用参数](#)。

请求接受采用 JSON 格式的以下数据。

#### [AgentArns](#)

指定DataSync要连接到 SMB 文件服务器的一个或多个代理。您可以使用代理的 Amazon 资源名称 (ARN) 指定代理项。

类型：字符串数组

数组成员：最少 1 项。最多项项项项项项项项项项。

长度约束：最大长度为 128。



模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\\-0-9]+:[0-9]{12}:agent/agent-[0-9a-z]{17}$`

必需：是

## Domain

指定您的 SMB 文件服务器所属的 Windows 域名称。

有关更多信息，请参阅 SMB 的[必需权限](#)。

类型：字符串

最大长度限制：最大长度

模式：`^[A-Za-z0-9](\\.|-+)?[A-Za-z0-9]{0,252}$`

必需：否

## MountOptions

指定用于访问SMB 文件服务器的 DataSync SMB 协议版本。

类型：[SmbMountOptions](#) 对象

必需：否

## Password

指定可以挂载要挂载 SMB 文件服务器的用户的密码。

有关更多信息，请参阅 SMB 的[必需权限](#)。

类型：字符串

最大长度限制：最大长度

模式：`^.{0,104}$`

必需：是

## ServerHostname

指定要挂载DataSync代理的 SMB 文件服务器的域名服务 ( DNS ) 地址。

**Note**

您无法指定 IP 版本 6 ( IPv6 ) 地址。

类型：字符串

长度限制：最大长度为 255。

模式：`^(([a-zA-Z0-9\-\_]*[a-zA-Z0-9])\.)*([A-Za-z0-9\-\_]*[A-Za-z0-9])$`

必需：是

**Subdirectory**

指定您的 SMB 文件服务器导出的共享的名称，该共享DataSync将在其中读取或写入数据。可以在共享路径中包含子目录（例如，`/path/to/subdirectory`）。确保网络中的其他 SMB 客户端也可以挂载此路径。

要复制指定子目录中的所有数据，DataSync必须能够装载 SMB 共享并访问其所有数据。有关更多信息，请参阅 SMB 的[必需权限](#)。

类型：字符串

最大长度限制：最大长度

模式：`^[a-zA-Z0-9_\-\+\.\^(\)\$\p{Zs}]+$`

必需：是

**Tags**

指定帮助您分类、筛选和搜索 AWS 资源的标签。我们建议至少为您的位置创建一个名称标签。

类型：[TagListEntry](#) 对象数组

数组成员：最少 0 项。最多 50 项。

必需：否

**User**

指定可以装载 SMB 文件服务器并有权访问传输中涉及的文件和文件夹的用户名。

有关为您的转移选择具有正确访问权限级别的用户的信息，请参阅 SMB 地点[所需的权限](#)。

类型：字符串

最大长度限制：最大长度

模式：`^[^\x5B\x5D\\\/:;|=,+*?]{1,104}$`

必需：是

## 响应语法

```
{  
  "LocationArn": "string"  
}
```

## 响应元素

如果此操作成功，则该服务将会发送回 HTTP 200 响应。

服务以 JSON 格式返回的以下数据。

### LocationArn

您创建的 SMB 的 ARN。

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\n-0-9]+:[0-9]{12}:location/loc-[0-9a-z]{17}$`

## 错误

有关所有操作常见错误的信息，请参阅[常见错误](#)。

### InternalException

当AWS DataSync服务中发生错误时，会引发此异常。

HTTP 状态代码：500

## InvalidRequestException

当客户端提交格式错误的请求时，会引发此异常。

HTTP 状态代码：400

### 示例

#### 示例请求

以下示例为 SMB 文件服务器创建一个位置。

```
{
  "AgentArns": [
    "arn:aws:datasync:us-east-2:111222333444:agent/agent-0b0addbeef44b3nfs",
    "arn:aws:datasync:us-east-2:111222333444:agent/agent-2345noo35nnee1123ovo3"
  ],
  "Domain": "AMAZON",
  "MountOptions": {
    "Version": "SMB3"
  },
  "Password": "string",
  "ServerHostname": "MyServer.amazon.com",
  "Subdirectory": "share",
  "Tags": [
    {
      "Key": "department",
      "Value": "finance"
    }
  ],
  "User": "user-1"
}
```

#### 示例响应

响应返回您的 SMB 文件服务器的位置 ARN。

```
{
  "LocationArn": "arn:aws:datasync:us-east-1:111222333444:location/loc-0f01451b140b2af49"
}
```

## 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [AWS 命令行界面](#)
- [适用于 .NET 的 AWS SDK](#)
- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [AWSSDK 用于 JavaScript](#)
- [适用于 PHP V3 的 AWS SDK](#)
- [适用于 Python 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)

# CreateTask

配置任务，定义在何处和如何AWS DataSync传输数据。

任务包括源位置、目标位置以及您希望如何和何时传输数据的首选项（例如带宽限制、日程安排和其他选项）。

## Important

如果您计划向 Amazon S3 地点传输数据或从 Amazon S3 站点传输数据，请在开始之前查看[DataSync如何影响您的 S3 请求费用](#)和[DataSync定价页面](#)。

## 请求语法

```
{
  "CloudWatchLogGroupArn": "string",
  "DestinationLocationArn": "string",
  "Excludes": [
    {
      "FilterType": "string",
      "Value": "string"
    }
  ],
  "Includes": [
    {
      "FilterType": "string",
      "Value": "string"
    }
  ],
  "Name": "string",
  "Options": {
    "Atime": "string",
    "BytesPerSecond": number,
    "Gid": "string",
    "LogLevel": "string",
    "Mtime": "string",
    "ObjectTags": "string",
    "OverwriteMode": "string",
    "PosixPermissions": "string",
    "PreserveDeletedFiles": "string",
    "PreserveDevices": "string",
```







必需：否

## Options

指定任务的配置选项。一些选项包括保留文件或对象元数据以及验证数据完整性。

您也可以在开始单独运行任务（也称为任务执行）之前覆盖这些选项。有关更多信息更多信息阅阅阅阅阅阅阅阅[StartTaskExecution](#)阅

类型：[Options](#) 对象

必需：否

## Schedule

指定用于定期将文件从源位置传输到目标位置的时间表。应以 UTC 时间来指定时间表。有关更多信息，请参阅[计划任务](#)。

类型：[TaskSchedule](#) 对象

必需：否

## SourceLocationArn

任务源位置的 Amazon 资源名称（ARN）。

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\\-0-9]+:[0-9]{12}:location/loc-[0-9a-z]{17}$`

必需：是

## Tags

指定要应用于代表任务的 Amazon 资源名称（ARN）的标签。

标签是键值对，帮助您管理、筛选和搜索DataSync资源。

类型：[TagListEntry](#) 对象数组

数组成员：最少 0 项。最多 50 项。

必需：否

## 响应语法

```
{  
  "TaskArn": "string"  
}
```

## 响应元素

如果此操作成功，则该服务将会发送回 HTTP 200 响应。

服务以 JSON 格式返回的以下数据。

### TaskArn

任务的 Amazon 资源名称 (ARN)。

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\\-0-9]*:[0-9]{12}:task/task-[0-9a-f]{17}$`

## 错误

有关所有操作常见错误的信息，请参阅[常见错误](#)。

### InternalException

当AWS DataSync服务中发生错误时，会引发此异常。

HTTP 状态代码：500

### InvalidRequestException

当客户端提交格式错误的请求时，会引发此异常。

HTTP 状态代码：400

## 示例

### 示例请求

以下示例使用源位置和目标位置创建任务：

```
{
  "Options": {
    "Atime": "BEST_EFFORT",
    "Gid": "NONE",
    "Mtime": "PRESERVE",
    "PosixPermissions": "PRESERVE",
    "PreserveDevices": "NONE",
    "PreserveDeletedFiles": "PRESERVE",
    "Uid": "NONE",
    "VerifyMode": "POINT_IN_TIME_CONSISTENT"
  },
  "Schedule": {
    "ScheduleExpression": "0 12 ? * SUN,WED *"
  },
  "CloudWatchLogGroupArn": "arn:aws:logs:us-east-2:111222333444:log-group",
  "DestinationLocationArn": "arn:aws:datsync:us-east-2:111222333444:location/loc-07db7abfc326c50fb",
  "Name": "MyTask",
  "SourceLocationArn": "arn:aws:datsync:us-east-2:111222333444:location/loc-0f01451b140b2af49",
  "Tags": [{
    "Key": "Name",
    "Value": "Task-1"
  }]
}
```

### 示例响应

以下响应将返回任务的 Amazon 资源名称 (ARN)。

```
{
  "TaskArn": "arn:aws:datsync:us-east-2:111222333444:task/task-08de6e6697796f026"
}
```

## 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [AWS 命令行界面](#)
- [适用于 .NET 的 AWS SDK](#)
- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [AWSSDK 用于 JavaScript](#)
- [适用于 PHP V3 的 AWS SDK](#)
- [适用于 Python 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)

## DeleteAgent

删除代理。要指定要删除的代理，请在请求中使用代理的 Amazon 资源名称 (ARN)。该操作会断开代理与您的关联AWS 账户。但是，它不会从本地环境中删除代理虚拟机 (VM)。

### 请求语法

```
{
  "AgentArn": "string"
}
```

### 请求参数

有关所有操作的通用参数的信息，请参阅[常用参数](#)。

请求接受采用 JSON 格式的以下数据。

#### AgentArn

要删除的代理的 Amazon 资源名称 (ARN)。使用 ListAgents 操作以返回账户和 AWS 区域的代理列表。

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\-\0-9]+:[0-9]{12}:agent/agent-[0-9a-z]{17}$`

必需：是

### 响应元素

如果此操作成功，则该服务会发送回带有空 HTTP 正文的 HTTP 200 响应。

### 错误

有关所有操作常见错误的信息，请参阅[常见错误](#)。

#### InternalException

当AWS DataSync服务中发生错误时，会抛出此异常。

HTTP 状态代码：500

InvalidRequestException

当客户端提交格式错误的请求时，会引发此异常。

HTTP 状态代码：400

## 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [AWS 命令行界面](#)
- [适用于 .NET 的 AWS SDK](#)
- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [AWS适用于JavaScript](#)
- [适用于 PHP V3 的 AWS SDK](#)
- [适用于 Python 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)

## DeleteLocation

删除使用的位置的配置AWS DataSync。

### 请求语法

```
{  
  "LocationArn": "string"  
}
```

### 请求参数

有关所有操作的通用参数的信息，请参阅[常用参数](#)。

请求接受采用 JSON 格式的以下数据。

#### LocationArn

要删除的 Amazon 资源名称 ( ARN ) 。

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\n-0-9]+:[0-9]{12}:location/loc-[0-9a-z]{17}$`

必需：是

### 响应元素

如果此操作成功，则该服务会发送回带有空 HTTP 正文的 HTTP 200 响应。

### 错误

有关所有操作常见错误的信息，请参阅[常见错误](#)。

#### InternalException

当AWS DataSync服务中发生错误时，会抛出此异常。

HTTP 状态代码：500

## InvalidRequestException

当客户端提交格式错误的请求时，会引发此异常。

HTTP 状态代码：400

### 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [AWS 命令行界面](#)
- [适用于 .NET 的 AWS SDK](#)
- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [AWS适用于JavaScript](#)
- [适用于 PHP V3 的 AWS SDK](#)
- [适用于 Python 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)



## DeleteTask

删除AWS DataSync任务。

### 请求语法

```
{
  "TaskArn": "string"
}
```

### 请求参数

有关所有操作的通用参数的信息，请参阅[通用参数](#)。

请求接受采用 JSON 格式的以下数据。

#### TaskArn

指定要使用任务的 Amazon Resource Name (ARN)。

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\-\0-9]*:[0-9]{12}:task/task-[0-9a-f]{17}$`

必需：是

### 响应元素

如果此操作成功，则该服务会发送回带有空 HTTP 正文的 HTTP 200 响应。

### 错误

有关所有操作常见错误的信息，请参阅[常见错误](#)。

#### InternalException

当AWS DataSync服务中发生错误时，会抛出此异常。

HTTP 状态代码：500

## InvalidRequestException

当客户端提交格式错误的请求时，会引发此异常。

HTTP 状态代码：400

## 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [AWS 命令行界面](#)
- [适用于 .NET 的 AWS SDK](#)
- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [AWS适用于 SDKJavaScript](#)
- [适用于 PHP V3 的 AWS SDK](#)
- [适用于 Python 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)

## DescribeAgent

返回有关AWS DataSync代理的元数据，例如其名称、端点类型和状态。

### 请求语法

```
{  
  "AgentArn": "string"  
}
```

### 请求参数

有关所有操作的通用参数的信息，请参阅[通用参数](#)。

请求接受采用 JSON 格式的以下数据。

#### AgentArn

指定DataSync代。

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\n-0-9]+:[0-9]{12}:agent/agent-[0-9a-z]{17}$`

必需：是

### 响应语法

```
{  
  "AgentArn": "string",  
  "CreationTime": number,  
  "EndpointType": "string",  
  "LastConnectionTime": number,  
  "Name": "string",  
  "PrivateLinkConfig": {  
    "PrivateLinkEndpoint": "string",  
    "SecurityGroupArns": [ "string" ],  
    "SubnetArns": [ "string" ],  
  },  
}
```

```
    "VpcEndpointId": "string"  
  },  
  "Status": "string"  
}
```

## 响应元素

如果此操作成功，则该服务将会发送回 HTTP 200 响应。

服务以 JSON 格式返回的以下数据。

### [AgentArn](#)

客服的 ARN。

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\\-0-9]+:[0-9]{12}:agent/agent-[0-9a-z]{17}$`

### [CreationTime](#)

代理被激活的时间（即在您的账户中创建）。

类型：时间戳

### [EndpointType](#)

代理连接到的端点的类型。如果端点是 VPC 端点，则无法通过公有互联网访问代理。

类型：字符串

有效值：PUBLIC | PRIVATE\_LINK | FIPS

### [LastConnectionTime](#)

代理最后一次连接的时间DataSync。

类型：时间戳

### [Name](#)

。

类型：字符串

长度限制：最小长度为 1。长度上限为 256。

模式：`^[a-zA-Z0-9\s+.=_:@/-]+$`

## [PrivateLinkConfig](#)

DataSync用于访问 VPC 终端节点的子网和安全组。

类型：[PrivateLinkConfig](#) 对象

## [Status](#)

。如果状态为 ONLINE，则代理已正确配置并可供使用。运行状态是代理的正常运行状态。如果状态为 OFFLINE，则代理的虚拟机已关闭或代理处于不健康状态。在解决导致不正常运行状态的问题后，代理将恢复为“ONLINE (联机)”状态。

类型：字符串

有效值：ONLINE | OFFLINE

## 错误

有关所有操作常见错误的信息，请参阅[常见错误](#)。

## InternalException

当AWS DataSync服务中发生错误时，会引发此异常。

HTTP 状态代码：500

## InvalidRequestException

当客户端提交格式错误的请求时，会引发此异常。

HTTP 状态代码：400

## 示例

### 示例

。

## 示例请求

```
{
  "AgentArn": "arn:aws:datsync:us-east-2:111222333444:agent/agent-0b0addbeef44baca3"
}
```

## 示例

此示例说明了的一种用法DescribeAgent。

## 示例响应

```
{
  "AgentArn": "arn:aws:datsync:us-east-2:111222333444:agent/agent-0b0addbeef44baca3",
  "CreationTime": "1532660733.39",
  "LastConnectionTime": "1532660733.39",
  "Name": "MyAgent",
  "Status": "ONLINE"
}
```

## 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [AWS 命令行界面](#)
- [适用于 .NET 的 AWS SDK](#)
- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [AWSSDK 用于 JavaScript](#)
- [适用于 PHP V3 的 AWS SDK](#)
- [适用于 Python 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)

## DescribeDiscoveryJob

返回有关DataSync发现作业的信息。

### 请求语法

```
{
  "DiscoveryJobArn": "string"
}
```

### 请求参数

有关所有操作的通用参数的信息，请参阅[通用参数](#)。

请求接受采用 JSON 格式的以下数据。

#### DiscoveryJobArn

指定要提供有关发现作业的 Amazon Resource Name ( ARN ) 。

类型：字符串

长度限制：最大长度为 256。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\n-0-9]+:[0-9]{12}:system/storage-system-[a-f0-9]{8}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{12}/job/discovery-job-[a-f0-9]{8}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{12}$`

必需：是

### 响应语法

```
{
  "CollectionDurationMinutes": number,
  "DiscoveryJobArn": "string",
  "JobEndTime": number,
  "JobStartTime": number,
  "Status": "string",
  "StorageSystemArn": "string"
}
```

```
}
```

## 响应元素

如果此操作成功，则该服务将会发送回 HTTP 200 响应。

服务以 JSON 格式返回的以下数据。

### CollectionDurationMinutes

发现作业运行的分钟数。

类型：整数

有效范围：最小值为 60。最大值为 44640。

### DiscoveryJobArn

发现作业的 ARN。

类型：字符串

长度限制：最大长度为 256。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\n-0-9]+:[0-9]{12}:system/storage-system-[a-f0-9]{8}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{12}/job/discovery-job-[a-f0-9]{8}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{12}$`

### JobEndTime

发现作业结束的时间。

类型：时间戳

### JobStartTime

发现作业的时间。

类型：时间戳

### Status

指示发现作业的状态。有关更多信息，请参阅“[发现作业的状态](#)”。

类型：字符串



有效值：RUNNING | WARNING | TERMINATED | FAILED | STOPPED | COMPLETED | COMPLETED\_WITH\_ISSUES

### StorageSystemArn

您正在运行发现任务的本地存储系统的 ARN。

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\-\0-9]+:[0-9]{12}:system/storage-system-[a-f0-9]{8}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{12}$`

### 错误

有关所有操作常见错误的信息，请参阅[常见错误](#)。

#### InternalException

当AWS DataSync服务中发生错误时，会引发此异常。

HTTP 状态代码：500

#### InvalidRequestException

当客户端提交格式错误的请求时，会引发此异常。

HTTP 状态代码：400

### 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [AWS 命令行界面](#)
- [适用于 .NET 的 AWS SDK](#)
- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)

- [AWSSDK 用于 JavaScript](#)
- [适用于 PHP V3 的 AWS SDK](#)
- [适用于 Python 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)

## DescribeLocationEfs

返回有关 Amazon EFS 文件系统AWS DataSync位置的 Matazon EFS 文件系统的

### 请求语法

```
{
  "LocationArn": "string"
}
```

### 请求参数

有关所有操作的通用参数的信息，请参阅[通用参数](#)。

请求接受采用 JSON 格式的以下数据。

#### LocationArn

您想要的 Amazon EFS 文件系统位置的 Amazon 资源名称 ( ARN ) 。

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\-\0-9]+:[0-9]{12}:location/loc-[0-9a-z]{17}$`

必需：是

### 响应语法

```
{
  "AccessPointArn": "string",
  "CreationTime": number,
  "Ec2Config": {
    "SecurityGroupArns": [ "string" ],
    "SubnetArn": "string"
  },
  "FileSystemAccessRoleArn": "string",
  "InTransitEncryption": "string",
  "LocationArn": "string",
```

```
"LocationUri": "string"  
}
```

## 响应元素

如果此操作成功，则该服务将会发送回 HTTP 200 响应。

服务以 JSON 格式返回的以下数据。

### [AccessPointArn](#)

DataSync用于访问 Amazon EFS 文件系统的访问点的 ARN。

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):elasticfilesystem:[a-z\-\0-9]+:[0-9]{12}:access-point/fsap-[0-9a-f]{8,40}$`

### [CreationTime](#)

位置的创建时间。

类型：时间戳

### [Ec2Config](#)

AWS DataSync 用于访问 Amazon EFS 文件系统的子网和安全组。

类型：[Ec2Config](#) 对象

### [FileSystemAccessRoleArn](#)

在挂载 Amazon EFS 文件系统时DataSync担任的AWS Identity and Access Management ( IAM ) 角色。

类型：字符串

长度限制：最大长度为 2048。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):iam::[0-9]{12}:role/.*$`

## InTransitEncryption

描述当 Amazon EFS 文件系统或从文件系统复制数据时，您是否DataSync使用传输层安全性协议 ( TLS ) 加密。

类型：字符串

有效值：NONE | TLS1\_2

## LocationArn

Amazon EFS 文件系统位置的 ARN。

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\-\0-9]+:[0-9]{12}:location/loc-[0-9a-z]{17}$`

## LocationUri

Amazon EFS 文件系统位置的 URL。

类型：字符串

长度限制：最大长度为 4360。

模式：`^(efs|nfs|s3|smb|hdfs|fsx[a-z0-9-]+)://[a-zA-Z0-9.:/\-]+$`

## 错误

有关所有操作常见错误的信息，请参阅[常见错误](#)。

### InternalException

当AWS DataSync服务中发生错误时，会引发此异常。

HTTP 状态代码：500

### InvalidRequestException

当客户端提交格式错误的请求时，会引发此异常。

HTTP 状态代码：400

## 示例

### 示例请求

以下示例显示如何获取有关特定 Amazon EFS 文件系统位置的信息。

```
{
  "LocationArn": "arn:aws:datasync:us-east-2:111222333444:location/
loc-12abcdef012345678"
}
```

### 示例响应

以下示例返回有关 Amazon EFS 文件系统的位置详细信息。

```
{
  "CreationTime": 1653319021.353,
  "Ec2Config": {
    "SubnetArn": "arn:aws:ec2:us-east-2:111222333444:subnet/
subnet-1234567890abcdef1",
    "SecurityGroupArns": [
      "arn:aws:ec2:us-east-2:111222333444:security-group/sg-1234567890abcdef2"
    ]
  },
  "LocationArn": "arn:aws:datasync:us-east-2:111222333444:location/loc-
abcdef01234567890",
  "LocationUri": "efs://us-east-2.fs-021345abcdef6789/"
}
```

示例响应：描述受限 Amazon EFS 文件系统的位置

以下示例返回有关具有受限访问权限的 Amazon EFS 文件系统的位置详细信息 `AccessPointArn`，包括 `FileSystemAccessRoleArn`、和 `InTransitEncryption` 元素。

```
{
  "CreationTime": 1653319021.353,
  "AccessPointArn": "arn:aws:elasticfilesystem:us-east-2:111222333444:access-point/
fsap-1234567890abcdef0",
  "Ec2Config": {
    "SubnetArn": "arn:aws:ec2:us-east-2:111222333444:subnet/
subnet-1234567890abcdef1",
    "SecurityGroupArns": [
```

```
        "arn:aws:ec2:us-east-2:111222333444:security-group/sg-1234567890abcdef2"
    ]
},
"FileSystemAccessRoleArn": "arn:aws:iam::111222333444:role/
AwsDataSyncFullAccessNew",
"InTransitEncryption": "TLS1_2",
"LocationArn": "arn:aws:datasync:us-east-2:111222333444:location/loc-
abcdef01234567890",
"LocationUri": "efs://us-east-2.fs-021345abcdef6789/",
"Subdirectory": "/mount/path",
"Tags": [{
    "Key": "Name",
    "Value": "ElasticFileSystem-1"
}]
}
```

## 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [AWS 命令行界面](#)
- [适用于 .NET 的 AWS SDK](#)
- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [AWSSDK 用于 JavaScript](#)
- [适用于 PHP V3 的 AWS SDK](#)
- [适用于 Python 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)

## DescribeLocationFsxLustre

提供有关如何配置 Amazon FSx for Lustre 文件系统AWS DataSync位置的详细信息。

### 请求语法

```
{  
  "LocationArn": "string"  
}
```

### 请求参数

有关所有操作的通用参数的信息，请参阅[通用参数](#)。

请求接受采用 JSON 格式的以下数据。

#### LocationArn

要描述的 FSx for Lustre 位置的 Amazon 资源名称 (ARN)。

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\n-0-9]+:[0-9]{12}:location/loc-[0-9a-z]{17}$`

必需：是

### 响应语法

```
{  
  "CreationTime": number,  
  "LocationArn": "string",  
  "LocationUri": "string",  
  "SecurityGroupArns": [ "string" ]  
}
```

### 响应元素

如果此操作成功，则该服务将会发送回 HTTP 200 响应。



服务以 JSON 格式返回的以下数据。

### CreationTime

创建 FsX for Lustre 位置的时间。

类型：时间戳

### LocationArn

所描述的 FSx for Lustre 位置的 Amazon 资源名称 (ARN)。

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\-\0-9]+:[0-9]{12}:location/loc-[0-9a-z]{17}$`

### LocationUri

所描述的 FSx for Stre 位置的 URI。

类型：字符串

长度约束：最大长度为 4360。

模式：`^(efs|nfs|s3|smb|hdfs|fsx[a-z0-9-]+)://[a-zA-Z0-9.\:/\-\-]+$`

### SecurityGroupArns

为 FSx for Lustre 文件系统配置的安全组的 Amazon 资源名称 (ARN)。

类型：字符串数组

数组成员：最少 1 项。最多 5 项。

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):ec2:[a-z\-\0-9]*:[0-9]{12}:security-group/sg-[a-f0-9]+$`

## 错误

有关所有操作常见错误的信息，请参阅[常见错误](#)。

## InternalException

当AWS DataSync服务中发生错误时，会抛出此异常。

HTTP 状态代码：500

## InvalidRequestException

当客户端提交格式错误的请求时，会引发此异常。

HTTP 状态代码：400

## 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [AWS 命令行界面](#)
- [适用于 .NET 的 AWS SDK](#)
- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [AWS适用于 SDKJavaScript](#)
- [适用于 PHP V3 的 AWS SDK](#)
- [适用于 Python 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)

## DescribeLocationFsxOntap

提供有关如何配置 Amazon FSx fNetApp or ONTAP 文件系统的AWS DataSync位置的详细信息。

### Note

如果您的位置使用 SMB，则该DescribeLocationFsxOntap操作实际上不会返回 aPassword。

### 请求语法

```
{  
  "LocationArn": "string"  
}
```

### 请求参数

有关所有操作的通用参数的信息，请参阅[常用参数](#)。

请求接受采用 JSON 格式的以下数据。

#### LocationArn

指定要访问 FSx for ONTAP 文件系统位置的 Amazon 资源名称 (ARN)。

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\\-0-9]+:[0-9]{12}:location/loc-[0-9a-z]{17}$`

必需：是

### 响应语法

```
{  
  "CreationTime": number,
```

```

    "FsxFilesystemArn": "string",
    "LocationArn": "string",
    "LocationUri": "string",
    "Protocol": {
      "NFS": {
        "MountOptions": {
          "Version": "string"
        }
      },
      "SMB": {
        "Domain": "string",
        "MountOptions": {
          "Version": "string"
        },
        "Password": "string",
        "User": "string"
      }
    },
    "SecurityGroupArns": [ "string" ],
    "StorageVirtualMachineArn": "string"
  }

```

## 响应元素

如果此操作成功，则该服务将会发送回 HTTP 200 响应。

服务以 JSON 格式返回的以下数据。

### CreationTime

位置的创建时间。

类型：时间戳

### FsxFilesystemArn

FSx for ONTAP 文件系统的 ARN。

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):fsx:[a-z\-\0-9]*:[0-9]{12}:file-system/fs-.*$`

## [LocationArn](#)

FSx for ONTAP 文件系统位置的 ARN。

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\\-0-9]+:[0-9]{12}:location/loc-[0-9a-z]{17}$`

## [LocationUri](#)

FSx for ONTAP 文件系统位置的统一资源标识符 (URI)。

类型：字符串

长度约束：最大长度为 4360。

模式：`^(efs|nfs|s3|smb|hdfs|fsx[a-z0-9-]+)://[a-zA-Z0-9.:/\\-]+$`

## [Protocol](#)

指定 AWS DataSync 用于访问您的 Amazon FSx 文件系统的数据传输协议。

类型：[FsxProtocol](#) 对象

## [SecurityGroupArns](#)

用于访问 FSx for ONTAP 文件系统的安全组。DataSync

类型：字符串数组

数组成员：最少 1 项。最多 5 项。

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):ec2:[a-z\\-0-9]*:[0-9]{12}:security-group/sg-[a-f0-9]+$`

## [StorageVirtualMachineArn](#)

要将数据复制到或从中复制数据的 FSx for ONTAP 文件系统上存储虚拟机 (SVM) 的 ARN。

类型：字符串

长度约束：最大长度为 162。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):fsx:[a-z\-\0-9]+:[0-9]{12}:storage-virtual-machine/fs-[0-9a-f]+/svm-[0-9a-f]{17,}$`

## 错误

有关所有操作常见错误的信息，请参阅[常见错误](#)。

### InternalException

当AWS DataSync服务中发生错误时，会抛出此异常。

HTTP 状态代码：500

### InvalidRequestException

当客户端提交格式错误的请求时，会引发此异常。

HTTP 状态代码：400

## 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [AWS 命令行界面](#)
- [适用于 .NET 的 AWS SDK](#)
- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [AWS适用于JavaScript](#)
- [适用于 PHP V3 的 AWS SDK](#)
- [适用于 Python 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)

## DescribeLocationFsxOpenZfs

提供有关如何配置 Amazon FSx for OpenZFS 文件系统的详细AWS DataSync位置。

### Note

该DescribeLocationFsxOpenZfs操作SMB不支持与相关的响应元素。

### 请求语法

```
{
  "LocationArn": "string"
}
```

### 请求参数

有关所有操作的通用参数的信息，请参阅[常用参数](#)。

请求接受采用 JSON 格式的以下数据。

#### LocationArn

要描述FSx for OpenZFS 位置的 Amazon 资源名称 ( ARN ) 。

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\-0-9]+:[0-9]{12}:location/loc-[0-9a-z]{17}$`

必需：是

### 响应语法

```
{
  "CreationTime": number,
  "LocationArn": "string",
  "LocationUri": "string",
  "Protocol": {
```

```

    "NFS": {
      "MountOptions": {
        "Version": "string"
      }
    },
    "SMB": {
      "Domain": "string",
      "MountOptions": {
        "Version": "string"
      },
      "Password": "string",
      "User": "string"
    }
  },
  "SecurityGroupArns": [ "string" ]
}

```

## 响应元素

如果此操作成功，则该服务将会发送回 HTTP 200 响应。

服务以 JSON 格式返回的以下数据。

### [CreationTime](#)

创建适用于 OpenZFS 的 FSx 位置的时间。

类型：时间戳

### [LocationArn](#)

所FSx ARN。

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\-\0-9]+:[0-9]{12}:location/loc-[0-9a-z]{17}$`

### [LocationUri](#)

所描述的 FSx for OpenZFS 位置的统一资源标度为 FSx。

示例：`fsxz://us-west-2.fs-1234567890abcdef02/fsx/folderA/folder`



类型：字符串

长度为 43。

模式：`^(efs|nfs|s3|smb|hdfs|fsx[a-z0-9-]+)://[a-zA-Z0-9.:/\-]+$`

## [Protocol](#)

AWS DataSync 用于访问文件系统的协议类型。

类型：[FsxProtocol](#) 对象

## [SecurityGroupArns](#)

配置 FSx for OpenZFS 文件系统的安全组的 ARN。

类型：字符串数组

数组成员：最少 1 项。最多 5 项。

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):ec2:[a-z\-\0-9]*:[0-9]{12}:security-group/sg-[a-f0-9]+$`

## 错误

有关所有操作常见错误的信息，请参阅[常见错误](#)。

### InternalException

当AWS DataSync服务中发生错误时，会抛出此异常。

HTTP 状态代码：500

### InvalidRequestException

当客户端提交格式错误的请求时，会引发此异常。

HTTP 状态代码：400

## 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [AWS 命令行界面](#)
- [适用于 .NET 的 AWS SDK](#)
- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [AWS适用于JavaScript](#)
- [适用于 PHP V3 的 AWS SDK](#)
- [适用于 Python 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)

## DescribeLocationFsxWindows

返回Amazon FSx for Windows File Server 位置的元数据，例如有关其路径的信息。

### 请求语法

```
{  
  "LocationArn": "string"  
}
```

### 请求参数

有关所有操作的通用参数的信息，请参阅[常用参数](#)。

请求接受采用 JSON 格式的以下数据。

#### LocationArn

要描述的 FSx for Windows File Server 位置的 Amazon 资源名称 ( ARN )。

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\n-0-9]+:[0-9]{12}:location/loc-[0-9a-z]{17}$`

必需：是

### 响应语法

```
{  
  "CreationTime": number,  
  "Domain": "string",  
  "LocationArn": "string",  
  "LocationUri": "string",  
  "SecurityGroupArns": [ "string" ],  
  "User": "string"  
}
```

## 响应元素

如果此操作成功，则该服务将会发送回 HTTP 200 响应。

服务以 JSON 格式返回的以下数据。

### CreationTime

FSx for Windows File Server 位置的创建时间。

类型：时间戳

### Domain

FSx for Windows File Server 所属的 Windows 域的名称。

类型：字符串

长度限制：最大长度为 253。

模式：`^[A-Za-z0-9](\\.|-+)?[A-Za-z0-9]{0,252}$`

### LocationArn

所描述的 FSx for Windows File Server 位置的 Amazon 资源名称 ( ARN )。

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\\-0-9]+:[0-9]{12}:location/loc-[0-9a-z]{17}$`

### LocationUri

描述的 FSx for Windows File Server 位置的 URL。

类型：字符串

长度限制：最大长度为 4360。

模式：`^(efs|nfs|s3|smb|hdfs|fsx[a-z0-9-]+)://[a-zA-Z0-9.:/\\-]+$`

### SecurityGroupArns

为 FSx for Windows File Server 文件系统的安全组的 Amazon 资源名称 ( ARN )。

类型：字符串数组

数组成员：最少 1 项。最多 5 项。

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):ec2:[a-z\-\0-9]*:[0-9]{12}:security-group/sg-[a-f0-9]+$`

## User

有权访问 FSx for Windows File Server 文件系统中的文件和文件夹的用户。

类型：字符串

长度限制：最大长度为 104。

模式：`^[^\x5B\x5D\\\/:;|=,+\*?]{1,104}$`

## 错误

有关所有操作常见错误的信息，请参阅[常见错误](#)。

### InternalException

当AWS DataSync服务中发生错误时，会抛出此异常。

HTTP 状态代码：500

### InvalidRequestException

当客户端提交格式错误的请求时，会引发此异常。

HTTP 状态代码：400

## 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [AWS 命令行界面](#)
- [适用于 .NET 的 AWS SDK](#)
- [适用于 C++ 的 AWS SDK](#)

- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [AWS适用于JavaScript](#)
- [适用于 PHP V3 的 AWS SDK](#)
- [适用于 Python 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)

## DescribeLocationHdfs

返回元数据，例如有关 Hadoop Distributed File System ( HDFS ) 信息。

### 请求语法

```
{
  "LocationArn": "string"
}
```

### 请求参数

有关所有操作的通用参数的信息，请参阅[通用参数](#)。

请求接受采用 JSON 格式的以下数据。

#### LocationArn

要描述的 HDFS 集群位置的 Amazon 资源名称 ( ARN )。

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\\-0-9]+:[0-9]{12}:location/loc-[0-9a-z]{17}$`

必需：是

### 响应语法

```
{
  "AgentArns": [ "string" ],
  "AuthenticationType": "string",
  "BlockSize": number,
  "CreationTime": number,
  "KerberosPrincipal": "string",
  "KmsKeyProviderUri": "string",
  "LocationArn": "string",
  "LocationUri": "string",
  "NameNodes": [
```

```
{
  {
    "Hostname": "string",
    "Port": number
  }
],
"QopConfiguration": {
  "DataTransferProtection": "string",
  "RpcProtection": "string"
},
"ReplicationFactor": number,
"SimpleUser": "string"
}
```

## 响应元素

如果此操作成功，则该服务将会发送回 HTTP 200 响应。

服务以 JSON 格式返回的以下数据。

### [AgentArns](#)

用于连接到 HDFS 集群的 ARN。

类型：字符串数组

数组成员：最少 1 项。最多

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\\-0-9]+:[0-9]{12}:agent/agent-[0-9a-z]{17}$`

### [AuthenticationType](#)

用于确定用户身份的身份验证类型。

类型：字符串

有效值：SIMPLE | KERBEROS

### [BlockSize](#)

要写入 HDFS 集群的大小大小。

类型：整数



有效范围：最多 最大值为

### [CreationTime](#)

HDFS 创建时间。

类型：时间戳

### [KerberosPrincipal](#)

有权访问 HDFS 集群上的文件和文件夹的 Kerberos 主体。如果定义为 `AuthenticationType` 则使用此参数 `KERBEROS`。

类型：字符串

长度限制：最小长度为 1。长度上限为 256。

模式：`^.+`

### [KmsKeyProviderUri](#)

HDFS 集群的密钥管理服务器 ( KMS ) 的 URI。

类型：字符串

长度限制：最小长度为 1。最大长度为 255。

模式：`^kms:\//http[s]?@((([a-zA-Z0-9\-\ ]*[a-zA-Z0-9])\.)*([A-Za-z0-9\-\ ]*[A-Za-z0-9])(;((([a-zA-Z0-9\-\ ]*[a-zA-Z0-9])\.)*([A-Za-z0-9\-\ ]*[A-Za-z0-9])))*:[0-9]{1,5}\//kms$`

### [LocationArn](#)

HDFS 集群的 ARN。

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-zA-Z0-9]+:[0-9]{12}:location/loc-[0-9a-z]{17}$`

### [LocationUri](#)

HDFS 集群位置的 URI。

类型：字符串

长度限制：最大长度为 430 0 0 0 0 0 0 0 0

模式：`^(efs|nfs|s3|smb|hdfs|fsx[a-z0-9-]+)://[a-zA-Z0-9.:/\-]+$`

## NameNodes

管理 H NameNode DFS 管理 HDFS。

类型：[HdfsNameNode](#) 对象数组

数组成员：最少 1 项。

## QopConfiguration

保护质量 ( QOP ) 配置指定在 Hadoop Distributed File System ( HDFS ) 集群上配置的远程程序调用 ( RPC ) 和数据传输保护设置。

类型：[QopConfiguration](#) 对象

## ReplicationFactor

写入 HDFS 集群时要将数据复制到的 HDFS 数量。DataNodes

类型：整数

有效范围：最小值为 1。最大值为 Futed

## SimpleUser

用于在主机操作系统上标识客户端的用户名。如果定义为 `AuthenticationType` 则使用此参数 `SIMPLE`。

类型：字符串

长度限制：最小长度为 1。长度上限为 256。

模式：`^[_.A-Za-z0-9][-_.A-Za-z0-9]*$`

## 错误

有关所有操作常见错误的信息，请参阅[常见错误](#)。

## InternalException

当 AWS DataSync 服务中发生错误时，会引发此异常。

HTTP 状态代码：500

InvalidRequestException

当客户端提交格式错误的请求时，会引发此异常。

HTTP 状态代码：400

## 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [AWS 命令行界面](#)
- [适用于 .NET 的 AWS SDK](#)
- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [AWSSDK 用于 JavaScript](#)
- [适用于 PHP V3 的 AWS SDK](#)
- [适用于 Python 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)

## DescribeLocationNfs

返回有关 NFS 位置的元数据，例如路径信息。

### 请求语法

```
{  
  "LocationArn": "string"  
}
```

### 请求参数

有关所有操作的通用参数的信息，请参阅[通用参数](#)。

请求接受采用 JSON 格式的以下数据。

#### LocationArn

要描述的 NFS 位置的 Amazon 资源名称 (ARN)。

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\-\0-9]+:[0-9]{12}:location/loc-[0-9a-z]{17}$`

必需：是

### 响应语法

```
{  
  "CreationTime": number,  
  "LocationArn": "string",  
  "LocationUri": "string",  
  "MountOptions": {  
    "Version": "string"  
  },  
  "OnPremConfig": {  
    "AgentArns": [ "string" ]  
  }  
}
```

```
}
```

## 响应元素

如果此操作成功，则该服务将会发送回 HTTP 200 响应。

服务以 JSON 格式返回的以下数据。

### CreationTime

创建 NFS 位置的时间。

类型：时间戳

### LocationArn

所描述的 NFS 位置的 Amazon 资源名称 (ARN)。

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\\-0-9]+:[0-9]{12}:location/loc-[0-9a-z]{17}$`

### LocationUri

指定的源 NFS 位置的 URL。

类型：字符串

长度约束：最大长度为 4360。

模式：`^(efs|nfs|s3|smb|hdfs|fsx[a-z0-9-]+)://[a-zA-Z0-9.:/\-]+$`

### MountOptions

DataSync用于挂载 NFS 共享的 NFS 挂载选项。

类型：[NfsMountOptions](#) 对象

### OnPremConfig

用于网络文件系统 (NFS) 位置的代理的 Amazon 资源名称 (ARN) 列表。

类型：[OnPremConfig](#) 对象

## 错误

有关所有操作常见错误的信息，请参阅[常见错误](#)。

### InternalException

当AWS DataSync服务中发生错误时，会抛出此异常。

HTTP 状态代码：500

### InvalidRequestException

当客户端提交格式错误的请求时，会引发此异常。

HTTP 状态代码：400

## 示例

### 示例

以下示例返回有关示例请求中指定的NFS位置的信息。

### 示例请求

```
{
  "LocationArn": "arn:aws:datsync:us-east-2:111222333444:location/
loc-07db7abfc326c50aa"
}
```

### 示例

此示例说明了的一种用法DescribeLocationNfs。

### 示例响应

```
{
  "CreationTime": 1532660733.39,
  "LocationArn": "arn:aws:datsync:us-east-2:111222333444:location/
loc-07db7abfc326c50aa",
  "LocationUri": "hostname.amazon.com",
  "OnPremConfig": {
    "AgentArns": [ "arn:aws:datsync:us-east-2:111222333444:agent/
agent-0b0addbeef44b3nfs" ]
  }
}
```

```
}  
}
```

## 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [AWS 命令行界面](#)
- [适用于 .NET 的 AWS SDK](#)
- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [AWS适用于 SDKJavaScript](#)
- [适用于 PHP V3 的 AWS SDK](#)
- [适用于 Python 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)

## DescribeLocationObjectStorage

返回有关您的对象存储系统AWS DataSync位置的元数据。

### 请求语法

```
{
  "LocationArn": "string"
}
```

### 请求参数

有关所有操作的通用参数的信息，请参阅[通用参数](#)。

请求接受采用 JSON 格式的以下数据。

#### LocationArn

要获取信息的对象存储系统位置的 Amazon 资源名称 (ARN)。

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\\-0-9]+:[0-9]{12}:location/loc-[0-9a-z]{17}$`

必需：是

### 响应语法

```
{
  "AccessKey": "string",
  "AgentArns": [ "string" ],
  "CreationTime": number,
  "LocationArn": "string",
  "LocationUri": "string",
  "ServerCertificate": blob,
  "ServerPort": number,
  "ServerProtocol": "string"
}
```



## 响应元素

如果此操作成功，则该服务将会发送回 HTTP 200 响应。

服务以 JSON 格式返回的以下数据。

### AccessKey

向对象存储系统进行身份验证所需的访问密钥（例如，用户名）。

类型：字符串

长度限制：最小长度为 1。最大长度为 200。

模式：`^.+`

### AgentArns

可以安全连接到您所在位置的 ARN。DataSync

类型：字符串数组

数组成员：最少 1 项。最多 4 项。

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\\-0-9]+:[0-9]{12}:agent/agent-[0-9a-z]{17}$`

### CreationTime

位置的 Amazon 资源创建时间。

类型：时间戳

### LocationArn

对象存储系统位置的 ARN。

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\\-0-9]+:[0-9]{12}:location/loc-[0-9a-z]{17}$`

## LocationUri

对象存储系统位置的 URL。

类型：字符串

长度约束：最大长度约束 4360。

模式：`^(efs|nfs|s3|smb|hdfs|fsx[a-z0-9-]+)://[a-zA-Z0-9.:/\-]+$`

## ServerCertificate

DataSync用于在对象存储系统上进行安全身份验证的自签名证书。

类型：Base64 编码的二进制数据对象

长度约束：最大长度约束 32768。

## ServerPort

对象存储服务器接受入站网络流量的端口（例如，端443）。

类型：整数

有效范围：最小值为 1。最大值为 65536。

## ServerProtocol

对象存储系统通信时使用的协议。

类型：字符串

有效值：HTTPS | HTTP

## 错误

有关所有操作常见错误的信息，请参阅[常见错误](#)。

### InternalException

当AWS DataSync服务中发生错误时，会引发此异常。

HTTP 状态代码：500

### InvalidRequestException

当客户端提交格式错误的请求时，会引发此异常。

HTTP 状态代码：400

## 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [AWS 命令行界面](#)
- [适用于 .NET 的 AWS SDK](#)
- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [AWSSDK 用于 JavaScript](#)
- [适用于 PHP V3 的 AWS SDK](#)
- [适用于 Python 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)

## DescribeLocationS3

返回有关 Amazon S3 存储桶位置的元数据，例如存储桶名称。

### 请求语法

```
{
  "LocationArn": "string"
}
```

### 请求参数

有关所有操作的通用参数的信息，请参阅[通用参数](#)。

请求接受采用 JSON 格式的以下数据。

#### LocationArn

要描述的 Amazon S3 存储桶位置的 Amazon 资源名称 (ARN)。

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\\-0-9]+:[0-9]{12}:location/loc-[0-9a-z]{17}$`

必需：是

### 响应语法

```
{
  "AgentArns": [ "string" ],
  "CreationTime": number,
  "LocationArn": "string",
  "LocationUri": "string",
  "S3Config": {
    "BucketAccessRoleArn": "string"
  },
  "S3StorageClass": "string"
}
```

## 响应元素

如果此操作成功，则该服务将会发送回 HTTP 200 响应。

服务以 JSON 格式返回的以下数据。

### AgentArns

如果您在AWS前哨基地DataSync上使用，则是在您的前哨基地上部署的 EC2 代理的亚马逊资源名称 (ARN)。有关在 AWS Outposts DataSync 上启动代理的更多信息，请参阅[在上AWS Outposts部署DataSync代理](#)。

类型：字符串数组

数组成员：最少 1 项。最多 4 项。

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\\-0-9]+:[0-9]{12}:agent/agent-[0-9a-z]{17}$`

### CreationTime

Amazon S3 存储桶位置的时间。

类型：时间戳

### LocationArn

Amazon S3 桶或接入点的 Amazon 资源名称 (ARN)。

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\\-0-9]+:[0-9]{12}:location/loc-[0-9a-z]{17}$`

### LocationUri

所描述的 Amazon S3 位置的 URL。

类型：字符串

长度限制：最大长度为 4360。

模式：`^(efs|nfs|s3|smb|hdfs|fsx[a-z0-9-]+)://[a-zA-Z0-9.:/\-]+$`

## [S3Config](#)

AWS Identity and Access Management ( IAM ) 角色用于访问 Amazon S3 桶的 Amazon 资源名称 ( ARN ) 。

有关使用此类角色的详细信息，请参阅《AWS DataSync 用户指南》中的[为 Amazon S3 创建位置](#)。

类型：[S3Config](#) 对象

## [S3StorageClass](#)

在将此位置用作任务目标时，您希望在其中存储文件的 Amazon S3 存储类。有关 S3 存储类的更多信息，请参阅[Amazon S3 存储类](#)。有些存储类别的行为会影响您的 S3 存储成本。有关详细信息，请参阅[中使用 S3 存储类时的注意事项DataSync](#)。

类型：字符串

有效值：STANDARD | STANDARD\_IA | ONEZONE\_IA | INTELLIGENT\_TIERING | GLACIER | DEEP\_ARCHIVE | OUTPOSTS | GLACIER\_INSTANT\_RETRIEVAL

## 错误

有关所有操作常见错误的信息，请参阅[常见错误](#)。

### InternalException

当AWS DataSync服务中发生错误时，会引发此异常。

HTTP 状态代码：500

### InvalidRequestException

当客户端提交格式错误的请求时，会引发此异常。

HTTP 状态代码：400

## 示例

### 示例

以下示例返回有关示例请求中指定的 S3 位置的信息。

## 示例请求

```
{
  "LocationArn": "arn:aws:datsync:us-east-2:111222333444:location/
loc-07db7abfc326c50s3"
}
```

## 示例

此示例说明了 DescribeLocation S3 的一种用法。

## 示例响应

```
{
  "CreationTime": 1532660733.39,
  "LocationArn": "arn:aws:datsync:us-east-2:111222333444:location/
loc-07db7abfc326c50s3",
  "LocationUri": "MyBucket.",
  "S3Config": {
    "BucketAccessRoleArn": "arn:aws:iam::111222333444:role/MyBucketAccessRole",
  }
  "S3StorageClass": "STANDARD"
}
```

## 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [AWS 命令行界面](#)
- [适用于 .NET 的 AWS SDK](#)
- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [AWS用于 SDK JavaScript](#)
- [适用于 PHP V3 的 AWS SDK](#)
- [适用于 Python 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)

## DescribeLocationSmb

返回有关 SMB 位置的元数据，例如路径和用户信息。

### 请求语法

```
{  
  "LocationArn": "string"  
}
```

### 请求参数

有关所有操作的通用参数的信息，请参阅[通用参数](#)。

请求接受采用 JSON 格式的以下数据。

#### LocationArn

要描述的 SMB 位置的 Amazon 资源名称 (ARN)。

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\n-0-9]+:[0-9]{12}:location/loc-[0-9a-z]{17}$`

必需：是

### 响应语法

```
{  
  "AgentArns": [ "string" ],  
  "CreationTime": number,  
  "Domain": "string",  
  "LocationArn": "string",  
  "LocationUri": "string",  
  "MountOptions": {  
    "Version": "string"  
  },  
  "User": "string"  
}
```



```
}
```

## 响应元素

如果此操作成功，则该服务将会发送回 HTTP 200 响应。

服务以 JSON 格式返回的以下数据。

### AgentArns

创建源 SMB 文件系统位置的 Amazon 资源名称 (ARN)。

类型：字符串数组

数组成员：最少 1 项。。

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\\-0-9]+:[0-9]{12}:agent/agent-[0-9a-z]{17}$`

### CreationTime

创建 SMB 位置的时间。

类型：时间戳

### Domain

SMB 服务器所属的 Windows 域的名称。

类型：字符串

长度约束：长度约束：。

模式：`^[A-Za-z0-9](\\.|-+)?[A-Za-z0-9]{0,252}$`

### LocationArn

描述的 SMB 位置的 Amazon 资源名称 (ARN)。

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\-\0-9]+:[0-9]{12}:location/loc-[0-9a-z]{17}$`

### LocationUri

指定的源 SMB 位置的 URL。

类型：字符串

长度约束：长度约束：。

模式：`^(efs|nfs|s3|smb|hdfs|fsx[a-z0-9-]+)://[a-zA-Z0-9./\-\-]+$`

### MountOptions

可供DataSync访问 SMB 位置的挂载选项。

类型：[SmbMountOptions](#) 对象

### User

可以挂载共享并对 SMB 共享中的文件和文件夹具有访问权限的用户。

类型：字符串

长度约束：长度约束：。

模式：`^[^\x5B\x5D\\\/:;|=,+\*?]{1,104}$`

## 错误

有关所有操作常见错误的信息，请参阅[常见错误](#)。

### InternalException

当AWS DataSync服务中发生错误时，会抛出此异常。

HTTP 状态代码：500

### InvalidRequestException

当客户端提交格式错误的请求时，会引发此异常。

HTTP 状态代码：400

## 示例

### 示例

此示例说明了的一种用法DescribeLocationSmb。

### 示例请求

```
{
  "arn:aws:datsync:us-east-1:111222333444:location/loc-0f01451b140b2af49"
}
```

### 示例

此示例说明了的一种用法DescribeLocationSmb。

### 示例响应

```
{
  "AgentArns": [
    "arn:aws:datsync:us-east-2:111222333444:agent/agent-0bc3b3dc9bbc15145",
    "arn:aws:datsync:us-east-2:111222333444:agent/agent-04b3fe3d261a18c8f"
  ],
  "CreationTime": "1532660733.39",
  "Domain": "AMAZON",
  "LocationArn": "arn:aws:datsync:us-east-1:111222333444:location/loc-0f01451b140b2af49",
  "LocationUri": "smb://hostname.amazon.com/share",
  "MountOptions": {
    "Version": "SMB3"
  },
  "User": "user-1"
}
```

## 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [AWS 命令行界面](#)
- [适用于 .NET 的 AWS SDK](#)
- [适用于 C++ 的 AWS SDK](#)

- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [AWS适用于 SDKJavaScript](#)
- [适用于 PHP V3 的 AWS SDK](#)
- [适用于 Python 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)

## DescribeStorageSystem

返回有关您在 DataSync Discovery 中使用的本地存储系统的信息。

### 请求语法

```
{
  "StorageSystemArn": "string"
}
```

### 请求参数

有关所有操作的通用参数的信息，请参阅[通用参数](#)。

请求接受采用 JSON 格式的以下数据。

#### StorageSystemArn

指定某个现有 DataSync Discovery 使用的本地存储系统的访问系统的 Amazon 资源名称 (ARN)。

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\\-0-9]+:[0-9]{12}:system/storage-system-[a-f0-9]{8}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{12}$`

必需：是

### 响应语法

```
{
  "AgentArns": [ "string" ],
  "CloudWatchLogGroupArn": "string",
  "ConnectivityStatus": "string",
  "CreationTime": number,
  "ErrorMessage": "string",
  "Name": "string",
  "SecretsManagerArn": "string",
```

```

  "ServerConfiguration": {
    "ServerHostname": "string",
    "ServerPort": number
  },
  "StorageSystemArn": "string",
  "SystemType": "string"
}

```

## 响应元素

如果此操作成功，则该服务将会发送回 HTTP 200 响应。

服务以 JSON 格式返回的以下数据。

### [AgentArns](#)

连接到您的本地存储系统并从中读取的DataSync代理的 ARN。

类型：字符串数组

数组成员：固定数量为 1 项。

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\-0-9]+:[0-9]{12}:agent/agent-[0-9a-z]{17}$`

### [CloudWatchLogGroupArn](#)

用于监控和记录发现任务事件的亚马逊CloudWatch日志组的 ARN。

类型：字符串

长度限制：最大长度为 562。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):logs:[a-z\-0-9]+:[0-9]{12}:log-group:([^\*]*)?(:\*)?$`

### [ConnectivityStatus](#)

表示您的DataSync代理是否可以连接到您的本地存储系统。

类型：字符串

有效值：PASS | FAIL | UNKNOWN

## CreationTime

将本地存储系统添加到 DataSync Discovery 的时间。

类型：时间戳

## ErrorMessage

描述DataSync代理在您的本地存储系统上遇到的连接错误。

类型：字符串

长度约束：最大长度为 128。

模式：.\*

## Name

您在将本地存储系统添加到 DataSync Discovery 时为其指定的名称。

类型：字符串

长度限制：最小长度为 1。长度上限为 256。

模式：`^[\\p{L}\\p{M}\\p{N}\\s+=._:@\\/-]+`

## SecretsManagerArn

存储本地存储系统凭证的密钥的 ARN。DataSyncDiscovery 将这些凭证存储在[AWS Secrets Manager](#)。

类型：字符串

长度限制：最大长度为 2048。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):secretsmanager:[a-z\\-0-9]+:[0-9]{12}:secret:.*`

## ServerConfiguration

连接到本地存储系统的管理界面所需的服务器名称和网络端口。

类型：[DiscoveryServerConfiguration](#) 对象

## StorageSystemArn

发现任务所查看的本地存储系统的 ARN。

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\\-0-9]+:[0-9]{12}:system/storage-system-[a-f0-9]{8}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{12}$`

## SystemType

本地存储系统的类型。

### Note

DataSyncDiscovery 目前仅支持运行 ON NetApp TAP 9.7 或更高版本的结构连接存储 (FAS) 和全闪存 FAS (AFF) 系统。

类型：字符串

有效值：NetAppONTAP

## 错误

有关所有操作常见错误的信息，请参阅[常见错误](#)。

### InternalException

当AWS DataSync服务中发生错误时，会引发此异常。

HTTP 状态代码：500

### InvalidRequestException

当客户端提交格式错误的请求时，会引发此异常。

HTTP 状态代码：400

## 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：



- [AWS 命令行界面](#)
- [适用于 .NET 的 AWS SDK](#)
- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [AWSSDK 用于 JavaScript](#)
- [适用于 PHP V3 的 AWS SDK](#)
- [适用于 Python 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)

## DescribeStorageSystemResourceMetrics

返回 DataSync Discovery 收集的有关本地存储系统中特定资源的信息，包括性能数据和容量使用情况。

### 请求语法

```
{
  "DiscoveryJobArn": "string",
  "EndTime": number,
  "MaxResults": number,
  "NextToken": "string",
  "ResourceId": "string",
  "ResourceType": "string",
  "StartTime": number
}
```

### 请求参数

有关所有操作的通用参数的信息，请参阅[通用参数](#)。

请求接受采用 JSON 格式的以下数据。

#### [DiscoveryJobArn](#)

指定您在其中计划的发现作业的 Amazon Resource Name ( ARN ) 。

类型：字符串

长度限制：最大长度为 256。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\n-0-9]+:[0-9]{12}:system/storage-system-[a-f0-9]{8}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{12}/job/discovery-job-[a-f0-9]{8}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{12}$`

必需：是

#### [EndTime](#)

指定发现作业运行总持续时间内的时间。要查看在特定时间范围内收集的信息，请将此参数与配合使用 `StartTime`。

类型：Timestamp

必需：否

### MaxResults

指定您想要在响应中得到多少结果。

类型：整数

有效范围：最小值为 1。最大值为 100。

必需：否

### NextToken

指定一个不透明的字符串，该字符串指示响应中下一个结果列表的开始位置。

类型：字符串

长长长长长长长长长长长长长长长长长长

模式：`[a-zA-Z0-9=_-]+`

必需：否

### ResourceId

指定您想要了解的存储系统资源的通用唯一标识符 (UUID)。

类型：字符串

模式：`[a-f0-9]{8}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{12}`

必需：是

### ResourceType

指定您想要了解的存储系统资源类型。

类型：字符串

有效值：SVM | VOLUME | CLUSTER

必需：是

## StartTime

指定发现作业运行总持续时间内的时间。要查看在特定时间范围内收集的信息，请将此参数与配合使用EndTime。

类型：Timestamp

必需：否

## 响应语法

```
{
  "Metrics": [
    {
      "Capacity": {
        "LogicalUsed": number,
        "Provisioned": number,
        "Used": number
      },
      "P95Metrics": {
        "IOPS": {
          "Other": number,
          "Read": number,
          "Total": number,
          "Write": number
        },
        "Latency": {
          "Other": number,
          "Read": number,
          "Write": number
        },
        "Throughput": {
          "Other": number,
          "Read": number,
          "Total": number,
          "Write": number
        }
      },
      "ResourceId": "string",
      "ResourceType": "string",
      "Timestamp": number
    }
  ],
}
```

```
"NextToken": "string"  
}
```

## 响应元素

如果此操作成功，则该服务将会发送回 HTTP 200 响应。

服务以 JSON 格式返回的以下数据。

### Metrics

您的发现任务收集的有关存储系统资源的详细信息。

类型：[ResourceMetrics](#) 对象数组

### NextToken

不透明的字符串，表示响应中下一个结果列表的开始位置。

类型：字符串

长长长长长长长长长长长长长长长长长长

模式：`[a-zA-Z0-9=_-]+`

## 错误

有关所有操作常见错误的信息，请参阅[常见错误](#)。

### InternalException

当AWS DataSync服务中发生错误时，会引发此异常。

HTTP 状态代码：500

### InvalidRequestException

当客户端提交格式错误的请求时，会引发此异常。

HTTP 状态代码：400

## 示例

### 示例请求

以下示例请求有关本地存储系统中卷的信息。

```
{
  "DiscoveryJobArn": "arn:aws:datasync:us-east-1:123456789012:system/storage-system-
  abcdef01234567890/job/discovery-job-12345678-90ab-cdef-0abc-021345abcdef6",
  "ResourceType": "VOLUME",
  "ResourceId": "a1b2c3d4-5678-90ab-cdef-EXAMPLE11111"
}
```

### 示例响应

以下示例响应包括有关该卷的性能和容量信息。

```
{
  "Metrics": [
    {
      "Timestamp": "2023-01-10T13:54:11-05:00",
      "P95Metrics": {
        "IOPS": {
          "Read": 251.0,
          "Write": 44.0,
          "Other": 17.0,
          "Total": 345.0
        },
        "Throughput": {
          "Read": 2.06,
          "Write": 0.88,
          "Other": 0.11,
          "Total": 2.17
        },
        "Latency": {
          "Read": 0.06,
          "Write": 0.07,
          "Other": 0.13
        }
      },
      "Capacity": {
        "Used": 409600,
        "Provisioned": 1099511627776
      }
    }
  ]
}
```

```
    },
    "ResourceId": "a1b2c3d4-5678-90ab-cdef-EXAMPLE11111",
    "ResourceType": "VOLUME"
  }
]
}
```

## 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [AWS 命令行界面](#)
- [适用于 .NET 的 AWS SDK](#)
- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [AWSSDK 用于 JavaScript](#)
- [适用于 PHP V3 的 AWS SDK](#)
- [适用于 Python 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)

## DescribeStorageSystemResources

返回 DataSync Discovery 收集的有关本地存储系统中资源的信息。

### 请求语法

```
{
  "DiscoveryJobArn": "string",
  "Filter": {
    "string": [ "string" ]
  },
  "MaxResults": number,
  "NextToken": "string",
  "ResourceIds": [ "string" ],
  "ResourceType": "string"
}
```

### 请求参数

有关所有操作的通用参数的信息，请参阅[通用参数](#)。

请求接受采用 JSON 格式的以下数据。

#### [DiscoveryJobArn](#)

指定您在其中计划的发现作业的 Amazon 资源名称 (ARN)。

类型：字符串

长度限制：最大长度为 256。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\n-0-9]+:[0-9]{12}:system/storage-system-[a-f0-9]{8}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{12}/job/discovery-job-[a-f0-9]{8}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{12}$`

必需：是

#### [Filter](#)

筛选要返回的存储系统资源。例如，这可能是与特定存储虚拟机 (SVM) 关联的卷。

类型：字符串到字符串数组映射



有效密钥：SVM

长度约束：最大长度为 1024。

模式：`^\{0,1024\}$`

必需：否

### MaxResults

指定要在响应中列出的存储系统资源的最大数量。

类型：整数

有效范围：最小值为 1。最大值为 100。

必需：否

### NextToken

指定一个不透明的字符串，该字符串指示响应中下一个结果列表的开始位置。

类型：字符串

长度：最大长度度度度度度度度度，最大长度度度度。

模式：`[a-zA-Z0-9=_-]+`

必需：否

### ResourceIds

指定您想要了解的存储系统资源的通用唯一标识符 (UUID)。不能将此参数 Resource Name Nesour  
Filter ce Name Name

类型：字符串数组

数组成员：最少 1 项。最多 100 项。

模式：`[a-f0-9]{8}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{12}`

必需：否

### ResourceType

指定需要有关哪种存储系统资源的信息。

类型：字符串

有效值：SVM | VOLUME | CLUSTER

必需：是

## 响应语法

```
{
  "NextToken": "string",
  "ResourceDetails": {
    "NetAppONTAPClusters": [
      {
        "CifsShareCount": number,
        "ClusterBlockStorageLogicalUsed": number,
        "ClusterBlockStorageSize": number,
        "ClusterBlockStorageUsed": number,
        "ClusterName": "string",
        "MaxP95Performance": {
          "IopsOther": number,
          "IopsRead": number,
          "IopsTotal": number,
          "IopsWrite": number,
          "LatencyOther": number,
          "LatencyRead": number,
          "LatencyWrite": number,
          "ThroughputOther": number,
          "ThroughputRead": number,
          "ThroughputTotal": number,
          "ThroughputWrite": number
        },
        "NfsExportedVolumes": number,
        "Recommendations": [
          {
            "EstimatedMonthlyStorageCost": "string",
            "StorageConfiguration": {
              "string": "string"
            },
            "StorageType": "string"
          }
        ],
        "RecommendationStatus": "string",
        "ResourceId": "string"
      }
    ],
  },
}
```

```

    "NetAppONTAPSVMS": [
      {
        "CifsShareCount": number,
        "ClusterUuid": "string",
        "EnabledProtocols": [ "string" ],
        "MaxP95Performance": {
          "IopsOther": number,
          "IopsRead": number,
          "IopsTotal": number,
          "IopsWrite": number,
          "LatencyOther": number,
          "LatencyRead": number,
          "LatencyWrite": number,
          "ThroughputOther": number,
          "ThroughputRead": number,
          "ThroughputTotal": number,
          "ThroughputWrite": number
        },
        "NfsExportedVolumes": number,
        "Recommendations": [
          {
            "EstimatedMonthlyStorageCost": "string",
            "StorageConfiguration": {
              "string" : "string"
            },
            "StorageType": "string"
          }
        ],
        "RecommendationStatus": "string",
        "ResourceId": "string",
        "SvmName": "string",
        "TotalCapacityProvisioned": number,
        "TotalCapacityUsed": number,
        "TotalLogicalCapacityUsed": number,
        "TotalSnapshotCapacityUsed": number
      }
    ],
    "NetAppONTAPVolumes": [
      {
        "CapacityProvisioned": number,
        "CapacityUsed": number,
        "CifsShareCount": number,
        "LogicalCapacityUsed": number,
        "MaxP95Performance": {

```

```

    "IopsOther": number,
    "IopsRead": number,
    "IopsTotal": number,
    "IopsWrite": number,
    "LatencyOther": number,
    "LatencyRead": number,
    "LatencyWrite": number,
    "ThroughputOther": number,
    "ThroughputRead": number,
    "ThroughputTotal": number,
    "ThroughputWrite": number
  },
  "NfsExported": boolean,
  "Recommendations": [
    {
      "EstimatedMonthlyStorageCost": "string",
      "StorageConfiguration": {
        "string" : "string"
      },
      "StorageType": "string"
    }
  ],
  "RecommendationStatus": "string",
  "ResourceId": "string",
  "SecurityStyle": "string",
  "SnapshotCapacityUsed": number,
  "SvmName": "string",
  "SvmUuid": "string",
  "VolumeName": "string"
}
]
}
}

```

## 响应元素

如果此操作成功，则该服务将会发送回 HTTP 200 响应。

服务以 JSON 格式返回的以下数据。

### [NextToken](#)

不透明的字符串，表示响应中下一个结果列表的开始位置。

类型：字符串

长度：最大长度度度度度度度度度，最大长度度度度度。

模式：[a-zA-Z0-9=\_-]+

## [ResourceDetails](#)

收集的有关存储系统资源的信息。响应还可以包括AWS存储服务建议。

有关更多信息，请参阅 DataSync Discovery 收集的[存储资源信息](#)和提供的[建议](#)。

类型：[ResourceDetails](#) 对象

## 错误

有关所有操作常见错误的信息，请参阅[常见错误](#)。

### InternalException

当AWS DataSync服务中发生错误时，会引发此异常。

HTTP 状态代码：500

### InvalidRequestException

当客户端提交格式错误的请求时，会引发此异常。

HTTP 状态代码：400

## 示例

### 示例请求

以下示例请求有关本地存储系统中卷的信息。

```
{
  "DiscoveryJobArn": "arn:aws:datasync:us-east-1:123456789012:system/storage-system-
  abcdef01234567890/job/discovery-job-12345678-90ab-cdef-0abc-021345abcdef6",
  "ResourceType": "VOLUME"
}
```

## 示例响应

以下示例响应包含有关本地存储系统中卷的信息，包括有关将卷数据移动到何处的建议AWS（例如适用于 NetApp ONTAP 的 Amazon FSx 或 Amazon EFS）。

```
{
  "ResourceDetails": {
    "NetAppONTAPVolumes": [
      {
        "VolumeName": "vol1",
        "ResourceId": "a1b2c3d4-5678-90ab-cdef-EXAMPLE11111",
        "CifsShareCount": 0,
        "SecurityStyle": "unix",
        "SvmUuid": "a1b2c3d4-5678-90ab-cdef-EXAMPLEaaaaa",
        "SvmName": "my-svm",
        "CapacityUsed": 409600,
        "CapacityProvisioned": 1099511627776,
        "LogicalCapacityUsed": 409600,
        "NfsExported": true,
        "SnapshotCapacityUsed": 573440,
        "MaxP95Performance": {
          "IopsRead": 251.0,
          "IopsWrite": 44.0,
          "IopsOther": 17.0,
          "IopsTotal": 345.0,
          "ThroughputRead": 2.06,
          "ThroughputWrite": 0.88,
          "ThroughputOther": 0.11,
          "ThroughputTotal": 2.17,
          "LatencyRead": 0.06,
          "LatencyWrite": 0.07,
          "LatencyOther": 0.13
        },
        "Recommendations": [
          {
            "StorageType": "fsxOntap",
            "StorageConfiguration": {
              "StorageCapacityGB": "1024",
              "ProvisionedIopsMode": "AUTOMATIC",
              "CapacityPoolGB": "0",
              "TotalIops": "0",
              "DeploymentType": "Multi-AZ",
              "ThroughputCapacity": "128"
            }
          }
        ]
      }
    ]
  }
}
```

```

        "EstimatedMonthlyStorageCost": "410.0"
    },
    {
        "StorageType": "efs",
        "StorageConfiguration": {
            "InfrequentAccessStorageGB": "1",
            "StandardStorageGB": "1",
            "InfrequentAccessRequests": "0",
            "ProvisionedThroughputMBps": "0",
            "PerformanceMode": "General Purpose",
            "ThroughputMode": "Bursting"
        },
        "EstimatedMonthlyStorageCost": "1.0"
    }
],
"RecommendationStatus": "COMPLETED"
},
{
    "VolumeName": "root_vol",
    "ResourceId": "a1b2c3d4-5678-90ab-cdef-EXAMPLE22222",
    "CifsShareCount": 0,
    "SecurityStyle": "unix",
    "SvmUuid": "a1b2c3d4-5678-90ab-cdef-EXAMPLEeaaaaa",
    "SvmName": "my-svm",
    "CapacityUsed": 462848,
    "CapacityProvisioned": 1073741824,
    "LogicalCapacityUsed": 462848,
    "NfsExported": true,
    "SnapshotCapacityUsed": 421888,
    "MaxP95Performance": {
        "IopsRead": 261.0,
        "IopsWrite": 53.0,
        "IopsOther": 23.0,
        "IopsTotal": 360.0,
        "ThroughputRead": 10.0,
        "ThroughputWrite": 2.0,
        "ThroughputOther": 4.0,
        "ThroughputTotal": 12.0,
        "LatencyRead": 0.25,
        "LatencyWrite": 0.3,
        "LatencyOther": 0.55
    },
    "Recommendations": [
        {

```

```
        "StorageType": "fsxOntap",
        "StorageConfiguration": {
            "StorageCapacityGB": "1024",
            "ProvisionedIOpsMode": "AUTOMATIC",
            "CapacityPoolGB": "0",
            "TotalIOps": "0",
            "DeploymentType": "Multi-AZ",
            "ThroughputCapacity": "128"
        },
        "EstimatedMonthlyStorageCost": "410.0"
    },
    {
        "StorageType": "efs",
        "StorageConfiguration": {
            "InfrequentAccessStorageGB": "1",
            "StandardStorageGB": "1",
            "InfrequentAccessRequests": "0",
            "ProvisionedThroughputMBps": "0",
            "PerformanceMode": "General Purpose",
            "ThroughputMode": "Bursting"
        },
        "EstimatedMonthlyStorageCost": "1.0"
    }
],
"RecommendationStatus": "COMPLETED"
}
]
}
```

## 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [AWS 命令行界面](#)
- [适用于 .NET 的 AWS SDK](#)
- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [AWSSDK 用于 JavaScript](#)
- [适用于 PHP V3 的 AWS SDK](#)



- [适用于 Python 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)

## DescribeTask

返回任务的相关元数据。

### 请求语法

```
{  
  "TaskArn": "string"  
}
```

### 请求参数

有关所有操作的通用参数的信息，请参阅[通用参数](#)。

请求接受采用 JSON 格式的以下数据。

#### TaskArn

要描述的任务的 Amazon 资源名称 (ARN)。

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\n-0-9]*:[0-9]{12}:task/task-[0-9a-f]{17}$`

必需：是

### 响应语法

```
{  
  "CloudWatchLogGroupArn": "string",  
  "CreationTime": number,  
  "CurrentTaskExecutionArn": "string",  
  "DestinationLocationArn": "string",  
  "DestinationNetworkInterfaceArns": [ "string" ],  
  "ErrorCode": "string",  
  "ErrorDetail": "string",  
  "Excludes": [  
    {  
      "FilterType": "string",  
      "Value": "string"  
    }  
  ]  
}
```

```

    }
  ],
  "Includes": [
    {
      "FilterType": "string",
      "Value": "string"
    }
  ],
  "Name": "string",
  "Options": {
    "Atime": "string",
    "BytesPerSecond": number,
    "Gid": "string",
    "LogLevel": "string",
    "Mtime": "string",
    "ObjectTags": "string",
    "OverwriteMode": "string",
    "PosixPermissions": "string",
    "PreserveDeletedFiles": "string",
    "PreserveDevices": "string",
    "SecurityDescriptorCopyFlags": "string",
    "TaskQueueing": "string",
    "TransferMode": "string",
    "Uid": "string",
    "VerifyMode": "string"
  },
  "Schedule": {
    "ScheduleExpression": "string"
  },
  "SourceLocationArn": "string",
  "SourceNetworkInterfaceArns": [ "string" ],
  "Status": "string",
  "TaskArn": "string"
}

```

## 响应元素

如果此操作成功，则该服务将会发送回 HTTP 200 响应。

服务以 JSON 格式返回的以下数据。

### [CloudWatchLogGroupArn](#)

Amazon CloudWatch 日志组（用于监控和记录任务中的事件）的 Amazon 资源名称（ARN）。

有关这些日志组的更多信息，请参阅《Amazon CloudWatch 用户指南》中的[使用日志组和日志流](#)。

类型：字符串

长度约束：最大长度为 562。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):logs:[a-z\-\0-9]+:[0-9]{12}:log-group:([^\:\*]*)(:\*)?$$`

### [CreationTime](#)

任务的创建时间。

类型：时间戳

### [CurrentTaskExecutionArn](#)

传输文件的任务的 Amazon 资源名称 ( ARN )。

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\-\0-9]*:[0-9]{12}:task/task-[0-9a-f]{17}/execution/exec-[0-9a-f]{17}$$`

### [DestinationLocationArn](#)

AWS 存储资源位置的 Amazon 资源名称 ( ARN )。

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\-\0-9]+:[0-9]{12}:location/loc-[0-9a-z]{17}$$`

### [DestinationNetworkInterfaceArns](#)

为目标位置创建的网络接口的 Amazon 资源名称 ( ARN )。有关更多信息，请参阅[网络接口要求](#)。

类型：字符串数组

长度约束：最大长度为 128。

模式：`^arn:aws[\-a-z]{0,}:ec2:[a-z\-\0-9]*:[0-9]{12}:network-interface/eni-[0-9a-f]+$`

### [ErrorCode](#)

执行任务期间AWS DataSync遇到的错误。您可以使用此错误代码来帮助解决问题。

类型：字符串

### [ErrorDetail](#)

对任务执行过程中遇到的错误的详细描述。您可以使用此信息来帮助排查问题。

类型：字符串

### [Excludes](#)

在传输过程中排除特定数据的筛选规则列表。有关更多信息和示例，请参阅 [Filtering by DataSferred DataSync by](#)

类型：[FilterRule](#) 对象数组

数组成员：最少 0 项。最多 1 项。

### [Includes](#)

在传输期间包含特定数据的筛选规则列表。有关更多信息和示例，请参阅 [Filtering by DataSferred DataSync by](#)

类型：[FilterRule](#) 对象数组

数组成员：最少 0 项。最多 1 项。

### [Name](#)

所描述的任务的名称。

类型：字符串

长度限制：最小长度为 1。长度上限为 256。

模式：`^[a-zA-Z0-9\s+=._:@/-]+$`

### [Options](#)

控制StartTaskExecution操作行为的配置选项。一些选项包括保留文件或对象元数据以及验证数据完整性。

每次执行任务时，您可以重写这些选项。有关更多信息，请参阅[StartTaskExecution](#)。

类型：[Options](#) 对象

### [Schedule](#)

用于定期将文件从源位置传输到目标位置的时间表。

类型：[TaskSchedule](#) 对象

### [SourceLocationArn](#)

源文件系统位置的 Amazon 资源名称 (ARN)。

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\\-0-9]+:[0-9]{12}:location/loc-[0-9a-z]{17}$`

### [SourceNetworkInterfaceArns](#)

为源位置创建的网络接口的 Amazon 资源名称 (ARN)。有关更多信息，请参阅[网络接口要求](#)。

类型：字符串数组

长度约束：最大长度为 128。

模式：`^arn:aws[\\-a-z]{0,}:ec2:[a-z\\-0-9]*:[0-9]{12}:network-interface/eni-[0-9a-f]+$`

### [Status](#)

所描述的任务的状态。

有关任务执行状态的详细信息，请参阅《AWS DataSync用户指南》中的“[了解任务状态](#)”。

类型：字符串

有效值：AVAILABLE | CREATING | QUEUED | RUNNING | UNAVAILABLE

### [TaskArn](#)

所描述的任务的 Amazon 资源名称 (ARN)。

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\-\0-9]*:[0-9]{12}:task/task-[0-9a-f]{17}$`

## 错误

有关所有操作常见错误的信息，请参阅[常见错误](#)。

### InternalException

当AWS DataSync服务中发生错误时，会引发此异常。

HTTP 状态代码：500

### InvalidRequestException

当客户端提交格式错误的请求时，会引发此异常。

HTTP 状态代码：400

## 示例

### 示例

以下示例返回有关示例请求中指定的任务的信息。

### 示例请求

```
{
  "TaskArn": "arn:aws:datasync:us-east-2:111222333444:task/task-08de6e6697796f026"
}
```

### 示例

此示例说明了的一种用法DescribeTask。

### 示例响应

```
{
  "CloudWatchLogGroupArn": "arn:aws:logs:us-east-2:111222333444:log-group"
```

```
"CreationTime": 1532660733.39,
"CurrentTaskExecutionArn": "arn:aws:datsync:us-east-2:111222333444:task/
task-08de6e6697796f026/execution/exec-04ce9d516d69bd52f",
"Options": {
  "Atime": "BEST_EFFORT",
  "BytesPerSecond": 1000,
  "Gid": "NONE",
  "Mtime": "PRESERVE",
  "PosixPermissions": "PRESERVE",
  "PreserveDevices": "NONE",
  "PreserveDeletedFiles": "PRESERVE",
  "Uid": "NONE",
  "VerifyMode": "POINT_IN_TIME_CONSISTENT"
},
"DestinationLocationArn": "arn:aws:datsync:us-east-2:111222333444:location/
loc-07db7abfc326c50fb",
"ErrorCode": "????????",
"ErrorDetail": "??????",
"Name": "MyTask",
"SourceLocationArn": "arn:aws:datsync:us-east-2:111222333444:location/
loc-07db7abfc326c50aa",
"Status": "CREATING",
"TaskArn": "arn:aws:datsync:us-east-2:111222333444:task/task-08de6e6697796f026"
}
```

## 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [AWS 命令行界面](#)
- [适用于 .NET 的 AWS SDK](#)
- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [AWSSDK 用于 JavaScript](#)
- [适用于 PHP V3 的 AWS SDK](#)
- [适用于 Python 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)



## DescribeTaskExecution

返回有关正在执行的任务的详细元数据。

### 请求语法

```
{  
  "TaskExecutionArn": "string"  
}
```

### 请求参数

有关所有操作的通用参数的信息，请参阅[通用参数](#)。

请求接受采用 JSON 格式的以下数据。

#### [TaskExecutionArn](#)

正在执行的任务的 Amazon 资源名称 ( ARN ) 。

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\n-0-9]*:[0-9]{12}:task/task-[0-9a-f]{17}/execution/exec-[0-9a-f]{17}$`

必需：是

### 响应语法

```
{  
  "BytesCompressed": number,  
  "BytesTransferred": number,  
  "BytesWritten": number,  
  "EstimatedBytesToTransfer": number,  
  "EstimatedFilesToTransfer": number,  
  "Excludes": [  
    {  
      "FilterType": "string",  
      "Value": "string"  
    }  
  ],  
}
```

```
"FilesTransferred": number,
"Includes": [
  {
    "FilterType": "string",
    "Value": "string"
  }
],
"Options": {
  "Atime": "string",
  "BytesPerSecond": number,
  "Gid": "string",
  "LogLevel": "string",
  "Mtime": "string",
  "ObjectTags": "string",
  "OverwriteMode": "string",
  "PosixPermissions": "string",
  "PreserveDeletedFiles": "string",
  "PreserveDevices": "string",
  "SecurityDescriptorCopyFlags": "string",
  "TaskQueueing": "string",
  "TransferMode": "string",
  "Uid": "string",
  "VerifyMode": "string"
},
"Result": {
  "ErrorCode": "string",
  "ErrorDetail": "string",
  "PrepareDuration": number,
  "PrepareStatus": "string",
  "TotalDuration": number,
  "TransferDuration": number,
  "TransferStatus": "string",
  "VerifyDuration": number,
  "VerifyStatus": "string"
},
"StartTime": number,
"Status": "string",
"TaskExecutionArn": "string"
}
```

## 响应元素

如果此操作成功，则该服务将会发送回 HTTP 200 响应。

服务以 JSON 格式返回的以下数据。

### [BytesCompressed](#)

应用压缩后通过网络传输的物理字节数。在大多数情况下，除非数据不可压缩，BytesTransferred否则此数字会小于。

类型：长整型

### [BytesTransferred](#)

传输中涉及的总字节数。有关通过网络发送的字节数，请参阅 [Filtering DataSyncs BytesCompressed](#) fan

类型：长整型

### [BytesWritten](#)

写入目标AWS存储资源的逻辑字节数。

类型：长整型

### [EstimatedBytesToTransfer](#)

通过网络传输的估计物理字节数。

类型：长整型

### [EstimatedFilesToTransfer](#)

通过网络传输的预期文件数。该值是在任务执行PREPARING阶段之前的TRANSFERRING阶段计算的。此值是预期要传输的文件数。它是在比较源位置和目标位置的内容并找到需要传输的增量的基础上计算得出的。

类型：长整型

### [Excludes](#)

在传输过程中排除特定数据的筛选规则列表。有关更多信息和示例，请参阅 [Filtering DataSyncsfer DataSync](#) red

类型：[FilterRule](#) 对象数组

数组成员：最少 0 项。最多 1 项。

## FilesTransferred

通过网络传输的实际文件数。该值是在任务执行TRANSFERRING阶段持续计算和更新的。当从源读取每个文件并通过网络发送时，它会定期更新。

如果传输期间发生故障，该值可能小于 `EstimatedFilesToTransfer`。在某些情况下，此值也可能大于 `EstimatedFilesToTransfer`。此元素是特定于某些位置类型的实施，因此不要将其用作正确文件编号的指示符或监控任务执行。

类型：长整型

## Includes

在传输过程中包含特定数据的筛选规则列表。有关更多信息和示例，请参阅 [Filtering DataSync red](#)

类型：[FilterRule](#) 对象数组

数组成员：最少 0 项。最多 1 项。

## Options

配置您的AWS DataSync任务设置。这些选项包括如何DataSync处理文件、对象及其关联的元数据。您还可以指定如何DataSync验证数据完整性、为任务设置带宽限制以及其他选项。

每个任务设置都有默认值。除非需要，否则Options在开始任务之前不必配置任何一项。

类型：[Options](#) 对象

## Result

任务执行的状态。

类型：[TaskExecutionResultDetail](#) 对象

## StartTime

开始执行任务的时间。

类型：时间戳

## Status

任务执行的状态。

有关任务执行状态的详细信息，请参阅[了解任务状态](#)。

类型：字符串

有效值：QUEUED | LAUNCHING | PREPARING | TRANSFERRING | VERIFYING | SUCCESS | ERROR

### [TaskExecutionArn](#)

所描述任务执行的 Amazon 资源名称 (ARN)。TaskExecutionArn 是分层 TaskArn 的，包括已执行的任务。

例如，具有 ARN 的 TaskExecution 值使用 ARN `arn:aws:datsync:us-east-1:111222333444:task/task-0208075f79cedf4a2/execution/exec-08ef1e88ec491019b` 执行任务 `arn:aws:datsync:us-east-1:111222333444:task/task-0208075f79cedf4a2`。

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datsync:[a-z\-\0-9]*:[0-9]{12}:task/task-[0-9a-f]{17}/execution/exec-[0-9a-f]{17}$`

## 错误

有关所有操作常见错误的信息，请参阅[常见错误](#)。

### InternalException

当 AWS DataSync 服务中发生错误时，会引发此异常。

HTTP 状态代码：500

### InvalidRequestException

当客户端提交格式错误的请求时，会引发此异常。

HTTP 状态代码：400

## 示例

### 示例请求

此示例说明了 DescribeTaskExecution 请求。

```
{
  "TaskExecutionArn": "arn:aws:datsync:us-east-1:111222333444:task/
task-08de6e6697796f026/execution/exec-04ce9d516d69bd52f"
}
```

## 示例响应

此示例说明了DescribeTaskExecution响应。

```
{
  "BytesCompressed": "3500",
  "BytesTransferred": "5000",
  "BytesWritten": "5000",
  "EstimatedBytesToTransfer": "5000",
  "EstimatedFilesToTransfer": "100",
  "FilesTransferred": "100",
  "Result": {
    "ErrorCode": "???????",
    "ErrorDetail": "???????",
    "PrepareDuration": "100",
    "PrepareStatus": "SUCCESS",
    "TransferDuration": "60",
    "TransferStatus": "AVAILABLE",
    "VerifyDuration": "30",
    "VerifyStatus": "SUCCESS"
  },
  "StartTime": "1532660733.39",
  "Status": "SUCCESS",
  "OverrideOptions": {
    "Atime": "BEST_EFFORT",
    "BytesPerSecond": "1000",
    "Gid": "NONE",
    "Mtime": "PRESERVE",
    "PosixPermissions": "PRESERVE",
    "PreserveDevices": "NONE",
    "PreserveDeletedFiles": "PRESERVE",
    "Uid": "NONE",
    "VerifyMode": "POINT_IN_TIME_CONSISTENT"
  },
  "TaskExecutionArn": "arn:aws:datsync:us-east-2:111222333444:task/
task-08de6e6697796f026/execution/exec-04ce9d516d69bd52f"
}
```

## 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [AWS 命令行界面](#)
- [适用于 .NET 的 AWS SDK](#)
- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [AWSSDK 用于 JavaScript](#)
- [适用于 PHP V3 的 AWS SDK](#)
- [适用于 Python 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)

## GenerateRecommendations

创建有关将数据迁移到何处的建议AWS。建议是基于 DataSync Discovery 收集的有关本地存储系统资源的信息生成的。有关更多信息，请参阅 [DataSyncDiscover Name 提供的建议](#)。

生成后，您可以使用 [DescribeStorageSystemResources](#) 操作查看您的建议。

### Note

如果您的 [发现任务成功完成](#)，则无需使用此操作。DataSyncDiscovery 会自动为您生成推荐。

## 请求语法

```
{
  "DiscoveryJobArn": "string",
  "ResourceIds": [ "string" ],
  "ResourceType": "string"
}
```

## 请求参数

有关所有操作的通用参数的信息，请参阅 [通用参数](#)。

请求接受采用 JSON 格式的以下数据。

### [DiscoveryJobArn](#)

指定您在其中计划的发现作业的 Amazon 资源名称 ( ARN )。

类型：字符串

长度限制：最大长度为 256。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\n-0-9]+:[0-9]{12}:system/storage-system-[a-f0-9]{8}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{12}/job/discovery-job-[a-f0-9]{8}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{12}$`

必需：是



## ResourceIds

为存储系统中要推荐的资源指定通用唯一标识符 (UUID)。

类型：字符串数组

数组成员：最少 1 项。最多 100 项。

模式：`[a-f0-9]{8}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{12}`

必需：是

## ResourceType

指定存储系统中需要推荐的资源类型。

类型：字符串

有效值：SVM | VOLUME | CLUSTER

必需：是

## 响应元素

如果此操作成功，则该服务会发送回带有空 HTTP 正文的 HTTP 200 响应。

## 错误

有关所有操作常见错误的信息，请参阅[常见错误](#)。

### InternalException

当AWS DataSync服务中发生错误时，会引发此异常。

HTTP 状态代码：500

### InvalidRequestException

当客户端提交格式错误的请求时，会引发此异常。

HTTP 状态代码：400

## 示例

### 示例请求

以下示例为本地AWS存储系统中的卷生成存储建议。

```
{
  "DiscoveryJobArn": "arn:aws:datsync:us-east-1:123456789012:system/storage-system-
  abcdef01234567890/job/discovery-job-12345678-90ab-cdef-0abc-021345abcdef6",
  "ResourceIds": [
    "a1b2c3d4-5678-90ab-cdef-EXAMPLE33333"
  ],
  "ResourceType": "VOLUME"
}
```

### 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [AWS 命令行界面](#)
- [适用于 .NET 的 AWS SDK](#)
- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [AWSSDK 用于 JavaScript](#)
- [适用于 PHP V3 的 AWS SDK](#)
- [适用于 Python 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)

## ListAgents

返回属于请求AWS 账户中AWS 区域指定的 a 的AWS DataSync代理列表。

通过分页，您可以减少响应中返回的代理数量。如果您在响应中收到截断的代理列表，则响应包含一个标记，您可以在下一个请求中指定该标记以获取下一页代理。

ListAgents结果。这意味着运行操作的结果可能无法反映出您刚刚创建或删除了代理。例如，如果您使用创建代理[CreateAgent](#)然后立即运行ListAgents，则该代理可能不会立即出现在列表中。在这种情况下，您可以随时使用来确认代理是否已创建（或删除）[DescribeAgent](#)。

### 请求语法

```
{
  "MaxResults": number,
  "NextToken": "string"
}
```

### 请求参数

有关所有操作的通用参数的信息，请参阅[通用参数](#)。

请求接受采用 JSON 格式的以下数据。

#### [MaxResults](#)

指定要在响应中列出的。DataSync默认情况下，响应最多显示 100 个代理。

类型：整数

有效范围：最小值为 0。最大值为 100。

必需：否

#### [NextToken](#)

指定一个不透明的字符串，该字符串指示响应中下一个结果列表的起始位置。

类型：字符串

长度束：最大长度度度度度度度束：最大长度度度

模式：[a-zA-Z0-9=\_-]+

必需：否

## 响应语法

```
{
  "Agents": [
    {
      "AgentArn": "string",
      "Name": "string",
      "Status": "string"
    }
  ],
  "NextToken": "string"
}
```

## 响应元素

如果此操作成功，则该服务将会发送回 HTTP 200 响应。

服务以 JSON 格式返回的以下数据。

### Agents

在请求AWS 账户中AWS 区域指定的您的DataSync代理列表。列表按代理商的 Amazon 资源名称 (ARN)。

类型：[AgentListEntry](#) 对象数组

### NextToken

不透明的字符串，表示响应中下一个结果列表的起始位置。

类型：字符串

长度束：最大长度度度度度度度度束：最大长度度度

模式：`[a-zA-Z0-9=_-]+`

## 错误

有关所有操作常见错误的信息，请参阅[常见错误](#)。

## InternalException

当AWS DataSync服务中发生错误时，会抛出此异常。

HTTP 状态代码：500

## InvalidRequestException

当客户端提交格式错误的请求时，会引发此异常。

HTTP 状态代码：400

## 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [AWS 命令行界面](#)
- [适用于 .NET 的 AWS SDK](#)
- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [AWS适用于 SDKJavaScript](#)
- [适用于 PHP V3 的 AWS SDK](#)
- [适用于 Python 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)

## ListDiscoveryJobs

提供中现有发现任务的列表AWS 区域以及您在AWS 账户何处使用DataSync Discovery。

### 请求语法

```
{  
  "MaxResults": number,  
  "NextToken": "string",  
  "StorageSystemArn": "string"  
}
```

### 请求参数

有关所有操作的通用参数的信息，请参阅[常用参数](#)。

请求接受采用 JSON 格式的以下数据。

#### [MaxResults](#)

指定您想要在响应中获得多少结果。

类型：整数

有效范围：最小值为 1。最大值为 100。

必需：否

#### [NextToken](#)

指定一个不透明的字符串，该字符串指示响应中下一个结果列表的起始位置。

类型：字符串

长度约束：最大长度为 65535。

模式：[a-zA-Z0-9=\_-]+

必需：否

#### [StorageSystemArn](#)

指定本地存储位置的 Amazon 资源位置名称 (ARN)。如果您只想列出与特定存储系统关联的发现任务，请使用此参数。

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\\-0-9]+:[0-9]{12}:system/storage-system-[a-f0-9]{8}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{12}$`

必需：否

## 响应语法

```
{
  "DiscoveryJobs": [
    {
      "DiscoveryJobArn": "string",
      "Status": "string"
    }
  ],
  "NextToken": "string"
}
```

## 响应元素

如果此操作成功，则该服务将会发送回 HTTP 200 响应。

服务以 JSON 格式返回的以下数据。

### [DiscoveryJobs](#)

你运行的发现任务。

类型：[DiscoveryJobListEntry](#) 对象数组

### [NextToken](#)

不透明的字符串，表示响应中下一个结果列表的起始位置。

类型：字符串

长度约束：最大长度为 65535。

模式：`[a-zA-Z0-9=_-]+`

## 错误

有关所有操作常见错误的信息，请参阅[常见错误](#)。

### InternalException

当AWS DataSync服务中发生错误时，会抛出此异常。

HTTP 状态代码：500

### InvalidRequestException

当客户端提交格式错误的请求时，会引发此异常。

HTTP 状态代码：400

## 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [AWS 命令行界面](#)
- [适用于 .NET 的 AWS SDK](#)
- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [AWS适用于JavaScript](#)
- [适用于 PHP V3 的 AWS SDK](#)
- [适用于 Python 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)



## ListLocations

返回源和目标位置列表。

如果您的位置多于响应中返回的地点（即，响应仅返回截断的代理列表），则响应包含一个令牌，您可以在下次请求中指定该令牌以获取下一页地点。

### 请求语法

```
{
  "Filters": [
    {
      "Name": "string",
      "Operator": "string",
      "Values": [ "string" ]
    }
  ],
  "MaxResults": number,
  "NextToken": "string"
}
```

### 请求参数

有关所有操作的通用参数的信息，请参阅[常用参数](#)。

请求接受采用 JSON 格式的以下数据。

#### [Filters](#)

您可以使用 API 过滤器来缩小返回的资源列表的范围 ListLocations。例如，要检索特定源位置上的所有任务，可以将过滤器名称 LocationType S3 和 ListLocations 一起使用 Operator Equals。

类型：[LocationFilter](#) 对象数组

必需：否

#### [MaxResults](#)

要返回的位置的最大数量。

类型：整数

有效范围：最小值为 0。最大值为 100。

必需：否

### [NextToken](#)

一个不透明的字符串，表示下一个位置列表的起始位置。

类型：字符串

长度限制：最大长度为 65535。

模式：`[a-zA-Z0-9=_-]+`

必需：否

## 响应语法

```
{
  "Locations": [
    {
      "LocationArn": "string",
      "LocationUri": "string"
    }
  ],
  "NextToken": "string"
}
```

## 响应元素

如果此操作成功，则该服务将会发送回 HTTP 200 响应。

服务以 JSON 格式返回的以下数据。

### [Locations](#)

包含位置列表的数组。

类型：[LocationListEntry](#) 对象数组

### [NextToken](#)

一个不透明的字符串，表示从哪个位置开始返回下一个位置列表。

类型：字符串

长度限制：最大长度为 65535。

模式：`[a-zA-Z0-9=_-]+`

## 错误

有关所有操作常见错误的信息，请参阅[常见错误](#)。

### InternalException

当AWS DataSync服务中发生错误时，会抛出此异常。

HTTP 状态代码：500

### InvalidRequestException

当客户端提交格式错误的请求时，会引发此异常。

HTTP 状态代码：400

## 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [AWS 命令行界面](#)
- [适用于 .NET 的 AWS SDK](#)
- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [AWS适用于JavaScript](#)
- [适用于 PHP V3 的 AWS SDK](#)
- [适用于 Python 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)

## ListStorageSystems

列出您在DataSync Discovery 中使用的本地存储系统。

### 请求语法

```
{
  "MaxResults": number,
  "NextToken": "string"
}
```

### 请求参数

有关所有操作的通用参数的信息，请参阅[常用参数](#)。

请求接受采用 JSON 格式的以下数据。

#### [MaxResults](#)

指定您想要在响应中获得多少结果。

类型：整数

有效范围：最小值为 1。最大值为 100。

必需：否

#### [NextToken](#)

指定一个不透明的字符串，该字符串指示响应中下一个结果列表的起始位置。

类型：字符串

长度约束：最大长度为 65535。

模式：`[a-zA-Z0-9=_-]+`

必需：否

### 响应语法

```
{
```

```
"NextToken": "string",  
"StorageSystems": [  
  {  
    "Name": "string",  
    "StorageSystemArn": "string"  
  }  
]  
}
```

## 响应元素

如果此操作成功，则该服务将会发送回 HTTP 200 响应。

服务以 JSON 格式返回的以下数据。

### NextToken

不透明的字符串，表示响应中下一个结果列表的起始位置。

类型：字符串

长度约束：最大长度为 65535。

模式：[a-zA-Z0-9=\_-]+

### StorageSystems

您在DataSync Discovery 中使用的本地存储系统的亚马逊资源名称 (ARN)。

类型：[StorageSystemListEntry](#) 对象数组

## 错误

有关所有操作常见错误的信息，请参阅[常见错误](#)。

### InternalException

当AWS DataSync服务中发生错误时，会抛出此异常。

HTTP 状态代码：500

### InvalidRequestException

当客户端提交格式错误的请求时，会引发此异常。

HTTP 状态代码：400

## 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [AWS 命令行界面](#)
- [适用于 .NET 的 AWS SDK](#)
- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [AWS适用于JavaScript](#)
- [适用于 PHP V3 的 AWS SDK](#)
- [适用于 Python 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)

## ListTagsForResource

返回与AWS资源关联的所有标签。

### 请求语法

```
{  
  "MaxResults": number,  
  "NextToken": "string",  
  "ResourceArn": "string"  
}
```

### 请求参数

有关所有操作的通用参数的信息，请参阅[通用参数](#)。

请求接受采用 JSON 格式的以下数据。

#### [MaxResults](#)

指定您想要在响应中获得多少结果。

类型：整数

有效范围：最小值为 0。最大值为 100。

必需：否

#### [NextToken](#)

指定一个不透明的字符串，该字符串指示响应中下一个结果列表的起始位置。

类型：字符串

长度限制：最大长度为 65535。

模式：[a-zA-Z0-9=\_-]+

必需：否

#### [ResourceArn](#)

指定要提供要使用标签信息的资源名称 (ARN)。

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\\-0-9]+:[0-9]{12}:(agent|task|location|system)/((agent|task|loc)-[a-f0-9]{17}|storage-system-[a-f0-9]{8}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{12})/(/execution/exec-[a-f0-9]{17})?&`

必需：是

## 响应语法

```
{
  "NextToken": "string",
  "Tags": [
    {
      "Key": "string",
      "Value": "string"
    }
  ]
}
```

## 响应元素

如果此操作成功，则该服务将会发送回 HTTP 200 响应。

服务以 JSON 格式返回的以下数据。

### [NextToken](#)

不透明的字符串，表示响应中下一个结果列表的起始位置。

类型：字符串

长度限制：最大长度为 65535。

模式：`[a-zA-Z0-9=_-]+`

### [Tags](#)

应用于指定资源的标签数组。



类型：[TagListEntry](#) 对象数组

数组成员：最少 0 项。最多 10 项项项项项项项项项项

## 错误

有关所有操作常见错误的信息，请参阅[常见错误](#)。

### InternalException

当AWS DataSync服务中发生错误时，会抛出此异常。

HTTP 状态代码：500

### InvalidRequestException

当客户端提交格式错误的请求时，会引发此异常。

HTTP 状态代码：400

## 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [AWS 命令行界面](#)
- [适用于 .NET 的 AWS SDK](#)
- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [AWS适用于 SDKJavaScript](#)
- [适用于 PHP V3 的 AWS SDK](#)
- [适用于 Python 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)

## ListTaskExecutions

返回执行任务列表。

### 请求语法

```
{
  "MaxResults": number,
  "NextToken": "string",
  "TaskArn": "string"
}
```

### 请求参数

有关所有操作的通用参数的信息，请参阅[通用参数](#)。

请求接受采用 JSON 格式的以下数据。

#### [MaxResults](#)

要列出的最大执行任务数量。

类型：整数

有效范围：最小值为 0。最大值为 100。

必需：否

#### [NextToken](#)

一个不透明的字符串，表示下一个已执行任务列表的起始位置。

类型：字符串

长度限制：最大长度为 65535。

模式：`[a-zA-Z0-9=_-]+`

必需：否

#### [TaskArn](#)

您要列出任务的 Amazon 资源名称 (ARN)。

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\\-0-9]*:[0-9]{12}:task/task-[0-9a-f]{17}$`

必需：否

## 响应语法

```
{
  "NextToken": "string",
  "TaskExecutions": [
    {
      "Status": "string",
      "TaskExecutionArn": "string"
    }
  ]
}
```

## 响应元素

如果此操作成功，则该服务将会发送回 HTTP 200 响应。

服务以 JSON 格式返回的以下数据。

### [NextToken](#)

一个不透明的字符串，表示从哪个位置开始返回下一个已执行任务列表。

类型：字符串

长度限制：最大长度为 65535。

模式：`[a-zA-Z0-9=_-]+`

### [TaskExecutions](#)

已执行任务的列表。

类型：[TaskExecutionListEntry](#) 对象数组

## 错误

有关所有操作常见错误的信息，请参阅[常见错误](#)。

### InternalException

当AWS DataSync服务中发生错误时，会抛出此异常。

HTTP 状态代码：500

### InvalidRequestException

当客户端提交格式错误的请求时，会引发此异常。

HTTP 状态代码：400

## 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [AWS 命令行界面](#)
- [适用于 .NET 的 AWS SDK](#)
- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [AWS适用于 SDKJavaScript](#)
- [适用于 PHP V3 的 AWS SDK](#)
- [适用于 Python 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)

## ListTasks

返回您创建的AWS DataSync任务的列表。

### 请求语法

```
{
  "Filters": [
    {
      "Name": "string",
      "Operator": "string",
      "Values": [ "string" ]
    }
  ],
  "MaxResults": number,
  "NextToken": "string"
}
```

### 请求参数

有关所有操作的通用参数的信息，请参阅[常用参数](#)。

请求接受采用 JSON 格式的以下数据。

#### [Filters](#)

您可以使用 API 过滤器来缩小返回的资源列表的范围ListTasks。例如，要检索特定源位置上的所有任务，您可以将筛选器名称LocationId和该位置的 ARNListTasksOperator Equals 一起使用。

类型：[TaskFilter](#) 对象数组

必需：否

#### [MaxResults](#)

要返回的任务最大数量。

类型：整数

有效范围：最小值为 0。最大值为 100。

必需：否

## NextToken

一个不透明的字符串，表示下一个任务列表的起始位置。

类型：字符串

长度约束：最大长度为 65535。

模式：`[a-zA-Z0-9=_-]+`

必需：否

## 响应语法

```
{
  "NextToken": "string",
  "Tasks": [
    {
      "Name": "string",
      "Status": "string",
      "TaskArn": "string"
    }
  ]
}
```

## 响应元素

如果此操作成功，则该服务将会发送回 HTTP 200 响应。

服务以 JSON 格式返回的以下数据。

## NextToken

一个不透明的字符串，表示从哪个位置开始返回下一个任务列表。

类型：字符串

长度约束：最大长度为 65535。

模式：`[a-zA-Z0-9=_-]+`

## Tasks

返回的所有任务的列表。

类型：[TaskListEntry](#) 对象数组

## 错误

有关所有操作常见错误的信息，请参阅[常见错误](#)。

### InternalException

当AWS DataSync服务中发生错误时，会抛出此异常。

HTTP 状态代码：500

### InvalidRequestException

当客户端提交格式错误的请求时，会引发此异常。

HTTP 状态代码：400

## 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [AWS 命令行界面](#)
- [适用于 .NET 的 AWS SDK](#)
- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [AWS适用于JavaScript](#)
- [适用于 PHP V3 的 AWS SDK](#)
- [适用于 Python 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)

## RemoveStorageSystem

从DataSync Discovery 中永久删除存储系统资源，包括相关的发现任务、收集的数据和建议。

### 请求语法

```
{
  "StorageSystemArn": "string"
}
```

### 请求参数

有关所有操作的通用参数的信息，请参阅[常用参数](#)。

请求接受采用 JSON 格式的以下数据。

#### StorageSystemArn

指定要从DataSync Discovery 中永久删除的存储系统的Amazon 资源名称 (ARN)。

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\n-0-9]+:[0-9]{12}:system/storage-system-[a-f0-9]{8}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{12}$`

必需：是

### 响应元素

如果此操作成功，则该服务会发送回带有空 HTTP 正文的 HTTP 200 响应。

### 错误

有关所有操作常见错误的信息，请参阅[常见错误](#)。

#### InternalException

当AWS DataSync服务中发生错误时，会抛出此异常。



HTTP 状态代码：500

InvalidRequestException

当客户端提交格式错误的请求时，会引发此异常。

HTTP 状态代码：400

## 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [AWS 命令行界面](#)
- [适用于 .NET 的 AWS SDK](#)
- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [AWS适用于JavaScript](#)
- [适用于 PHP V3 的 AWS SDK](#)
- [适用于 Python 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)

## StartDiscoveryJob

在本地存储系统上运行DataSync发现任务。如果您尚未将存储系统添加到DataSync Discovery，请先使用操作执行此[AddStorageSystem](#)操作。

### 请求语法

```
{
  "ClientToken": "string",
  "CollectionDurationMinutes": number,
  "StorageSystemArn": "string",
  "Tags": [
    {
      "Key": "string",
      "Value": "string"
    }
  ]
}
```

### 请求参数

有关所有操作的通用参数的信息，请参阅[常用参数](#)。

请求接受采用 JSON 格式的以下数据。

#### [ClientToken](#)

指定客户端令牌以确保使用此 API 操作的请求是幂等的。如果您没有指定客户名称，则 100000000000DataSync000000000000

类型：字符串

模式：[a-f0-9]{8}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{12}

必需：是

#### [CollectionDurationMinutes](#)

指定您希望的预签名 (Resource Name)。

**Note**

要获得更准确的建议，我们建议的持续时间至少为 14 天。更长的持续时间可以留出时间收集足够数量的数据点，并真实地呈现存储性能和利用率。

类型：整数

有效范围：最小值为 60。最大值为 40000 0 0 0 0 0 0 0 0

必需：是

### StorageSystemArn

指定要用于运行发现任务的 Amazon Resource Name (ARN)。

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\n-0-9]+:[0-9]{12}:system/storage-system-[a-f0-9]{8}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{12}$`

必需：是

### Tags

指定帮助您分类、筛选和搜索 AWS 资源的标签。

类型：[TagListEntry](#) 对象数组

数组成员：最少 0 项。最多 50 项。

必需：否

### 响应语法

```
{
  "DiscoveryJobArn": "string"
}
```

## 响应元素

如果此操作成功，则该服务将会发送回 HTTP 200 响应。

服务以 JSON 格式返回的以下数据。

### DiscoveryJobArn

您启动的发现任务的 ARN。

类型：字符串

长度限制：最大长度为 256。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\n-0-9]+:[0-9]{12}:system/storage-system-[a-f0-9]{8}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{12}/job/discovery-job-[a-f0-9]{8}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{12}$`

## 错误

有关所有操作常见错误的信息，请参阅[常见错误](#)。

### InternalException

当AWS DataSync服务中发生错误时，会抛出此异常。

HTTP 状态代码：500

### InvalidRequestException

当客户端提交格式错误的请求时，会引发此异常。

HTTP 状态代码：400

## 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [AWS 命令行界面](#)
- [适用于 .NET 的 AWS SDK](#)

- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [AWS适用于JavaScript](#)
- [适用于 PHP V3 的 AWS SDK](#)
- [适用于 Python 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)

## StartTaskExecution

启动AWS DataSync任务。对于每个任务，您一次只能运行一个任务执行任务。

任务执行分为几个阶段。有关更多信息，请参阅[任务执行状态](#)。

### Important

如果您计划向 Amazon S3 地点传输数据或从 Amazon S3 站点传输数据，请在开始之前查看[DataSync如何影响您的 S3 请求费用](#)和[DataSync定价页面](#)。

## 请求语法

```
{
  "Excludes": [
    {
      "FilterType": "string",
      "Value": "string"
    }
  ],
  "Includes": [
    {
      "FilterType": "string",
      "Value": "string"
    }
  ],
  "OverrideOptions": {
    "Atime": "string",
    "BytesPerSecond": number,
    "Gid": "string",
    "LogLevel": "string",
    "Mtime": "string",
    "ObjectTags": "string",
    "OverwriteMode": "string",
    "PosixPermissions": "string",
    "PreserveDeletedFiles": "string",
    "PreserveDevices": "string",
    "SecurityDescriptorCopyFlags": "string",
    "TaskQueueing": "string",
    "TransferMode": "string",
    "Uid": "string",
  }
}
```

```
    "VerifyMode": "string"
  },
  "Tags": [
    {
      "Key": "string",
      "Value": "string"
    }
  ],
  "TaskArn": "string"
}
```

## 请求参数

有关所有操作的通用参数的信息，请参阅[通用参数](#)。

请求接受采用 JSON 格式的以下数据。

### Excludes

指定要从任务中排除哪些文件的筛选规则列表。该列表包含一个由要排除的模式组成的筛选字符串。模式由“|”（即竖线）分隔，例如“/folder1|/folder2”。

类型：[FilterRule](#) 对象数组

数组成员：最少 0 项。最多 1 项。

必需：否

### Includes

指定确定运行任务时要包含哪些文件的筛选规则列表。该模式应包含一个由要包含的模式组成的筛选字符串。模式由“|”（即竖线）分隔，例如“/folder1|/folder2”。

类型：[FilterRule](#) 对象数组

数组成员：最少 0 项。最多 1 项。

必需：否

### OverrideOptions

配置您的 AWS DataSync 任务设置。这些选项包括如何 DataSync 处理文件、对象及其关联的元数据。您还可以指定如何 DataSync 验证数据完整性、为任务设置带宽限制以及其他选项。

每个任务设置都有一个默认值。除非需要，否则 Options 在开始任务之前不必配置任何一项。

类型：[Options](#) 对象

必需：否

## [Tags](#)

指定要应用于代表任务执行的 Amazon 资源名称 (ARN) 的标签。

标签是键值对，帮助您管理、筛选和搜索DataSync资源。

类型：[TagListEntry](#) 对象数组

数组成员：最少 0 项。最多 50 项。

必需：否

## [TaskArn](#)

指定要启动的任务的 Amazon Resource Name (ARN)。

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\\-0-9]*:[0-9]{12}:task/task-[0-9a-f]{17}$`

必需：是

## 响应语法

```
{
  "TaskExecutionArn": "string"
}
```

## 响应元素

如果此操作成功，则该服务将会发送回 HTTP 200 响应。

服务以 JSON 格式返回的以下数据。

## [TaskExecutionArn](#)

运行任务执行的 ARN。



类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\\-0-9]*:[0-9]{12}:task/task-[0-9a-f]{17}/execution/exec-[0-9a-f]{17}$`

## 错误

有关所有操作常见错误的信息，请参阅[常见错误](#)。

### InternalException

当AWS DataSync服务中发生错误时，会引发此异常。

HTTP 状态代码：500

### InvalidRequestException

当客户端提交格式错误的请求时，会引发此异常。

HTTP 状态代码：400

## 示例

### 示例请求

以下示例使用指定任务的默认选项开始执行任务。

```
{
  "OverrideOptions": {
    "Atime": "BEST_EFFORT",
    "BytesPerSecond": 1000,
    "Gid": "NONE",
    "Mtime": "PRESERVE",
    "PosixPermissions": "PRESERVE",
    "PreserveDevices": "NONE",
    "PreserveDeletedFiles": "PRESERVE",
    "Uid": "NONE",
    "VerifyMode": "POINT_IN_TIME_CONSISTENT"
  },
  "TaskArn": "arn:aws:datasync:us-east-2:111222333444:task/task-08de6e6697796f026"
```

```
}
```

## 示例响应

此示例说明了的一种用法StartTaskExecution。

```
{  
  "TaskExecutionArn": "arn:aws:datsync:us-east-2:111222333444:task/  
task-08de6e6697796f026/execution/exec-04ce9d516d69bd52f"  
}
```

## 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [AWS 命令行界面](#)
- [适用于 .NET 的 AWS SDK](#)
- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [AWSSDK 用于 JavaScript](#)
- [适用于 PHP V3 的 AWS SDK](#)
- [适用于 Python 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)

## StopDiscoveryJob

停止正在运行的DataSync发现任务。

您可以随时停止发现任务。在计划结束之前停止的任务可能会为您提供一些有关本地存储系统资源的信息。要获取有关已停止作业的建议，必须使用该[GenerateRecommendations](#)操作。

### 请求语法

```
{  
  "DiscoveryJobArn": "string"  
}
```

### 请求参数

有关所有操作的通用参数的信息，请参阅[通用参数](#)。

请求接受采用 JSON 格式的以下数据。

#### [DiscoveryJobArn](#)

指定要停止发现任务的 Amazon Resource Name (ARN)。

类型：字符串

长度限制：最大长度为 256。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\n-0-9]+:[0-9]{12}:system/storage-system-[a-f0-9]{8}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{12}/job/discovery-job-[a-f0-9]{8}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{12}$`

必需：是

### 响应元素

如果此操作成功，则该服务会发送回带有空 HTTP 正文的 HTTP 200 响应。

### 错误

有关所有操作常见错误的信息，请参阅[常见错误](#)。

## InternalException

当AWS DataSync服务中发生错误时，会抛出此异常。

HTTP 状态代码：500

## InvalidRequestException

当客户端提交格式错误的请求时，会引发此异常。

HTTP 状态代码：400

## 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [AWS 命令行界面](#)
- [适用于 .NET 的 AWS SDK](#)
- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [AWS适用于 SDKJavaScript](#)
- [适用于 PHP V3 的 AWS SDK](#)
- [适用于 Python 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)

# TagResource

将标签应用于AWS资源。标签是键值对，帮助您管理、筛选和搜索资源。

其中包括AWS DataSync资源，例如位置、任务和任务执行。

## 请求语法

```
{
  "ResourceArn": "string",
  "Tags": [
    {
      "Key": "string",
      "Value": "string"
    }
  ]
}
```

## 请求参数

有关所有操作的通用参数的信息，请参阅[通用参数](#)。

请求接受采用 JSON 格式的以下数据。

### [ResourceArn](#)

指定要应用标签可帮助您的 Amazon Name ARN ame Name Name Name Name Name Name

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z \-0-9]+:[0-9]{12}:(agent|task|location|system)/((agent|task|loc)-[a-f0-9]{17}|storage-system-[a-f0-9]{8}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{12})/(/execution/exec-[a-f0-9]{17})?&`

必需：是

### [Tags](#)

指定要应用于资源的标签。

类型：[TagListEntry](#) 对象数组

数组成员：最少 0 项。最多 50 项。

必需：是

## 响应元素

如果此操作成功，则该服务会发送回带有空 HTTP 正文的 HTTP 200 响应。

## 错误

有关所有操作常见错误的信息，请参阅[常见错误](#)。

### InternalException

当AWS DataSync服务中发生错误时，会抛出此异常。

HTTP 状态代码：500

### InvalidRequestException

当客户端提交格式错误的请求时，会引发此异常。

HTTP 状态代码：400

## 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [AWS 命令行界面](#)
- [适用于 .NET 的 AWS SDK](#)
- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [AWS适用于 SDKJavaScript](#)
- [适用于 PHP V3 的 AWS SDK](#)
- [适用于 Python 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)



# UntagResource

从AWS资源中删除标签。

## 请求语法

```
{
  "Keys": [ "string" ],
  "ResourceArn": "string"
}
```

## 请求参数

有关所有操作的通用参数的信息，请参阅[常用参数](#)。

请求接受采用 JSON 格式的以下数据。

### Keys

指定要删除的标签中删除的密钥。

类型：字符串数组

数组成员：最少 1 项。最多 50 项。

长度限制：最小长度为 1。长度上限为 256。

模式：`^[a-zA-Z0-9\s+=. _:/-]+$`

必需：是

### ResourceArn

指定要从中删除标签的资源的 Amazon Resource Name ( ARN ) 。

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\n-0-9]+:[0-9]{12}:(agent|task|location|system)/((agent|task|loc)-[a-f0-9]{17}|storage-system-[a-f0-9]{8}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{12})/(/execution/exec-[a-f0-9]{17})?$`



必需：是

## 响应元素

如果此操作成功，则该服务会发送回带有空 HTTP 正文的 HTTP 200 响应。

## 错误

有关所有操作常见错误的信息，请参阅[常见错误](#)。

### InternalException

当AWS DataSync服务中发生错误时，会抛出此异常。

HTTP 状态代码：500

### InvalidRequestException

当客户端提交格式错误的请求时，会引发此异常。

HTTP 状态代码：400

## 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [AWS 命令行界面](#)
- [适用于 .NET 的 AWS SDK](#)
- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [AWS适用于JavaScript](#)
- [适用于 PHP V3 的 AWS SDK](#)
- [适用于 Python 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)

# UpdateAgent

更新代理的名称。

## 请求语法

```
{  
  "AgentArn": "string",  
  "Name": "string"  
}
```

## 请求参数

有关所有操作的通用参数的信息，请参阅[常用参数](#)。

请求接受采用 JSON 格式的以下数据。

### AgentArn

要更新的代理的 Amazon 资源名称 (ARN)。

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\-\0-9]+:[0-9]{12}:agent/agent-[0-9a-z]{17}$`

必需：是

### Name

要用于配置代理的 Agent。

类型：字符串

长度限制：最小长度为 1。长度上限为 256。

模式：`^[a-zA-Z0-9\s+=. _:@/-]+$`

必需：否

## 响应元素

如果此操作成功，则该服务会发送回带有空 HTTP 正文的 HTTP 200 响应。

## 错误

有关所有操作常见错误的信息，请参阅[常见错误](#)。

### InternalException

当AWS DataSync服务中发生错误时，会抛出此异常。

HTTP 状态代码：500

### InvalidRequestException

当客户端提交格式错误的请求时，会引发此异常。

HTTP 状态代码：400

## 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [AWS 命令行界面](#)
- [适用于 .NET 的 AWS SDK](#)
- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [AWS适用于JavaScript](#)
- [适用于 PHP V3 的 AWS SDK](#)
- [适用于 Python 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)

# UpdateDiscoveryJob

编辑DataSync发现任务配置。

## 请求语法

```
{  
  "CollectionDurationMinutes": number,  
  "DiscoveryJobArn": "string"  
}
```

## 请求参数

有关所有操作的通用参数的信息，请参阅[通用参数](#)。

请求接受采用 JSON 格式的以下数据。

### [CollectionDurationMinutes](#)

以分钟为单位指定您希望发现作业运行多长时间。（您不能将此参数设置为小于作业已经运行的分钟数。）

类型：整数

有效范围：最小值为 60。最大

必需：是

### [DiscoveryJobArn](#)

指定要更新发现任务的 Amazon Resource Name (ARN)。

类型：字符串

长度限制：最大长度为 256。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\n-0-9]+:[0-9]{12}:system/storage-system-[a-f0-9]{8}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{12}/job/discovery-job-[a-f0-9]{8}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{12}$`

必需：是

## 响应元素

如果此操作成功，则该服务会发送回带有空 HTTP 正文的 HTTP 200 响应。

## 错误

有关所有操作常见错误的信息，请参阅[常见错误](#)。

### InternalException

当AWS DataSync服务中发生错误时，会抛出此异常。

HTTP 状态代码：500

### InvalidRequestException

当客户端提交格式错误的请求时，会引发此异常。

HTTP 状态代码：400

## 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [AWS 命令行界面](#)
- [适用于 .NET 的 AWS SDK](#)
- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [AWS适用于 SDKJavaScript](#)
- [适用于 PHP V3 的 AWS SDK](#)
- [适用于 Python 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)

## UpdateLocationHdfs

更新先前为 Hadoop 分布式文件系统集群创建的位置的某些参数。

### 请求语法

```
{
  "AgentArns": [ "string" ],
  "AuthenticationType": "string",
  "BlockSize": number,
  "KerberosKeytab": blob,
  "KerberosKrb5Conf": blob,
  "KerberosPrincipal": "string",
  "KmsKeyProviderUri": "string",
  "LocationArn": "string",
  "NameNodes": [
    {
      "Hostname": "string",
      "Port": number
    }
  ],
  "QopConfiguration": {
    "DataTransferProtection": "string",
    "RpcProtection": "string"
  },
  "ReplicationFactor": number,
  "SimpleUser": "string",
  "Subdirectory": "string"
}
```

### 请求参数

有关所有操作的通用参数的信息，请参阅[通用参数](#)。

请求接受采用 JSON 格式的以下数据。

#### [AgentArns](#)

用于连接到 HDFS 集群的代理的 ARN。

类型：字符串数组

数组成员：最少 1 项。最多 4 项。

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\-\0-9]+:[0-9]{12}:agent/agent-[0-9a-z]{17}$`

必需：否

### AuthenticationType

用于确定用户身份的身份验证类型。

类型：字符串

有效值：SIMPLE | KERBEROS

必需：否

### BlockSize

要写入 HDFS 集群的数据块大小。

类型：整数

有效范围：有效范围：有效范围：有效范围：最小值为 1048576。最大值为 1073741824。

必需：否

### KerberosKeytab

Kerberos 密钥表 (keytab)，其中包含定义的 Kerberos 主体和加密密钥之间的映射。您可以通过提供文件的地址从文件加载密钥表。如果您使用 AWS CLI，它将执行 base64 编码。否则，请提供 base64 编码文本。

类型：Base64 编码的二进制数据对象

长度约束：长度约束：长度限制：长度约束：长度

必需：否

### KerberosKrb5Conf

包含 Kerberos 配置信息的 `krb5.conf` 文件。您可以通过提供 `krb5.conf` 文件的地址来加载文件。如果您使用的 AWS CLI，它将执行 base64 编码。否则，请提供 base64 编码文本。

类型：Base64 编码的二进制数据对象

长度约束：长度约束：长度限制：长度约束：长度

必需：否

### KerberosPrincipal

有权访问 HDFS 集群上的文件和文件夹的 Kerberos 主体。

类型：字符串

长度限制：最小长度为 1。长度上限为 256。

模式：`^.+`\$

必需：否

### KmsKeyProviderUri

HDFS 集群的密钥管理服务器 ( KMS ) 的 URI。

类型：字符串

长度限制：最小长度为 1。最大长度为 255。

模式：`^kms:\//http[s]?@((([a-zA-Z0-9-]*[a-zA-Z0-9])\.)*([A-Za-z0-9-]*[A-Za-z0-9]))(;\((([a-zA-Z0-9-]*[a-zA-Z0-9])\.)*([A-Za-z0-9-]*[A-Za-z0-9])))*: [0-9]{1,5}\//kms$`

必需：否

### LocationArn

源 HDFS 集群位置的 Amazon 资源名称 ( ARN )。

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\-\0-9]+:[0-9]{12}:location/loc-[0-9a-z]{17}$`

必需：是

### NameNodes

管理 HDFS 命名空间的。NameNodeNameNode 执行打开、关闭和重命名文件和目录等操作。NameNode 包含将数据块映射到 1 000 DataNodes。您只能使用一个 NameNode。



类型：[HdfsNameNode](#) 对象数组

数组成员：最少 1 项。

必需：否

### [QopConfiguration](#)

保护质量 ( QOP ) 配置指定在 Hadoop Distributed File System ( HDFS ) 集群上配置的远程程序调用 ( RPC ) 和数据传输隐私设置。

类型：[QopConfiguration](#) 对象

必需：否

### [ReplicationFactor](#)

写入 HDFS 集群时要将数据复制到的数量。DataNodes

类型：整数

有效范围：最小值为 1。最大值为 512。

必需：否

### [SimpleUser](#)

用于在主机操作系统上标识客户端的用户名。

类型：字符串

长度限制：最小长度为 1。长度上限为 256。

模式：`^[_.A-Za-z0-9][_.A-Za-z0-9]*$`

必需：否

### [Subdirectory](#)

HDFS 集群中的子目录。此子目录用于从 HDFS 集群读取数据或向其写入数据。

类型：字符串

长度约束：长度约束：长度约束：长度约束

模式：`^[a-zA-Z0-9_-\+\.\/(\\)\$\p{Zs}]+$`

必需：否

## 响应元素

如果此操作成功，则该服务会发送回带有空 HTTP 正文的 HTTP 200 响应。

## 错误

有关所有操作常见错误的信息，请参阅[常见错误](#)。

### InternalException

当AWS DataSync服务中发生错误时，会引发此异常。

HTTP 状态代码：500

### InvalidRequestException

当客户端提交格式错误的请求时，会引发此异常。

HTTP 状态代码：400

## 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [AWS 命令行界面](#)
- [适用于 .NET 的 AWS SDK](#)
- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [AWSSDK 用于 JavaScript](#)
- [适用于 PHP V3 的 AWS SDK](#)
- [适用于 Python 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)

## UpdateLocationNfs

更新先前创建的网络文件系统 (NFS) 访问位置的一些参数。有关创建 NFS 位置的信息，请参阅[为 NFS 创建位置](#)。

### 请求语法

```
{
  "LocationArn": "string",
  "MountOptions": {
    "Version": "string"
  },
  "OnPremConfig": {
    "AgentArns": [ "string" ]
  },
  "Subdirectory": "string"
}
```

### 请求参数

有关所有操作的通用参数的信息，请参阅[通用参数](#)。

请求接受采用 JSON 格式的以下数据。

#### [LocationArn](#)

要更新的 NFS 位置的 Amazon 资源名称 (ARN)。

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\\-0-9]+:[0-9]{12}:location/loc-[0-9a-z]{17}$`

必需：是

#### [MountOptions](#)

指定 DataSync 如何使用 NFS 协议访问某个位置。

类型：[NfsMountOptions](#) 对象

必需：否

## [OnPremConfig](#)

用于网络文件系统 (NFS) 位置的代理的 Amazon 资源名称 (ARN) 列表。

类型：[OnPremConfig](#) 对象

必需：否

## [Subdirectory](#)

NFS 文件系统中的子目录，用于从 NFS 源位置读取数据或将数据写入 NFS 目标。NFS 路径应该是由 NFS 服务器导出的路径，或者是该路径的一个子目录。该路径应使其能由网络中的其他 NFS 客户端挂载。

要查看 NFS 服务器导出的所有路径，请从可访问服务器的 NFS 客户端运行“`showmount -e nfs-server-name`”。您可以指定显示在结果中的任何目录以及该目录的任何子目录。确保无需 Kerberos 身份验证即可访问 NFS 导出。

要传输指定文件夹中的所有数据，DataSync 必须有权读取所有数据。为确保这一点，请为 NFS 导出配置 `no_root_squash`，或者确保需要 NFS 导出配置，或者 DataSync 确保需要 NFS 导出配置，或者确保需要 App。通过执行上述任一选项，将允许代理读取这些文件。要使代理能够访问目录，您还必须启用全部执行访问权限。

如果您要将数据复制到设备或从 AWS Snowcone 设备复制数据，请参阅 [NFS Server](#)。AWS Snowcone

有关 NFS 导出配置的信息，请参阅 [18.7。Red Hat Enterprise Linux 文档中的 /etc/exports 配置文件](#)。

类型：字符串

长度约束：最大长度为 4096。

模式：`^[a-zA-Z0-9_-\+\.\/(\\)\p{Zs}]+$`

必需：否

## 响应元素

如果此操作成功，则该服务会发送回带有空 HTTP 正文的 HTTP 200 响应。

## 错误

有关所有操作常见错误的信息，请参阅[常见错误](#)。

### InternalException

当AWS DataSync服务中发生错误时，会抛出此异常。

HTTP 状态代码：500

### InvalidRequestException

当客户端提交格式错误的请求时，会引发此异常。

HTTP 状态代码：400

## 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [AWS 命令行界面](#)
- [适用于 .NET 的 AWS SDK](#)
- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [AWS适用于 SDKJavaScript](#)
- [适用于 PHP V3 的 AWS SDK](#)
- [适用于 Python 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)

## UpdateLocationObjectStorage

更新用于传输的现有对象存储位置AWS DataSync的某些参数。有关创建自管对象存储位置的信息，请参阅[创建对象存储位置](#)。

### 请求语法

```
{
  "AccessKey": "string",
  "AgentArns": [ "string" ],
  "LocationArn": "string",
  "SecretKey": "string",
  "ServerCertificate": blob,
  "ServerPort": number,
  "ServerProtocol": "string",
  "Subdirectory": "string"
}
```

### 请求参数

有关所有操作的通用参数的信息，请参阅[通用参数](#)。

请求接受采用 JSON 格式的以下数据。

#### [AccessKey](#)

如果需要凭证来向对象存储服务器进行身份验证，则指定访问密钥（例如，用户名）。

类型：字符串

长度限制：最小长度为 1。最大长度为 200。

模式：`^.+`

必需：否

#### [AgentArns](#)

指定可以安全连接到您所在位置的 Amazon 资源名称（ARN）。DataSync

类型：字符串数组

数组成员：最少 1 项。Maximum number of 4 items.

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\-\0-9]+:[0-9]{12}:agent/agent-[0-9a-z]{17}$`

必需：否

### LocationArn

指定您正在更新的对象存储系统位置的 ARN。

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\-\0-9]+:[0-9]{12}:location/loc-[0-9a-z]{17}$`

必需：是

### SecretKey

如果需要凭证来向对象存储服务器进行身份验证，则指定私有密钥（例如，密码）。

类型：字符串

长度限制：最小长度为 1。最大长度为 200。

模式：`^.+`

必需：否

### ServerCertificate

指定要使用私有或自签名证书颁发机构 (CA) 的对象存储系统进行身份验证的证书。您必须指定一个 Base64 编码的 .pem 文件（例如，`file:///home/user/.ssh/storage_sys_certificate.pem`）。证书最多可以是 32768 字节（在 Base64 编码之前）。

要使用此参数，请配置 `ServerProtocol` 为 HTTPS。

更新证书不会干扰您正在进行的任务。

类型：Base64 编码的二进制数据对象

为 324。

必需：否

### ServerPort

指定对象存储服务器接受入站网络流量的端口（例如，端口 443）。

类型：整数

有效范围：最小值为 1。Maximum value of 65536.

必需：否

### ServerProtocol

指定对象存储服务器通信时使用的协议。

类型：字符串

有效值：HTTPS | HTTP

必需：否

### Subdirectory

指定对象存储服务器的对象前缀。如果这是源位置，则DataSync仅复制具有此前缀的对象。如果这是目标位置，则DataSync写入具有此前缀的所有对象。

类型：字符串

为 400。

模式：`^[a-zA-Z0-9_\-\.\/\(\)\p{Zs}]*$`

必需：否

## 响应元素

如果此操作成功，则该服务会发送回带有空 HTTP 正文的 HTTP 200 响应。

## 错误

有关所有操作常见错误的信息，请参阅[常见错误](#)。

## InternalException

当AWS DataSync服务中发生错误时，会引发此异常。



HTTP 状态代码：500

InvalidRequestException

当客户端提交格式错误的请求时，会引发此异常。

HTTP 状态代码：400

## 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [AWS 命令行界面](#)
- [适用于 .NET 的 AWS SDK](#)
- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [AWSSDK 用于 JavaScript](#)
- [适用于 PHP V3 的 AWS SDK](#)
- [适用于 Python 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)

## UpdateLocationSmb

更新先前为服务器消息块 (SMB) 文件系统访问权限的部分。有关创建 SMB 位置的信息，请参阅[为 SMB 创建位置](#)。

### 请求语法

```
{
  "AgentArns": [ "string" ],
  "Domain": "string",
  "LocationArn": "string",
  "MountOptions": {
    "Version": "string"
  },
  "Password": "string",
  "Subdirectory": "string",
  "User": "string"
}
```

### 请求参数

有关所有操作的通用参数的信息，请参阅[常用参数](#)。

请求接受采用 JSON 格式的以下数据。

#### [AgentArns](#)

用于 Simple Message Block (SMB) 位置的代理的 Amazon 资源名称 (ARN)。

类型：字符串数组

数组成员：最少 1 项。最。。

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\-0-9]+:[0-9]{12}:agent/agent-[0-9a-z]{17}$`

必需：否

#### [Domain](#)

SMB 服务器所属的 Windows 域的名称。

类型：字符串

长约束：长约束：最大约束：最大约束：最大

模式：`^[A-Za-z0-9](\.|-+)?[A-Za-z0-9]{0,252}$`

必需：否

### LocationArn

要更新的SMB位置的 Amazon 资源名称 (ARN)。

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\-0-9]+:[0-9]{12}:location/loc-[0-9a-z]{17}$`

必需：是

### MountOptions

指定 AWS DataSync 用于访问服务器消息块 (SMB) 文件服务器的 SMB 协议版本。

类型：[SmbMountOptions](#) 对象

必需：否

### Password

可以挂载共享并对 SMB 共享共享中的文件和文件夹具有访问权限的用户的密码。

类型：字符串

长约束：长约束：最大约束：最大约束：最大

模式：`^.{0,104}$`

必需：否

### Subdirectory

SMB 文件系统中用于从 SMB 源位置读取数据或将数据写入 SMB 目标的子目录。SMB 路径应该是 SMB 服务器导出的路径，或者是该路径的一个子目录。此路径应使其能由网络中的其他 SMB 客户端挂载。

**Note**

必须使用正斜杠指定 Subdirectory。例如，/path/to/folder。

要传输指定文件夹中的数据，DataSync 必须具有挂载 SMB 共享和访问该共享中数据的权限。为确保这一点，请执行以下任一操作：

- 确保指定以下用户的用户/密码：可以挂载共享的用户、对您要 DataSync 访问的所有文件和目录具有适当权限的用户。
- 使用 Backup 操作员组成员的凭据来装载共享。

执行任一操作都可以使代理访问数据。要使代理访问目录，您还必须启用所有执行访问权限。

类型：字符串

长约束：长约束：最大约束：最大长约约束：最大

模式：`^[a-zA-Z0-9_-\+\.\(\)\$\p{Zs}]+$`

必需：否

**User**

可以挂载共享并对 SMB 共享中的文件和文件夹具有访问权限的用户。

类型：字符串

长约束：长约束：最大约束：最大约束：最大

模式：`^[^\x5B\x5D\\\/;|=, +*?]{1,104}$`

必需：否

**响应元素**

如果此操作成功，则该服务会发送回带有空 HTTP 正文的 HTTP 200 响应。

**错误**

有关所有操作常见错误的信息，请参阅[常见错误](#)。

## InternalException

当AWS DataSync服务中发生错误时，会抛出此异常。

HTTP 状态代码：500

## InvalidRequestException

当客户端提交格式错误的请求时，会引发此异常。

HTTP 状态代码：400

## 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [AWS 命令行界面](#)
- [适用于 .NET 的 AWS SDK](#)
- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [AWS适用于JavaScript](#)
- [适用于 PHP V3 的 AWS SDK](#)
- [适用于 Python 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)

## UpdateStorageSystem

修改您在DataSync Discovery 中使用的本地存储系统资源的某些配置。

### 请求语法

```
{
  "AgentArns": [ "string" ],
  "CloudWatchLogGroupArn": "string",
  "Credentials": {
    "Password": "string",
    "Username": "string"
  },
  "Name": "string",
  "ServerConfiguration": {
    "ServerHostname": "string",
    "ServerPort": number
  },
  "StorageSystemArn": "string"
}
```

### 请求参数

有关所有操作的通用参数的信息，请参阅[常用参数](#)。

请求接受采用 JSON 格式的以下数据。

#### [AgentArns](#)

指定用于连接和读本地存储系统的DataSync代理的 Amazon Rame ( ARN ) 。

类型：字符串数组

数组成员：固定数量为 1 项。

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\n-0-9]+:[0-9]{12}:agent/agent-[0-9a-z]{17}$`

必需：否

## [CloudWatchLogGroupArn](#)

指定用于监控和记录发现任务事件的 AmazonCloudWatch 日志组的 ARN。

类型：字符串

长度限制：最大长度为 562。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):logs:[a-z\-\0-9]+:[0-9]{12}:log-group:([^\:\*]*)(:\*)?$`

必需：否

## [Credentials](#)

指定用于访问本地存储系统的管理界面的用户名和密码。

类型：[Credentials](#) 对象

必需：否

## [Name](#)

为您的本地存储系统指定一个熟悉的名称。

类型：字符串

长度限制：最小长度为 1。长度上限为 256。

模式：`^[\\p{L}\\p{M}\\p{N}\\s+=. _:@\\/-]+`

必需：否

## [ServerConfiguration](#)

指定连接本地存储系统的管理接口所需的服务器名称和网络端口。

类型：[DiscoveryServerConfiguration](#) 对象

必需：否

## [StorageSystemArn](#)

指定要重新配置的本地存储系统的 ARN。

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\\-0-9]+:[0-9]{12}:system/storage-system-[a-f0-9]{8}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{12}$`

必需：是

## 响应元素

如果此操作成功，则该服务会发送回带有空 HTTP 正文的 HTTP 200 响应。

## 错误

有关所有操作常见错误的信息，请参阅[常见错误](#)。

### InternalException

当AWS DataSync服务中发生错误时，会抛出此异常。

HTTP 状态代码：500

### InvalidRequestException

当客户端提交格式错误的请求时，会引发此异常。

HTTP 状态代码：400

## 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [AWS 命令行界面](#)
- [适用于 .NET 的 AWS SDK](#)
- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [AWS适用于JavaScript](#)
- [适用于 PHP V3 的 AWS SDK](#)



- [适用于 Python 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)

# UpdateTask

更新与名称。

请求语法

```
{
  "CloudWatchLogGroupArn": "string",
  "Excludes": [
    {
      "FilterType": "string",
      "Value": "string"
    }
  ],
  "Includes": [
    {
      "FilterType": "string",
      "Value": "string"
    }
  ],
  "Name": "string",
  "Options": {
    "Atime": "string",
    "BytesPerSecond": number,
    "Gid": "string",
    "LogLevel": "string",
    "Mtime": "string",
    "ObjectTags": "string",
    "OverwriteMode": "string",
    "PosixPermissions": "string",
    "PreserveDeletedFiles": "string",
    "PreserveDevices": "string",
    "SecurityDescriptorCopyFlags": "string",
    "TaskQueueing": "string",
    "TransferMode": "string",
    "Uid": "string",
    "VerifyMode": "string"
  },
  "Schedule": {
    "ScheduleExpression": "string"
  },
  "TaskArn": "string"
}
```

## 请求参数

有关所有操作的通用参数的信息，请参阅[通用参数](#)。

请求接受采用 JSON 格式的以下数据。

### [CloudWatchLogGroupArn](#)

Amazon 资源名称 ( ARNCloudWatch ) 。

类型：字符串

长度约束：最大长度为 562。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):logs:[a-z\-\0-9]+:[0-9]{12}:log-group:([^\:]*)(:\:)*$`

必需：否

### [Excludes](#)

指定在传输过程中排除特定数据的筛选规则列表。有关更多信息和示例，请参阅 [Filtering by transferredDataSync by](#)

类型：[FilterRule](#) 对象数组

数组成员：最少 0 项。最多 1 项。

必需：否

### [Includes](#)

指定在传输期间包含特定数据的筛选规则列表。有关更多信息和示例，请参阅 [Filtering by transferredDataSync by](#)

类型：[FilterRule](#) 对象数组

数组成员：最少 0 项。最多 1 项。

必需：否

### [Name](#)

要更新的名称。

类型：字符串

长度限制：最小长度为 1。长度上限为 256。

模式：`^[a-zA-Z0-9\s+.=_:@/-]+$`

必需：否

## Options

配置您的AWS DataSync任务设置。这些选项包括如何DataSync处理文件、对象及其关联的元数据。您还可以指定如何DataSync验证数据完整性、为任务设置带宽限制以及其他选项。

每个设置都有默认值。除非需要，否则Options在开始任务之前不必对其进行任何配置。

类型：[Options](#) 对象

必需：否

## Schedule

指定用于定期将文件从源位置传输到目标位置的时间表。您可以将任务配置为每小时、每天、每周或每周的特定日期执行。您可以控制要在一天或一小时内执行任务的时间。您指定的时间是 UTC 时间。有关更多信息，请参阅[计划任务](#)。

类型：[TaskSchedule](#) 对象

必需：否

## TaskArn

要更新的资源名称 ( ARN )。

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\-\0-9]*:[0-9]{12}:task/task-[0-9a-f]{17}$`

必需：是

## 响应元素

如果此操作成功，则该服务会发送回带有空 HTTP 正文的 HTTP 200 响应。

## 错误

有关所有操作常见错误的信息，请参阅[常见错误](#)。

### InternalException

当AWS DataSync服务中发生错误时，会抛出此异常。

HTTP 状态代码：500

### InvalidRequestException

当客户端提交格式错误的请求时，会引发此异常。

HTTP 状态代码：400

## 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [AWS 命令行界面](#)
- [适用于 .NET 的 AWS SDK](#)
- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [AWS适用于 SDKJavaScript](#)
- [适用于 PHP V3 的 AWS SDK](#)
- [适用于 Python 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)

## UpdateTaskExecution

修改正在运行的AWS DataSync任务。

### Note

目前，您唯一可以修改OptionUpdateTaskExecution的是 [BytesPerSecond](#)，它会限制正在运行或排队的任务的带宽。

### 请求语法

```
{
  "Options": {
    "Atime": "string",
    "BytesPerSecond": number,
    "Gid": "string",
    "LogLevel": "string",
    "Mtime": "string",
    "ObjectTags": "string",
    "OverwriteMode": "string",
    "PosixPermissions": "string",
    "PreserveDeletedFiles": "string",
    "PreserveDevices": "string",
    "SecurityDescriptorCopyFlags": "string",
    "TaskQueueing": "string",
    "TransferMode": "string",
    "Uid": "string",
    "VerifyMode": "string"
  },
  "TaskExecutionArn": "string"
}
```

### 请求参数

有关所有操作的通用参数的信息，请参阅[常用参数](#)。

请求接受采用JSON格式的以下数据。

## Options

配置您的AWS DataSync任务设置。这些选项包括如何DataSync处理文件、对象及其关联的元数据。您还可以指定如何DataSync验证数据完整性、为任务设置带宽限制以及其他选项。

每个任务设置，具有默认值。除非需要，否则Options在开始任务之前不必对其进行任何配置。

类型：[Options](#) 对象

必需：是

## TaskExecutionArn

指定要更新的任务执行的 Amazon 资源名称 ( ARN ) 。

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\n-0-9]*:[0-9]{12}:task/task-[0-9a-f]{17}/execution/exec-[0-9a-f]{17}$`

必需：是

## 响应元素

如果此操作成功，则该服务会发送回带有空 HTTP 正文的 HTTP 200 响应。

## 错误

有关所有操作常见错误的信息，请参阅[常见错误](#)。

### InternalException

当AWS DataSync服务中发生错误时，会抛出此异常。

HTTP 状态代码：500

### InvalidRequestException

当客户端提交格式错误的请求时，会引发此异常。

HTTP 状态代码：400

## 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [AWS 命令行界面](#)
- [适用于 .NET 的 AWS SDK](#)
- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [AWS适用于JavaScript](#)
- [适用于 PHP V3 的 AWS SDK](#)
- [适用于 Python 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)

## 数据类型

支持以下数据类型：

- [AgentListEntry](#)
- [Capacity](#)
- [Credentials](#)
- [DiscoveryJobListEntry](#)
- [DiscoveryServerConfiguration](#)
- [Ec2Config](#)
- [FilterRule](#)
- [FsxProtocol](#)
- [FsxProtocolNfs](#)
- [FsxProtocolSmb](#)
- [HdfsNameNode](#)
- [IOPS](#)
- [Latency](#)
- [LocationFilter](#)



- [LocationListEntry](#)
- [MaxP95Performance](#)
- [NetAppONTAPCluster](#)
- [NetAppONTAPSVM](#)
- [NetAppONTAPVolume](#)
- [NfsMountOptions](#)
- [OnPremConfig](#)
- [Options](#)
- [P95Metrics](#)
- [PrivateLinkConfig](#)
- [QopConfiguration](#)
- [Recommendation](#)
- [ResourceDetails](#)
- [ResourceMetrics](#)
- [S3Config](#)
- [SmbMountOptions](#)
- [StorageSystemListEntry](#)
- [TagListEntry](#)
- [TaskExecutionListEntry](#)
- [TaskExecutionResultDetail](#)
- [TaskFilter](#)
- [TaskListEntry](#)
- [TaskSchedule](#)
- [Throughput](#)

## AgentListEntry

表示调用[ListAgents](#)操作时AWS DataSync代理列表（或数组）中的单个条目。

### 目录

#### AgentArn

DataSync代理的 Amazon 资源名称 (ARN)。

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\-\0-9]+:[0-9]{12}:agent/agent-[0-9a-z]{17}$`

必需：否

#### Name

代理的名称。

类型：字符串

长度限制：最小长度为 1。长度上限为 256。

模式：`^[a-zA-Z0-9\s+=. _:@/-]+$`

必需：否

#### Status

代理的状态。有关更多信息，请参阅[DataSync代理的状态](#)。

类型：字符串

有效值：ONLINE | OFFLINE

必需：否

### 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)

# Capacity

本地存储系统资源（例如，卷）的存储容量。

## 目录

### LogicalUsed

在不考虑压缩或重复数据删除的情况下，存储系统资源中使用的空间量。

类型：长整型

有效范围：最小值为 0。

必需：否

### Provisioned

存储系统资源中可用的空间总量。

类型：长整型

有效范围：最小值为 0。

必需：否

### Used

存储系统资源中使用的空间量。

类型：长整型

有效范围：最小值为 0。

必需：否

## 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)

- [适用于 Ruby V3 的 AWS SDK](#)

# Credentials

为本地存储系统管理界面提供DataSync Discovery 读取访问权限的证书。

DataSyncDiscovery 将这些凭证存储在[AWS Secrets Manager](#)。有关更多信息，请参阅[访问您的本地存储系统](#)。

## 目录

### Password

为存储系统的管理接口指定密码。

类型：字符串

长度约束：最大长度为 1024。

模式：`^(?!.*[:\"])[^:"]*$).+$`

必需：是

### Username

为存储系统的管理界面指定用户名。

类型：字符串

长度约束：最大长度为 1024。

模式：`^(?!.*[:\"])[^:"]*$).+$`

必需：是

## 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)



# DiscoveryJobListEntry

任务DataSync的。

## 目录

### DiscoveryJobArn

发现作业的 Amazon 资源名称 ( ARN ) 。

类型：字符串

长度限制：最大长度为 256。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\n-0-9]+:[0-9]{12}:system/storage-system-[a-f0-9]{8}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{12}/job/discovery-job-[a-f0-9]{8}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{12}$`

必需：否

### Status

发现作业的。Resour [ce 的。](#)

类型：字符串

有效值：RUNNING | WARNING | TERMINATED | FAILED | STOPPED | COMPLETED | COMPLETED\_WITH\_ISSUES

必需：否

## 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)



# DiscoveryServerConfiguration

DataSyncDiscovery 用于连接本地存储系统的管理接口的网络设置。

## 目录

### ServerHostname

存储系统管理接口的域名或 IP 地址。

类型：字符串

长度限制：最大长度为 255。

模式：`^(([a-zA-Z0-9\-]*[a-zA-Z0-9])\.)*([A-Za-z0-9\-]*[A-Za-z0-9])$`

必需：是

### ServerPort

用于访问存储系统管理接口的网络端口。

类型：整数

有效范围：最小值为 1。最大值为 65535。

必需：否

## 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)

## Ec2Config

AWS DataSync 用于访问 Amazon EFS 文件系统的子网和安全组。

### 目录

#### SecurityGroupArns

指定与 Amazon EFS 文件系统的挂载目标关联的安全组的 Amazon 资源名称 (ARN)。

类型：字符串数组

数组成员：最少 1 项。最多 5 项。

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):ec2:[a-z\-\0-9]*:[0-9]{12}:security-group/sg-[a-f0-9]+$`

必需：是

#### SubnetArn

指定[网络接口用于管理传输过程中的流量的子网](#)的 ARN。DataSync

子网必须位于：

- 与 Amazon EFS 文件系统位于同一虚拟私有云 (VPC)。
- 至少与 Amazon EFS 文件系统的一个挂载目标位于同一个可用区。

#### Note

您无需指定包含文件系统挂载目标的子网。

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):ec2:[a-z\-\0-9]*:[0-9]{12}:subnet/.*$`

必需：是

## 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)

## FilterRule

指定在将文件从源传输到目标时要包括或排除哪些文件、文件夹和对象。

### 目录

#### FilterType

要应用的筛选规则的类型。AWS DataSync 仅支持 SIMPLE\_PATTERN 规则类型。

类型：字符串

长度约束：最大长度为 128。

模式：`^[A-Z0-9_]+$`

有效值：SIMPLE\_PATTERN

必需：否

#### Value

由要包含或排除的模式组成的单个筛选条件字符串。模式用 "|" ( 即管道 ) 分隔，例如：`/folder1|/folder2`

类型：字符串

长度约束：最大长度为 102400。

模式：`^[^\x00]+$`

必需：否

### 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)



## FsxProtocol

指定 AWS DataSync 用于访问您的 Amazon FSx 文件系统的数据传输协议。

### 目录

#### NFS

指定用于访问 FSx for ONTAP 文件系统的存储虚拟机 ( SVM ) 的网络文件系统 ( NFS ) 协议配置。DataSync

类型 : [FsxProtocolNfs](#) 对象

必需 : 否

#### SMB

指定用于访问 FSx for ONTAP 文件系统 SVM 的服务器消息块 ( SMB ) 协议配置。DataSync

类型 : [FsxProtocolSmb](#) 对象

必需 : 否

### 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)

# FsxProtocolNfs

指定用于访问您的 Amazon FSx for NetApp ONTAP 文件系统的网络文件系统 ( NFSx S ) 协议配置。AWS DataSync

## 目录

### MountOptions

指定 DataSync 如何使用 NFS 协议访问某个位置。

类型：[NfsMountOptions](#) 对象

必需：否

## 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)

# FsxProtocolSmb

指定用于访问 Amazon FSx for NetApp ONTAP 文件系统的服务器消息块 (SMB) 协议配置。AWS DataSync 有关更多信息，请参阅 [访问 FSx for ONTAP 文件系统](#)。

## 目录

### Password

指定有权访问您的 SVM 的用户的密码。

类型：字符串

长度为 104。

模式：`^[0,104]$`

必需：是

### User

指定一个用户名，该用户名可以挂载位置并访问 SVM 中所需的文件、文件夹和元数据。

如果您在 Active Directory 中提供用户，请注意以下事项：

- 如果您使用 AWS Directory Service for Microsoft Active Directory，则用户必须是 AWS 委托 FSx 管理员组的成员。
- 如果您使用的是自行管理的 Active Directory，则用户必须是域管理员组或在创建文件系统时为文件系统管理指定的自定义组的成员。

确保用户具有复制所需数据所需的权限：

- SE\_TCB\_NAME：设置对象所有权和文件元数据所必需的。使用此权限，您还可以复制 NTFS 自由访问列表 (DACL)。
- SE\_SECURITY\_NAME：可能需要复制 NTFS 系统访问控制列表 (SACL)。此操作特别需要 Windows 权限，该权限授予域管理员组的成员。如果将任务配置为复制 SACL，请确保用户具有所需的权限。有关复制 SACL 的信息，请参阅 [所有权和权限相关的选项](#)。

类型：字符串

长度为 104。

模式：`^[^x5B\x5D\\\/:;|=,+*?]{1,104}$`



必需：是

## Domain

指定存储虚拟机 ( SVM ) 所属的 Microsoft Active Directory 的完全限定域名 ( FQDN ) 。

类型：字符串

长度为 253。

模式：`^[A-Za-z0-9](\.|-+)?[A-Za-z0-9]{0,252}$`

必需：否

## MountOptions

指定 AWS DataSync 用于访问服务器消息块 ( SMB ) 文件服务器的 SMB 协议版本。

类型：[SmbMountOptions](#) 对象

必需：否

## 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)

## HdfsNameNode

Hadoop Distributed File System ( HDFS ) NameNode NameNode 管理文件系统的命名空间。NameNode 执行打开、关闭和重命名文件和目录等操作。NameNode 包含将数据块映射到的信息 DataNodes。

### 目录

#### Hostname

HDFS 集群 NameNode 中 HDFS 集群中的主机名。它是 IP 地址或域名服务 (DNS) 名称 NameNode。在本地安装的代理使用该主机名与网络 NameNode 中的服务器通信。

类型：字符串

长度限制：最小长度为 1。最大长度为 255。

模式：`^(([a-zA-Z0-9\-\_]*[a-zA-Z0-9])\.)*([A-Za-z0-9\-\_]*[A-Za-z0-9])$`

必需：是

#### Port

NameNode 用来侦听客户端请求的端口。

类型：整数

有效范围：最小值为 1。最大值为 50。

必需：是

### 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)

# IOPS

本地存储系统资源的 IOPS 峰值。每个数据点表示 1 小时间隔内的第 95 个百分位峰值。

## 目录

### Other

峰值 IOPS 与读取和写入操作无关。

类型：双精度

有效范围：最小值为 0。

必需：否

### Read

与读取操作相关的峰值 IOPS。

类型：双精度

有效范围：最小值为 0。

必需：否

### Total

本地存储系统资源的峰值总 IOPS。

类型：双精度

有效范围：最小值为 0。

必需：否

### Write

与写入操作相关的峰值 IOPS。

类型：双精度

有效范围：最小值为 0。

必需：否

## 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)

# Latency

本地存储系统资源的延迟峰值。每个数据点表示 1 小时间隔内的第 95 个百分位峰值。

## 目录

### Other

与读取和写入操作无关的操作的峰值延迟。

类型：双精度

有效范围：最小值为 0。

必需：否

### Read

读取操作的峰值延迟。

类型：双精度

有效范围：最小值为 0。

必需：否

### Write

写操作的峰值延迟。

类型：双精度

有效范围：最小值为 0。

必需：否

## 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)

- [适用于 Ruby V3 的 AWS SDK](#)

## LocationFilter

缩小返回的资源列表的范围ListLocations。例如，要查看您的所有 Amazon S3 位置，请使用 "Name": "LocationType""Operator": "Equals"、和创建筛选条件 "Values": "S3"。

有关更多信息，请参阅[器的名称](#)。

### 目录

#### Name

正在使用的名称。每个 API 调用都支持可用的过滤器列表 ( LocationType例如，用于ListLocations )。

类型：字符串

有效值：LocationUri | LocationType | CreationTime

必需：是

#### Operator

用于比较筛选值的运算符 ( 例如，Equals或Contains )。

类型：字符串

有效值：Equals | NotEquals | In | LessThanOrEqual | LessThan | GreaterThanOrEqual | GreaterThan | Contains | NotContains | BeginsWith

必需：是

#### Values

您可以制作器的名称。例如，您可以执行以下Amazon S3。

类型：字符串数组

长度限制：最小长度为 1。最大长度为 255。

模式：`^[0-9a-zA-Z_ \ -:\*\.\ \\/\?-\]*$`

必需：是

## 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)



## LocationListEntry

表示位置列表中的一个条目。LocationListEntry返回一个包含调用[ListLocations](#)操作时位置列表的数组。

### 目录

#### LocationArn

Outpost 的 Amazon 资源名称 ( ARN )。对于网络文件系统 (NFS) 或 Amazon EFS，位置是导出路径。对于 Amazon S3，位置是您要安装并用作该位置根目录的前缀路径。

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\-0-9]+:[0-9]{12}:location/loc-[0-9a-z]{17}$`

必需：否

#### LocationUri

表示某个位置的 URI 列表。LocationUri返回一个包含调用[ListLocations](#)操作时位置列表的数组。

格式：`TYPE://GLOBAL_ID/SUBDIR`。

TYPE 指定位置的类型 ( 例如，`nfs`或`s3` )。

GLOBAL\_ID 是支持该位置的资源的全局唯一标识符。EFS 的一个例子是`us-east-2.fs-abcd1234`。Amazon S3 的一个示例是存储桶名称，例如`myBucket`。NFS 的一个示例是有效的 IPv4 地址或符合域名服务 (DNS) 的主机名。

SUBDIR 是有效的文件系统路径，与\*nix 约定一样，由正斜杠分隔。对于 NFS 和 Amazon EFS，它是安装该位置的导出路径。对于 Amazon S3，它是您装载到的前缀路径，并将其视为该位置的根目录。

类型：字符串

长度约束：最大长度为 4360。

模式：`^(efs|nfs|s3|smb|hdfs|fsx[a-z0-9-]+)://[a-zA-Z0-9.:/\-]+$`

必需：否

## 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)

# MaxP95Performance

DataSyncDiscovery 收集的有关本地存储系统资源的性能数据。

## 目录

### lopsOther

与读取和写入操作无关的峰值 IOPS。

类型：双精度

有效范围：最小值为 0。

必需：否

### lopsRead

与读取操作相关的峰值 IOPS。

类型：双精度

有效范围：最小值为 0。

必需：否

### lopsTotal

本地存储系统资源的峰值总 IOPS。

类型：双精度

有效范围：最小值为 0。

必需：否

### lopsWrite

与写入操作相关的峰值 IOPS。

类型：双精度

有效范围：最小值为 0。

必需：否

## LatencyOther

与读取和写入操作无关的操作的峰值延迟。

类型：双精度

有效范围：最小值为 0。

必需：否

## LatencyRead

读取操作的峰值延迟。

类型：双精度

有效范围：最小值为 0。

必需：否

## LatencyWrite

写操作的峰值延迟。

类型：双精度

有效范围：最小值为 0。

必需：否

## ThroughputOther

与读取和写入操作无关的峰值吞吐量。

类型：双精度

有效范围：最小值为 0。

必需：否

## ThroughputRead

与读取操作相关的峰值吞吐量。

类型：双精度

有效范围：最小值为 0。

必需：否

## ThroughputTotal

本地存储系统资源的峰值总吞吐量。

类型：双精度

有效范围：最小值为 0。

必需：否

## ThroughputWrite

与写入操作相关的峰值吞吐量。

类型：双精度

有效范围：最小值为 0。

必需：否

## 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)

# NetAppONTAPCluster

DataSyncDiscovery 收集的有关本地存储系统集群的信息。

## 目录

### CifsShareCount

集群中的 CIFS 共享数。

类型：长整型

有效范围：最小值为 0。

必需：否

### ClusterBlockStorageLogicalUsed

集群中使用的存储空间，不考虑压缩或去重。

类型：长整型

有效范围：最小值为 0。

必需：否

### ClusterBlockStorageSize

集群中可用的总存储空间。

类型：长整型

有效范围：最小值为 0。

必需：否

### ClusterBlockStorageUsed

集群中使用的存储空间。

类型：长整型

有效范围：最小值为 0。

必需：否

## ClusterName

集群的名称。

类型：字符串

长度约束：最大长度为 1024。

模式：`^\{0,1024\}$`

必需：否

## MaxP95Performance

DataSyncDiscovery 收集的有关集群的性能数据。

类型：[MaxP95Performance](#) 对象

必需：否

## NfsExportedVolumes

集群中 NFS 卷的数量。

类型：长整型

有效范围：最小值为 0。

必需：否

## Recommendations

DataSyncDiscovery 为集群推荐的AWS存储服务。有关更多信息，请参阅[DataSync发现提供的建议](#)。

类型：[Recommendation](#) 对象数组

必需：否

## RecommendationStatus

表示集群的DataSync发现建议是已准备就绪、不完整还是无法确定。

有关更多信息，请参阅[推荐状态](#)。

类型：字符串

有效值：NONE | IN\_PROGRESS | COMPLETED | FAILED

必需：否

### ResourceId

集群的通用唯一标识符 (UUID)。

类型：字符串

模式：`[a-f0-9]{8}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{12}`

必需：否

### 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)



# NetAppONTAPSVM

DataSyncDiscovery 收集的有关本地存储系统中存储虚拟机 (SVM) 的信息。

## 目录

### CifsShareCount

SVM 中 CIFS 共享的数量。

类型：长整型

有效范围：最小值为 0。

必需：否

### ClusterUuid

与 SVM 关联的集群的通用唯一标识符 (UUID)。

类型：字符串

模式：`[a-f0-9]{8}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{12}`

必需：否

### EnabledProtocols

为 SVM 配置的数据传输协议 (例如 NFS)。

类型：字符串数组

长度约束：最大长度为 1024。

模式：`^.{0,1024}$`

必需：否

### MaxP95Performance

DataSyncDiscovery 收集的有关 SVM 的性能数据。

类型：[MaxP95Performance](#) 对象

必需：否

## NfsExportedVolumes

SVM 中的 NFS 卷数。

类型：长整型

有效范围：最小值为 0。

必需：否

## Recommendations

DataSyncDiscovery 为 SVM 推荐的 AWS 存储服务。有关更多信息，请参阅 [DataSyncDiscovery 提供的建议](#)。

类型：[Recommendation](#) 对象数组

必需：否

## RecommendationStatus

指示 SVM 的 DataSync 发现建议是已准备好查看、不完整还是无法确定。

有关更多信息，请参阅[推荐状态](#)。

类型：字符串

有效值：NONE | IN\_PROGRESS | COMPLETED | FAILED

必需：否

## ResourceId

SVM 的 UID。

类型：字符串

模式：`[a-f0-9]{8}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{12}`

必需：否

## SvmName

SVM 名称

类型：字符串

长度约束：最大长度为 1024。

模式：`^\{0,1024\}$`

必需：否

#### TotalCapacityProvisioned

SVM 中可用的总存储空间。

类型：长整型

有效范围：最小值为 0。

必需：否

#### TotalCapacityUsed

SVM 中使用的存储空间。

类型：长整型

有效范围：最小值为 0。

必需：否

#### TotalLogicalCapacityUsed

SVM 中使用的存储空间，不考虑压缩或去重。

类型：长整型

有效范围：最小值为 0。

必需：否

#### TotalSnapshotCapacityUsed

SVM 中用于快照的存储量。

类型：长整型

有效范围：最小值为 0。

必需：否

## 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)

# NetAppONTAPVolume

DataSyncDiscovery 收集的有关本地存储系统中某个卷的信息。

## 目录

### CapacityProvisioned

卷中可用的总存储空间。

类型：长整型

有效范围：最小值为 0。

必需：否

### CapacityUsed

卷中使用的存储空间。

类型：长整型

有效范围：最小值为 0。

必需：否

### CifsShareCount

卷中 CIFS 份额的数量。

类型：长整型

有效范围：最小值为 0。

必需：否

### LogicalCapacityUsed

卷中使用的存储空间，不考虑压缩或重复数据删除。

类型：长整型

有效范围：最小值为 0。

必需：否

## MaxP95Performance

DataSyncDiscovery 收集的有关该卷的性能数据。

类型：[MaxP95Performance](#) 对象

必需：否

## NfsExported

卷中 NFS 卷的数量。

类型：布尔值

必需：否

## Recommendations

DataSyncDiscovery 为该卷推荐的AWS存储服务。有关更多信息，请参阅 [DataSyncDiscovery 提供的建议](#)。

类型：[Recommendation](#) 对象数组

必需：否

## RecommendationStatus

指明该卷的 DataSync Discovery 推荐是否已准备好查看、不完整还是无法确定。

有关更多信息，请参阅[推荐状态](#)。

类型：字符串

有效值：NONE | IN\_PROGRESS | COMPLETED | FAILED

必需：否

## ResourceId

卷的通用唯一标识符 ( UUID ) 。

类型：字符串

模式：`[a-f0-9]{8}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{12}`

必需：否

## SecurityStyle

该卷的安全风格 ( 例如 Unix 或 NTFS ) 。

类型 : 字符串

长度约束 : 最大长度为 1024。

模式 : `^\{0,1024\}`\$

必需 : 否

## SnapshotCapacityUsed

卷中用于快照的存储量。

类型 : 长整型

有效范围 : 最小值为 0。

必需 : 否

## SvmName

卷关联的 SVM 的名称。

类型 : 字符串

长度约束 : 最大长度为 1024。

模式 : `^\{0,1024\}`\$

必需 : 否

## SvmUuid

与该卷关联的存储虚拟机 (SVM) 的 UUID。

类型 : 字符串

模式 : `[a-f0-9]{8}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{12}`

必需 : 否

## VolumeName

卷的名称。

类型：字符串

长度约束：最大长度为 1024。

模式：`^[0,1024]$`

必需：否

## 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)



# NfsMountOptions

指定 DataSync 如何使用 NFS 协议访问某个位置。

## 目录

### Version

指定您希望 NFS 共享时 DataSync 使用的 NFS 版本。如果服务器拒绝使用指定的版本，则任务将失败。

可以指定以下选项：

- **AUTOMATIC** (默认)：DataSync 选择 NFS 版本 4.1。
- **NFS3**：允许在服务器上进行异步写入的无状态协议版本。
- **NFSv4\_0**：支持委托和伪文件系统的有状态、防火墙友好的协议版本。
- **NFSv4\_1**：支持会话、目录委托和并行数据处理的有状态协议版本。NFS 4.1 版还包括 4.0 版中提供的所有功能。

#### Note

DataSync 目前仅支持 Amazon FSx for Net 位置的 NFS 版本 3。NetApp

类型：字符串

有效值：AUTOMATIC | NFS3 | NFS4\_0 | NFS4\_1

必需：否

## 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)

# OnPremConfig

用于网络文件系统 (NFS) 位置的代理的 Amazon 资源名称 (ARN) 列表。

## 目录

### AgentArns

用于 NFS 位置的代理的 ARN。

类型：字符串数组

数组成员：最少 1 项。最多 4 项。

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\-\0-9]+:[0-9]{12}:agent/agent-[0-9a-z]{17}$`

必需：是

## 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)

## Options

配置您的AWS DataSync任务设置。这些选项包括如何DataSync处理文件、对象及其关联的元数据。您还可以指定如何DataSync验证数据完整性、为任务设置带宽限制以及其他选项。

每个任务设置都有默认值。除非需要，否则Options在开始任务之前不必配置任何一项。

### 目录

#### Atime

指定是否保留表示上次读取或写入文件的元数据。如果您设置为 `AtimeBEST Effort`，将DataSync尝试在所有源文件（即任务执行PREPARING阶段之前的版本）上保留原始Atime属性。

#### Note

各平台的行为Atime并不是完全标准的，因此DataSync只能在尽最大努力的基础上做到这一点。

默认值：`BEST Effort`

`BEST Effort`：尝试保留每个文件的 Atime 值（推荐）。

`NONE`：忽略 Atime。

#### Note

如果 Atime 被设置为 `BEST Effort`，则 Mtime 必须被设置为 `PRESERVE`。  
如果 Atime 设置为 `NONE`，则 Mtime 也必须为 `NONE`。

类型：字符串

有效值：`NONE` | `BEST Effort`

必需：否

#### BytesPerSecond

限制DataSync任务使用的带宽。例如，如果您DataSync想最多使用 1 MB，请将此值设置为 `1048576 (=1024*1024)`。

类型：长整型

有效范围：最小值为 -1。

必需：否

## Gid

指定文件拥有者的 POSIX 组 ID (GID)。

有关更多信息，请参阅[复制的元数据DataSync](#)。

默认值：INT\_VALUE。这将保留 ID 的整数值。

INT\_VALUE：保留用户 ID ( UID ) 和 GID ( GID ) 的整数值 ( 推荐 )。

NONE：忽略 UID 和 GID。

类型：字符串

有效值：NONE | INT\_VALUE | NAME | BOTH

必需：否

## LogLevel

指定DataSync发布到 Amazon Lo CloudWatch gs 日志组的日志类型。要指定日志组，请参阅[CloudWatchLogGroupArn](#)。

如果将 LogLevel 设置为 OFF，则不会发布任何日志。BASIC 发布传输的单个文件的错误日志。TRANSFER 为传输并检查完整性的每个文件或对象发布日志。

类型：字符串

有效值：OFF | BASIC | TRANSFER

必需：否


## Mtime

指定是否保留指明在任务执行PREPARING阶段之前最后一次写入文件的时间。当您需要多次运行 a 任务时，必须使用此选项。

默认值: PRESERVE

PRESERVE : 保留原始 Mtime ( 推荐 )

NONE : 忽略 Mtime。

 Note

如果 Mtime 被设置为 PRESERVE , 则 Atime 必须被设置为 BEST\_EFFORT。  
如果 Mtime 被设置为 NONE , 则 Atime 也必须被设置为 NONE。

类型 : 字符串

有效值 : NONE | PRESERVE

必需 : 否

### ObjectTags

指定在对象存储系统之间传输时是否保留对象标签。如果您希望DataSync任务忽略对象NONE标签。

默认值: PRESERVE

类型 : 字符串

有效值 : PRESERVE | NONE

必需 : 否

### OverwriteMode

指定是否覆盖或保留目标位置的数据。例如，如果设置为 NEVER，则目标文件不会被源文件替换（即使目标文件与源文件不同）。如果修改目标位置的文件并同步文件，则可以使用此值来防止覆盖这些更改。

有些存储类别的特定行为会影响 Amazon S3 存储成本。有关详细信息，请参阅[使用 Amazon S3 存储类别时的注意事项DataSync](#)。

类型 : 字符串

有效值 : ALWAYS | NEVER

必需 : 否

## PosixPermissions

指定哪些用户或组可以出于特定目的（例如读取、写入或执行文件）访问文件。

有关更多信息，请参阅[复制的元数据DataSync](#)。

默认值：PRESERVE

PRESERVE：保留 POSIX 式权限（推荐）。

NONE：忽略权限。

### Note

AWS DataSync 可以保留源位置的现有权限。

类型：字符串

有效值：NONE | PRESERVE

必需：否

## PreserveDeletedFiles

指定是否应保留源位置中不存在的目标位置的文件。此选项可能会影响您的 Amazon S3 存储成本。如果您的任务删除对象，则您可能需要为某些存储类别支付最低存储期限费用。有关详细信息，请参阅[使用 Amazon S3 存储类别时的注意事项DataSync](#)。

默认值：PRESERVE

PRESERVE：忽略此类目标文件。

REMOVE：移除源中不存在的目标文件。

### Note

如果将此参数设置为REMOVE，则无法设置TransferMode为ALL。当你传输所有数据时，DataSync不会扫描你的目标位置，也不知道要删除什么。

类型：字符串

有效值：PRESERVE | REMOVE

必需：否

### PreserveDevices

指定是否DataSync应在源位置保留数据块设备和字符设备的元数据，并在目标上使用该设备名称和元数据重新创建文件。DataSync仅复制此类设备的名称和元数据。

#### Note

DataSync无法复制这些设备的实际内容，因为它们是非终端设备，不返回 e nd-of-file (EOF) 标记。

默认值：NONE

NONE：忽略特殊设备（推荐）。

PRESERVE：保留字符和数据块设备元数据。Amazon EFS 目前不支持此选项。

类型：字符串

有效值：NONE | PRESERVE

必需：否

### SecurityDescriptorCopyFlags

指定将 SMB 安全描述符的哪些组件从源对象复制到目标对象。

此值仅用于 SMB 和 Amazon FSx for Windows File Server 位置之间的传输，或两个 FSx for Windows File Server 位置之间的传输。有关更多信息，请参阅[如何DataSync处理元数据](#)。

默认值：OWNER\_DACL

OWNER\_DACL：对于每个复制的对象，DataSync复制以下元数据：

- 对象所有者。
- NTFS 自由访问控制列表 ( DACL )，用于确定是否向对象授予访问权限。

DataSync不会使用此选项复制 NTFS 系统访问控制列表 ( SACL )。

OWNER\_DACL\_SACL：对于每个复制的对象，DataSync复制以下元数据：

- 对象所有者。
- NTFS 自由访问控制列表 ( DACL ) ，用于确定是否向对象授予访问权限。
- SACL ，管理员将使用 SACL 记录对受保护对象的访问尝试。

复制 SACL 需要向DataSync用于访问您的 SMB 位置的 Windows 用户授予其他权限。有关如何选择可确保对文件、文件夹和元数据具有足够权限的用户的信息，请参阅[用户](#)。

NONE：不会复制任何 SMB 安全描述符组件。目标对象由为访问目标位置而提供的用户所有。DACL 和 SACL 根据目标服务器的配置进行设置。

类型：字符串

有效值：NONE | OWNER\_DACL | OWNER\_DACL\_SACL

必需：否

### TaskQueueing

指定在[运行多个任务的某些情况下是否应将您的传输任务](#)放入队列。这是ENABLED默认的。

类型：字符串

有效值：ENABLED | DISABLED

必需：否

### TransferMode

确定仅DataSync传输源位置和目标位置之间不同的数据和元数据，，或源位置的所有内容（不与目标位置进行比较的内容）。

CHANGED：仅DataSync将新内容或不同内容的数据或元数据从源位置复制到目标位置。

ALL：DataSync将所有源位置内容复制到目标位置（不与目标中的内容进行比较）。

类型：字符串

有效值：CHANGED | ALL

必需：否

### Uid

指定文件拥有者的 POSIX 用户 ID (UID)。



有关更多信息，请参阅[复制的元数据DataSync](#)。

默认值：INT\_VALUE。这将保留 ID 的整数值。

INT\_VALUE：保留 UID 和组 ID ( GID ) 的整数值 ( 推荐 )。

NONE：忽略 UID 和 GID。

类型：字符串

有效值：NONE | INT\_VALUE | NAME | BOTH

必需：否

## VerifyMode

指定在传输过程中如何和何时DataSync检查数据的完整性。

默认值：POINT\_IN\_TIME\_CONSISTENT

ONLY\_FILES\_TRANSFERRED ( 推荐 )：DataSync计算源位置传输的文件和元数据的校验和。在传输结束时，将DataSync此校验和与在目标位置的这些文件上计算的校验和进行比较。

在传输到 S3 Glacier Deep Archive 存储类时，我们建议使用此选项。有关更多信息，请参阅[Amazon S3 位置的存储类别注意事项](#)。

POINT\_IN\_TIME\_CONSISTENT：在传输结束时，DataSync扫描整个源位置和目标以验证两个位置是否完全同步。

在传输到 S3 Glacier Deep Archive 存储类时，您不能使用此选项。有关更多信息，请参阅[Amazon S3 位置的存储类别注意事项](#)。

NONE: 在传输结束时DataSync不进行额外验证。在传输过程中，仍将通过校验和验证来检查所有数据传输的完整性。

类型：字符串

有效值：POINT\_IN\_TIME\_CONSISTENT | ONLY\_FILES\_TRANSFERRED | NONE

必需：否

## 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)

## P95Metrics

DataSyncDiscovery 收集的有关本地存储系统资源的性能数据类型。

### 目录

#### IOPS

本地存储系统资源的 IOPS 峰值。每个数据点表示 1 小时间隔内的第 95 个百分位峰值。

类型：[IOPS](#) 对象

必需：否

#### Latency

本地存储系统资源的延迟峰值。每个数据点表示 1 小时间隔内的第 95 个百分位峰值。

类型：[Latency](#) 对象

必需：否

#### Throughput

本地存储系统资源的吞吐量峰值。每个数据点表示 1 小时间隔内的第 95 个百分位峰值。

类型：[Throughput](#) 对象

必需：否

### 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)



模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):ec2:[a-z\-\0-9]*:[0-9]{12}:subnet/.*$`

必需：否

### VpcEndpointId

为 Aget 配置的 VPC 端点的 ID 端点的 ID 端点的 ID。为 VPC 端点配置为 VPC 端点访问度为 VPC 端点访问度为 VPC 端点配置的 Aget t t

类型：字符串

模式：`^vpce-[0-9a-f]{17}$`

必需：否

### 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)

## QopConfiguration

保护质量 ( QOP ) 配置指定在 Hadoop Distributed File System ( HDFS ) 集群上配置的远程程序调用 ( RPC ) 和数据传输隐私设置。

### 目录

#### DataTransferProtection

在 HDFS 集群上配置的数据传输保护设置。此设置对应于 Hadoop 集群上的 `hdfs-site.xml` 文件中的 `dfs.data.transfer.protection` 设置。

类型：字符串

有效值：DISABLED | AUTHENTICATION | INTEGRITY | PRIVACY

必需：否

#### RpcProtection

在 HDFS 集群上配置的 RPC 保护设置。此设置对应于您在 Hadoop 集群上的 `core-site.xml` 文件中的 `hadoop.rpc.protection` 设置。

类型：字符串

有效值：DISABLED | AUTHENTICATION | INTEGRITY | PRIVACY

必需：否

### 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)

## Recommendation

DataSyncDiscovery 为本地AWS存储系统中的资源推荐的存储服务的详细信息。

有关更多信息，[请参阅服务DataSync发现](#)。

### 目录

#### EstimatedMonthlyStorageCost

推荐AWS存储服务的估计每月成本。

类型：字符串

长度约束：最大长度为 1024。

模式：`^\{0,1024\}$`

必需：否

#### StorageConfiguration

有关如何设置推荐的AWS存储服务的信息。

类型：字符串到字符串映射

密钥长度约束：最大长度为 1024。

密钥模式：`^\{0,1024\}$`

值长度约束：最大长度为 1024。

价值模式：`^\{0,1024\}$`

必需：否

#### StorageType

一项推荐的AWS存储服务，您可以根据DataSync Discovery 收集的有关本地存储系统的信息将数据迁移到该服务。

类型：字符串

长度约束：最大长度为 1024。

模式：`^\{0,1024\}$`

必需：否

## 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)



## ResourceDetails

DataSyncDiscovery 提供的有关本地存储系统中资源的信息。

### 目录

#### NetAppONTAPClusters

DataSyncDiscovery 收集的有关本地存储系统中集群的信息。

类型：[NetAppONTAPCluster](#) 对象数组

必需：否

#### NetAppONTAPSVMs

DataSyncDiscovery 收集的有关本地存储系统中存储虚拟机 (SVM) 的信息。

类型：[NetAppONTAPSVM](#) 对象数组

必需：否

#### NetAppONTAPVolumes

DataSyncDiscovery 收集的有关本地存储系统中卷的信息。

类型：[NetAppONTAPVolume](#) 对象数组

必需：否

### 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)

# ResourceMetrics

DataSyncDiscovery 提供的有关本地存储系统中资源的信息，包括性能数据和容量使用情况。

## 目录

### Capacity

本地存储系统资源的存储容量。

类型：[Capacity](#) 对象

必需：否

### P95Metrics

DataSyncDiscovery 收集的有关本地存储系统资源的性能数据类型。

类型：[P95Metrics](#) 对象

必需：否

### ResourceId

本地存储系统资源的通用唯一标识符 (UID)，用于本地存储系统资源。

类型：字符串

模式：`[a-f0-9]{8}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{12}`

必需：否

### ResourceType

本地存储系统资源的类型。

类型：字符串

有效值：SVM | VOLUME | CLUSTER

必需：否

### Timestamp

DataSyncDiscovery 从资源收集这些信息的时间。

类型：Timestamp

必需：否

## 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)

## S3Config

AWS Identity and Access Management ( IAM ) 角色用于访问 Amazon S3 桶的 Amazon 资源名称 ( ARN )。

有关使用此类角色的详细信息，请参阅《AWS DataSync 用户指南》中的[为 Amazon S3 创建位置](#)。

### 目录

#### BucketAccessRoleArn

访问 S3 桶的 IAM 角色的 ARN。

类型：字符串

长度限制：最大长度为 2048。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):iam::[0-9]{12}:role/.*$`

必需：是

### 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)

# SmbMountOptions

指定 AWS DataSync 用于访问服务器消息块 ( SMB ) 文件服务器的 SMB 协议版本。

## 目录

### Version

默认情况下，DataSync会根据与 SMB 文件服务器的协商自动选择 SMB 文件服务器协商自动选择 SMB 文件版本。您也可以配置DataSync为使用特定的 SMB 版本，但我们建议仅在 SMB 配置为仅在 SMB 文件服务器无法自动与 SMB 文件服务器协商的情况下DataSync才这样做。

以下是用于配置 SMB 版本的选项：

- AUTOMATIC ( 默认 ) : DataSync并且 SMB 文件服务器协商它们在 2.1 和 3.1.1 之间相互支持的 SMB 的最高版本。

这是建议的选项。如果您改为选择文件服务器不支持的特定版本，则可能会出现 Operation Not Supported 错误。

- SMB3 : 将协议协商限制为仅 SMB 版本 3.0.2。
- SMB2 : 将协议协商限制为仅 SMB 版本 2.0.2。
- SMB2\_0 : 将协议协商限制为仅 SMB 版本 2.0.2。
- SMB1 : 将协议协商限制为仅 SMB 版本 1.0.0。



#### Note

为 [NetAppONTAP 位置创建 Amazon FSx 时](#)，该SMB1选项不可用。

类型：字符串

有效值：AUTOMATIC | SMB2 | SMB3 | SMB1 | SMB2\_0

必需：否

## 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [适用于 C++ 的 AWS SDK](#)

- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)

# StorageSystemListEntry

用于识别您在DataSync Discovery 中使用的本地存储系统的信息。

## 目录

### Name

您添加到DataSync Discovery 的本地存储系统的名称。

类型：字符串

长度限制：最小长度为 1。长度上限为 256。

模式：`^[\\p{L}\\p{M}\\p{N}\\s+=. _:@\\/-]+`\$

必需：否

### StorageSystemArn

您添加到DataSync Discovery 的本地存储系统的亚马逊资源名称 (ARN)。

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\\-0-9]+:[0-9]{12}:system/storage-system-[a-f0-9]{8}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{12}$`

必需：否

## 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)

## TagListEntry

代表应用于AWS资源的单个标签的键值对。

### 目录

#### Key

AWS资源标签的密钥。

类型：字符串

长度限制：最小长度为 1。长度上限为 256。

模式：`^[a-zA-Z0-9\s+=. _:/-]+$`

必需：是

#### Value

AWS资源标签的值。

类型：字符串

长度限制：最小长度为 1。长度上限为 256。

模式：`^[a-zA-Z0-9\s+=. _:@/-]+$`

必需：是

### 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)



# TaskExecutionListEntry

表示任务执行单个条目。TaskExecutionListEntry返回一个数组，其中包含调用[ListTaskExecutions](#)操作时任务的特定调用列表。

## 目录

### Status

任务执行任务执行状态。

类型：字符串

有效值：QUEUED | LAUNCHING | PREPARING | TRANSFERRING | VERIFYING | SUCCESS | ERROR

必需：否

### TaskExecutionArn

任务的 Amazon 资源名称 ( ARN ) 。

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\-\0-9]*:[0-9]{12}:task/task-[0-9a-f]{17}/execution/exec-[0-9a-f]{17}$`

必需：否

## 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)

## TaskExecutionResultDetail

描述TaskExecution操作的详细结果。此结果包括每个阶段花费的时间（以毫秒为单位）、任务执行状态以及遇到的错误。

### 目录

#### ErrorCode

执行任务期间AWS DataSync遇到的错误。您可以使用此错误代码来帮助解决问题。

类型：字符串

必需：否

#### ErrorDetail

对任务执行过程中遇到的错误的详细描述。您可以使用此信息来帮助排查问题。

类型：字符串

必需：否

#### PrepareDuration

在准备阶段AWS DataSync花费的总时间（以毫秒为单位）。

类型：长整型

有效范围：最小值为 0。

必需：否

#### PrepareStatus

准备阶段的状态。

类型：字符串

有效值：PENDING | SUCCESS | ERROR

必需：否

#### TotalDuration

将文件从源位置传输到目标位置所AWS DataSync花费的总时间（以毫秒为单位）。

类型：长整型

有效范围：最小值为 0。

必需：否

### TransferDuration

在传输阶段AWS DataSync花费的总时间（以毫秒为单位）。

类型：长整型

有效范围：最小值为 0。

必需：否

### TransferStatus

转移阶段的状态。

类型：字符串

有效值：PENDING | SUCCESS | ERROR

必需：否

### VerifyDuration

在验证阶段AWS DataSync花费的总时间（以毫秒为单位）。

类型：长整型

有效范围：最小值为 0。

必需：否

### VerifyStatus

验证阶段的状态。

类型：字符串

有效值：PENDING | SUCCESS | ERROR

必需：否

## 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)

# TaskFilter

您可以使用 API 筛选条件的范围迁移ListTasks。例如，要检索源位置上的所有任务，您可以ListTasks使用过滤器名称LocationId和Operator Equals位置的 ARN。

有关更多信息，请参阅[筛选条件的DataSync内容](#)。

## 目录

### Name

正在使用的筛选条件的名称。每个 API 调用都支持可用的过滤器列表。例如，ListTasks 使用 LocationId。

类型：字符串

有效值：LocationId | CreationTime

必需：是

### Operator

用于比较筛选值的运算符（例如，Equals或Contains）。

类型：字符串

有效值：Equals | NotEquals | In | LessThanOrEqual | LessThan | GreaterThanOrEqual | GreaterThan | Contains | NotContains | BeginsWith

必需：是

### Values

您想要筛选条件的值。例如，您可能希望显示任务以制作特定目标位置的任务。

类型：字符串数组

长度限制：最小长度为 1。最大长度为 255。

模式：`^[0-9a-zA-Z_ \ -\:\*\.\ \\/\?-\]*$`

必需：是

## 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)

## TaskListEntry

表示任务列表中的单个条目。TaskListEntry调用[ListTasks](#)操作时返回一个包含任务列表的数组。任务包括要同步的源文件系统和目标文件系统以及用于任务的选项。

### 目录

#### Name

任务的名称。

类型：字符串

长度限制：最小长度为 1。长度上限为 256。

模式：`^[a-zA-Z0-9\s+.=_:@/-]+$`

必需：否

#### Status

任务的状态。

类型：字符串

有效值：AVAILABLE | CREATING | QUEUED | RUNNING | UNAVAILABLE

必需：否

#### TaskArn

任务的 Amazon 资源名称 (ARN)。

类型：字符串

长度约束：最大长度为 128。

模式：`^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):datasync:[a-z\-\0-9]*:[0-9]{12}:task/task-[0-9a-f]{17}$`

必需：否

### 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)



# TaskSchedule

指定希望任务用于重复执行的计划。有关更多信息，请参阅[规则的计划表达式](#)。

## 目录

### ScheduleExpression

指定 AWS DataSync 何时启动从源到目标位置的计划传输的 cron 表达式。

类型：字符串

长度限制：最大长度为 256。

模式：`^[a-zA-Z0-9\ \_]*\?[,|\/\^-\|\/#\s\(\)\+]*$`

必需：是

## 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)

# Throughput

本地存储系统卷的吞吐量峰值。每个数据点表示 1 小时间隔内的第 95 个百分位峰值。

## 目录

### Other

峰值吞吐量与读取和写入操作无关。

类型：双精度

有效范围：最小值为 0。

必需：否

### Read

与读取操作相关的峰值吞吐量。

类型：双精度

有效范围：最小值为 0。

必需：否

### Total

本地存储系统资源的峰值总吞吐量。

类型：双精度

有效范围：最小值为 0。

必需：否

### Write

与写入操作相关的峰值吞吐量。

类型：双精度

有效范围：最小值为 0。

必需：否

## 另请参阅

有关在特定语言的 AWS 软件开发工具包中使用此 API 的更多信息，请参阅以下内容：

- [适用于 C++ 的 AWS SDK](#)
- [适用于 Go 的 AWS SDK](#)
- [适用于 Java V2 的 AWS SDK](#)
- [适用于 Ruby V3 的 AWS SDK](#)

## 常见错误

本部分列出了所有 API 操作错误AWS服务。对于特定于此服务的 API 操作的错误，请参阅该 API 操作的主题。

### AccessDeniedException

您没有足够的访问权限，无法执行该操作。

HTTP 状态代码：400

### IncompleteSignature

请求签名不符合AWS标准。

HTTP 状态代码：400

### InternalFailure

由于未知错误、异常或故障，请求处理失败。

HTTP 状态代码：500

### InvalidAction

所请求的操作无效。验证操作是否已正确键入。

HTTP 状态代码：400

### InvalidClientTokenId

在我们的记录中没有所提供的 X.509 证书或 AWS 访问密钥 ID。

HTTP 状态代码：403

## InvalidParameterCombination

不得共用的参数被一起使用。

HTTP 状态代码：400

## InvalidParameterValue

为输入参数提供的值无效或超出范围。

HTTP 状态代码：400

## InvalidQueryParameter

这些区域有：AWS查询字符串格式异常或未遵循AWS标准。

HTTP 状态代码：400

## MalformedQueryString

查询字符串包含语法错误。

HTTP 状态代码：404

## MissingAction

请求中遗漏了一个操作或必需参数。

HTTP 状态代码：400

## MissingAuthenticationToken

请求中必须包含有效的（已注册的）AWS 访问密钥 ID 或 X.509 证书。

HTTP 状态代码：403

## MissingParameter

未提供用于指定操作的必需参数。

HTTP 状态代码：400

## NotAuthorized

您没有执行该操作的权限。

HTTP 状态代码：400

## OptInRequired

AWS 访问密钥 ID 需要订阅服务。

HTTP 状态代码：403

## RequestExpired

请求到达服务的时间超过请求上的日期戳或请求到期日期 (如针对预签名 URL) 15 分钟，或者请求上的日期戳离到期还有 15 分钟以上。

HTTP 状态代码：400

## ServiceUnavailable

由于服务器发生临时故障而导致请求失败。

HTTP 状态代码：503

## ThrottlingException

由于请求限制而导致请求被拒绝。

HTTP 状态代码：400

## ValidationError

输入未能满足指定的约束。AWS服务。

HTTP 状态代码：400

## 常见参数

以下列表包含所有操作用于使用查询字符串对 Signature Version 4 请求进行签名的参数。任何特定于操作的参数都列在该操作的主题中。有关签名版本 4 的更多信息，请参阅 IAM 用户指南中的 [AWSAPI 请求签名](#)。

### Action

要执行的操作。

类型：字符串。

必需：是

## Version

编写请求所针对的 API 版本，格式为 YYYY-MM-DD。

类型：字符串。

必需：是

## X-Amz-Algorithm

您用于创建请求签名的哈希算法。

条件：当您在查询字符串中而不是 HTTP 授权标头中包括身份验证信息时，请指定此参数。

类型：字符串

有效值：AWS4-HMAC-SHA256

必需：条件

## X-Amz-Credential

凭证范围值，该值是一个字符串，其中包含您的访问密钥、日期、您要定位的区域、您请求的服务以及终止字符串（“aws4\_request”）。值采用以下格式表示：access\_key/YYYYMMDD/region/service/aws4\_request。

有关更多信息，请参阅 IAM 用户指南中的[创建签名AWS API 请求](#)。

条件：当您在查询字符串中而不是 HTTP 授权标头中包括身份验证信息时，请指定此参数。

类型：字符串

必需：条件

## X-Amz-Date

用于创建签名的日期。格式必须为 ISO 8601 基本格式 (YYYYMMDD'T'HHMMSS'Z')。例如，以下日期时间是有效的 X-Amz-Date 值：20120325T120000Z。

条件：X-Amz-Date 对于所有请求都是可选的；它可以用于覆盖对请求签名所使用的日期。如果以 ISO 8601 基本格式指定 Date 标头，则不需要 X-Amz-Date。使用 X-Amz-Date 时，它始终会覆盖 Date 标头的值。有关更多信息，请参阅 IAM 用户指南中的[AWSAPI 请求签名元素](#)。

类型：字符串

必需：条件

### X-Amz-Security-Token

通过调用AWS Security Token Service (AWS STS) 获得的临时安全令牌。有关支持来自AWS STS 的临时安全凭证的服务列表，请转[AWS 服务到《IAM 用户指南》](#)中的使用 IAM 的服务。

条件：如果您使用来自的临时安全凭证AWS STS，则必须包含安全令牌。

类型：字符串

必需：条件

### X-Amz-Signature

指定从要签名的字符串和派生的签名密钥计算的十六进制编码签名。

条件：当您在查询字符串中而不是 HTTP 授权标头中包括身份验证信息时，请指定此参数。

类型：字符串

必需：条件

### X-Amz-SignedHeaders

指定作为规范请求的一部分包含的所有 HTTP 标头。有关指定签名标头的更多信息，请参阅 IAM 用户指南中的[创建签名AWS API 请求](#)。

条件：当您在查询字符串中而不是 HTTP 授权标头中包括身份验证信息时，请指定此参数。

类型：字符串

必需：条件

## 文档历史记录

下表介绍 AWS DataSync 文档的重要补充部分。我们还经常更新文档来处理您发送给我们的反馈意见。

如需有关文档更新的通知，您可以订阅 RSS 源。

变更	说明	日期
<a href="#">全新 AWS 区域</a>	AWS DataSync 现已在亚太地区（墨尔本）区域推出。	2023 年 5 月 24 日
<a href="#">Support 在 Snowball Edge 上使用兼容 S3 的存储</a>	您可以使用 AWS DataSync 在 Amazon S3 兼容存储 AWS Snowball Edge 和 AWS 存储服务之间传输数据。	2023 年 5 月 18 日
<a href="#">AWS 托管策略更新 – 对现有策略的更新</a>	该 AWSDataSyncFullAccess 策略为可使用的服务提供了新权限 DataSync。	2023 年 5 月 2 日
<a href="#">的正式上市 AWS DataSync 数据查找服务</a>	使用 DataSync Discovery 来帮助加快迁移到 AWS。	2023 年 4 月 25 日
<a href="#">微软 Azure Blob 存储支持的公开发布预览版</a>	AWS DataSync 现在可以从微软 Azure Blob 存储中传输对象。	2023 年 3 月 29 日
<a href="#">新的 IAM 政策</a>	要支持 DataSync 发现功能，请 DataSync 使用名为 AWSServiceRoleForDataSyncDiscovery 的服务相关角色。	2023 年 3 月 21 日
<a href="#">全新 AWS 区域</a>	AWS DataSync 现已在以下区域推出 AWS 区域：亚太地区（海得拉巴）、欧洲（西班牙	2023 年 2 月 6 日



	牙)、欧洲(西班牙)和欧洲(苏黎世)。	
<a href="#">在任务执行中使用标签</a>	现在,您可以标记AWS DataSync任务执行情况。	2022年12月16日
<a href="#">Support S3 Glacier Glacier</a>	现在,您可以将对象直接传输到S3 Glacier Glacier Griever存储类中。	2022年12月16日
<a href="#">复制对象系统元数据</a>	AWS DataSync现在可以在对象存储系统和Amazon S3之间传输时复制系统元数据。	2022年12月16日
<a href="#">全新 AWS 区域</a>	AWS DataSync 现已在中国(北京)和中国(宁夏)区域推出。	2022年12月14日
<a href="#">全新 AWS 区域</a>	AWS DataSync现已在中东(阿联酋)区域推出。	2022年11月16日
<a href="#">Support 带有对象存储位置的自签名证书</a>	AWS DataSync可以连接到使用自签名或私有证书的对象存储位置。	2022年10月25日
<a href="#">获取数据压缩信息</a>	AWS DataSync可以提供应用压缩后通过网络传输的物理字节数。	2022年10月25日
<a href="#">的公开发布预览版 AWS DataSync 数据查找服务</a>	使用 DataSync Discovery 来帮助加快迁移到AWS。	2022年9月21日
<a href="#">将数据迁移到或迁出 Google 云存储的新选项</a>	您可以在 Google Cloud 中部署AWS DataSync代理,将数据传输到 Google 云存储或从 Google 云存储中传输数据。	2022年7月21日

<a href="#">Support NetApp ONTAP 文件系统的亚马逊 FSx</a>	AWS DataSync现在可以向 ONTAP 文件系统的 FsX 传输文件和文件夹，或从 FsX 传输文件和文件夹。	2022 年 6 月 28 日
<a href="#">亚马逊 EFS 地点的新安全选项</a>	AWS DataSync可以使用 TLS、接入点和 IAM 角色访问亚马逊 EFS 文件系统。	2022 年 5 月 31 日
<a href="#">将数据迁入或迁出谷歌云存储和 Azure 文件</a>	使用AWS DataSync，你可以将数据传输到 Google 云存储和 Azure 文件或从中传输数据。有关更多信息，请参阅 <a href="#">创建对象存储位置</a> 和为 <a href="#">SMB 创建位置</a> 。	2022 年 5 月 24 日
<a href="#">新AWS DataSync任务设置</a>	使用复制对象标签选项，您可以指定在对象存储系统之间传输时是否保留对象标签。	2022 年 5 月 5 日
<a href="#">全新 AWS 区域</a>	AWS DataSync现已在亚太地区（雅加达）区域推出。	2022 年 4 月 19 日
<a href="#">Support Amazon FSx for OpenZFS 文件系统</a>	AWS DataSync现在可以在 OpenZFSx S 文件系统之间传输文件和文件夹。	2022 年 4 月 5 日
<a href="#">Support Amazon FSx for Lustre 文件系统</a>	AWS DataSync现在可以将文件和文件夹传输到 FsX for Lustre 文件系统，或从 FsX 传输文件和文件夹。	2021 年 2 月 12 月 12 月
<a href="#">Support Hadoop 分布式文件系统 (HDFS)</a>	AWS DataSync现在支持在 HDFS 集群之间传输文件和文件夹。	2021 年 11 月 3 日
<a href="#">全新 AWS 区域</a>	AWS DataSync现已在亚太地区（大阪）区域推出。	2021 年 7 月 28 日

<a href="#">AWS存储服务之间的全自动传输</a>	AWS DataSync现在，只需在 DataSync控制台中单击几下，即可在 Amazon S3、Amazon EFS 或 Amazon EFS 或 FSx for Windows File Server 中。	2020 年 11 月 9 日
<a href="#">调整正在运行的任务使用的网络带宽</a>	AWS DataSync现在，客户可以调整正在运行的DataSync任务所使用的网络带宽。当任务跨越多天时，这有助于最大限度地减少对其他用户或应用程序的影响。	2020 年 11 月 9 日
<a href="#">增强对本地DataSync虚拟机 (VM) 功能的支持</a>	AWS DataSync代理虚拟机主机控制台现在支持增强功能，包括从本地控制台激活代理。	2020 年 10 月 19 日
<a href="#">AWS DataSync现在可以往返传输数据 AWS Outposts</a>	DataSync现在支持在 Amazon S3 之间传输对象AWS Outposts。	2020 年 9 月 30 日
<a href="#">Support API 筛选</a>	AWS DataSync现在支持筛选ListTasks 和 ListLocations API 调用，使您能够使用过滤器（例如数据传输的来源或目标）轻松检索数据传输任务的配置。	2020 年 8 月 8 月 8 月
<a href="#">Support 从自行管理的对象存储中复制数据</a>	AWS DataSync现在支持在自管对象存储与 Amazon S3、Amazon Elastic File System 或 FsX for Windows 文件服务器之间传输数据。	2020 年 7 月 27 日

<a href="#">Support 基于 Linux 内核的虚拟机 (KVM) 和微软 Hyper-V 虚拟机管理程序</a>	AWS DataSync除了现有的 VMware 和 Amazon EC2 选项外，现在还允许在 KVM 和微软 Hyper-V 虚拟化平台上部署本地代理。	2020 年 7 月 1 日
<a href="#">AWS DataSync现在可以自动配置您的 Amazon CloudWatch Logs 配置</a>	使用时DataSync，您现在可以选择自动生成发布数据传输 CloudWatch日志所需的日志组和资源策略，从而简化任务创建和监控设置。	2020 年 7 月 1 日
<a href="#">AWS DataSync现在可以往返传输数据 AWS Snowcone</a>	DataSync现在支持在边缘计算 AWS Snowcone和数据传输设备的最小成员之间传输文件。AWS Snow FamilySnowcone 便于携带、坚固耐用、安全——小巧轻便，可以装在背包里，能够承受恶劣的环境。	2020 年 6 月 17 日
<a href="#">全新 AWS 区域</a>	AWS DataSync现已在非洲（开普敦）和欧洲（斯兰）区域推出。	2020 年 6 月 16 日
<a href="#">通过文件级日志增强监控功能</a>	现在，您可以对在 NFS 服务器、SMB 服务器、Amazon S3 存储桶、亚马逊 EFS 文件系统和适用于 Windows 文件服务器文件系统的 FsX 之间复制的文件和对象启用详细日志记录。	2020 年 4 月 24 日
<a href="#">Support Amazon FSx for Windows File Server 中，复制数据。</a>	现在，你可以在 SMB 共享和 FSx for Windows 文件服务器之间复制数据。	2020 年 1 月 24 日
<a href="#">Support 调度任务</a>	您现在可以手动运行任务，也可以根据指定计划安排任务运行。	2019 年 11 月 20 日



[Support 筛选](#)

现在，当您将数据从源传输到目标位置时，您可以应用筛选条件以仅传输源位置中的文件子集。

2019 年 5 月 22 日

[的首次发布 AWS DataSync](#)

AWS DataSync 服务的一般发布版本。

2018 年 11 月 26 日

# AWS词汇表

有关最新 AWS 术语，请参阅《AWS 一般参考》中的 [AWS 术语表](#)。

本文属于机器翻译版本。若本译文内容与英语原文存在差异，则一律以英文原文为准。