



User Guide

# AWS DataSync



# AWS DataSync: User Guide

Copyright © 2024 Amazon Web Services, Inc. and/or its affiliates. All rights reserved.

Amazon's trademarks and trade dress may not be used in connection with any product or service that is not Amazon's, in any manner that is likely to cause confusion among customers, or in any manner that disparages or discredits Amazon. All other trademarks not owned by Amazon are the property of their respective owners, who may or may not be affiliated with, connected to, or sponsored by Amazon.

---

# Table of Contents

<b>What is AWS DataSync?</b> .....	<b>1</b>
Use cases .....	2
Benefits .....	3
Additional resources .....	3
<b>How it works</b> .....	<b>4</b>
DataSync transfer architecture .....	4
Transferring between on-premises storage and AWS .....	4
Transferring between AWS storage services .....	5
Transferring between cloud storage systems and AWS storage services .....	6
Concepts and terminology .....	6
Agent .....	7
Location .....	7
Task .....	7
Task execution .....	7
How DataSync transfers files, objects, and directories .....	7
Open and locked files .....	8
Data integrity .....	8
Recurring transfers .....	9
<b>Setting up</b> .....	<b>10</b>
Sign up for an AWS account .....	10
Create a user with administrative access .....	10
Required IAM permissions for using DataSync .....	12
AWS managed policies .....	12
Customer managed policies .....	12
Where can I use DataSync? .....	12
How can I use DataSync? .....	12
How much will DataSync cost? .....	13
Do I need an agent? .....	13
Situations when you need a DataSync agent .....	13
Situations when you don't need a DataSync agent .....	14
Using multiple DataSync agents .....	14
Next steps .....	15
Agent requirements .....	15
Hypervisor requirements .....	15

Agent requirements for DataSync transfers .....	16
Agent requirements for DataSync Discovery .....	17
Agent requirements for partitions .....	17
Agent management requirements .....	17
Deploying your agent .....	17
Deploying your agent on VMware .....	18
Deploying your agent on KVM .....	18
Deploying your agent on Microsoft Hyper-V .....	19
Deploying your Amazon EC2 agent .....	20
Deploying your agent on AWS Snowball Edge .....	25
Deploying your agent on AWS Snowcone .....	25
Deploying your agent on AWS Outposts .....	26
Choosing a service endpoint for your agent .....	26
Using a VPC service endpoint with DataSync .....	26
Using a public service endpoint with DataSync .....	29
Using a FIPS service endpoint with DataSync .....	30
Activating your agent .....	30
<b>Networking .....</b>	<b>32</b>
1. Network connection between your storage system and agent .....	32
2. Network connection between your agent and DataSync service .....	32
Connecting your storage network to AWS .....	33
Choosing a service endpoint .....	33
3. Network connection between DataSync service and AWS storage service .....	33
Networking when you don't need a DataSync agent .....	34
How and where DataSync traffic flows through the network .....	34
Network security for DataSync .....	34
Network requirements .....	34
Network requirements for on-premises, self-managed, other cloud, and edge storage .....	34
Network requirements for AWS storage services .....	37
Network requirements for VPC service endpoints .....	38
Network requirements for public or FIPS service endpoints .....	41
Network interfaces for data transfers .....	47
Network interfaces for transfers with agents .....	48
Network interfaces for transfers without agents .....	48
Viewing your network interfaces .....	49
Architecture and routing examples with Direct Connect .....	50



Using Direct Connect with a DataSync VPC service endpoint .....	50
Using Direct Connect with a DataSync public or FIPS service endpoint .....	53
Next steps .....	54
Configuring your agent for multiple NICs .....	54
<b>Transferring data .....</b>	<b>56</b>
Where can I transfer my data? .....	56
DataSync transfer requirements .....	56
Supported transfers in the same AWS account .....	56
Supported transfers across AWS accounts .....	59
Supported transfers in the same AWS Region .....	60
Supported transfers between AWS Regions .....	60
Transferring to or from on-premises storage .....	61
Configuring transfers with an NFS file server .....	61
Configuring transfers with an SMB file server .....	65
Configuring transfers with HDFS .....	69
Configuring transfers with an object storage system .....	73
Transferring to or from AWS storage .....	75
Configuring transfers with Amazon S3 .....	75
Configuring transfers with Amazon EFS .....	103
Configuring transfers with FSx for Windows File Server .....	108
Configuring transfers with FSx for Lustre .....	112
Configuring transfers with FSx for OpenZFS .....	113
Configuring transfers with FSx for ONTAP .....	116
Transferring to or from other cloud storage .....	122
Configuring transfers with Google Cloud Storage .....	122
Configuring transfers with Microsoft Azure Blob Storage .....	130
Configuring transfers with Microsoft Azure Files .....	148
Configuring transfers with other cloud object storage .....	151
Transferring to or from edge storage .....	157
Configuring transfers with S3 compatible storage on Snowball Edge .....	157
Configuring transfers with Snowcone .....	163
Configuring what DataSync transfers .....	164
Using a manifest .....	164
Using filters .....	178
Metadata copied by DataSync .....	183
Links and directories copied by DataSync .....	193

Configuring how to handle files, objects, and metadata .....	194
Verifying data integrity .....	198
Working with transfer tasks .....	200
Setting bandwidth limits .....	200
Scheduling your task .....	202
Tagging your tasks .....	206
Starting your task .....	209
Cancelling your task execution .....	210
Task statuses .....	211
<b>Discovering storage .....</b>	<b>213</b>
How it works .....	213
DataSync Discovery architecture .....	213
Concepts and terminology .....	214
Limitations .....	216
Adding your on-premises storage system .....	216
Accessing your on-premises storage system .....	216
Adding your on-premises storage system .....	217
Removing your on-premises storage system .....	219
Logging DataSync Discovery activity to Amazon CloudWatch .....	219
Working with discovery jobs .....	220
Starting a discovery job .....	220
Stopping a discovery job .....	221
Viewing storage resource information .....	221
Viewing information collected about your storage system .....	222
Getting recommendations .....	224
What's included in the recommendations? .....	225
What's not included in the recommendations? .....	225
Getting recommendations .....	226
DataSync Discovery statuses .....	228
Discovery job statuses .....	228
Recommendation statuses .....	229
<b>Monitoring .....</b>	<b>231</b>
Monitoring with task reports .....	231
Use cases .....	232
Summary only task reports .....	232
Standard task reports .....	232

---

Example task reports .....	235
Prerequisites .....	237
Creating a summary only task report .....	239
Creating a standard task report .....	243
Viewing your task reports .....	247
Limitations .....	247
Monitoring with CloudWatch .....	248
CloudWatch metrics for DataSync .....	248
Allowing DataSync to upload logs to CloudWatch log groups .....	250
Configuring logging for your DataSync transfer task .....	252
Logging with CloudTrail .....	254
Working with DataSync information in CloudTrail .....	254
Understanding DataSync log file entries .....	255
Monitoring with EventBridge .....	256
DataSync transfer events .....	257
DataSync Discovery events .....	258
Monitoring with manual tools .....	259
Monitoring your transfer by using the DataSync console .....	259
Monitoring your transfer by using the AWS CLI .....	260
Monitoring your transfer by using the watch utility .....	262
<b>Managing .....</b>	<b>263</b>
Managing your DataSync agent .....	263
Working with your DataSync agent's local console .....	263
Replacing your DataSync agent .....	263
Cleaning up DataSync resources .....	263
Reusing a DataSync agent's infrastructure .....	263
Managing your agent .....	263
Agent software updates .....	264
Agent statuses .....	264
Troubleshooting your agent .....	265
Working with your agent's local console .....	265
Accessing the agent's local console .....	265
Getting an agent activation key .....	266
Configuring your agent network settings .....	267
Testing your agent's connection to AWS .....	270
Testing your agent's connection to your storage .....	271

Checking your agent's system resources .....	272
Synchronizing the time on your VMware agent .....	273
Configuring other agent settings .....	275
Getting help with your agent from AWS Support .....	276
Replacing your agent .....	277
Creating a new agent .....	277
Updating your location with the new agent .....	278
Next steps .....	283
Cleaning up resources .....	283
Deleting a DataSync agent .....	283
Deleting a DataSync location .....	284
Deleting a DataSync task .....	284
Reusing agent infrastructure .....	286
<b>Security .....</b>	<b>287</b>
Data protection .....	287
Encryption in transit .....	288
Encryption at rest .....	291
Internetwork traffic privacy .....	292
Identity and access management .....	292
Access management .....	293
AWS managed policies .....	298
Customer managed policies .....	302
Using service-linked roles .....	305
Tagging resources during creation .....	308
Cross-service confused deputy prevention .....	310
API permissions reference .....	311
Compliance validation .....	321
Resilience .....	322
Infrastructure security .....	322
<b>Quotas .....</b>	<b>324</b>
Storage system, file, and object limits .....	324
DataSync task quotas .....	324
DataSync Discovery quotas .....	329
Request a quota increase .....	329
<b>Troubleshooting .....</b>	<b>330</b>
Agents .....	330

How do I connect to an Amazon EC2 agent's local console? .....	330
What does the Failed to retrieve agent activation key error mean? .....	330
I still can't activate an agent by using a VPC service endpoint .....	331
What do I do if my agent is offline? .....	331
I don't know what's going on with my agent. Can someone help me? .....	332
Data transfers .....	332
How do I configure DataSync to use a specific NFS or SMB version to mount my file share? .....	332
Error: Invalid SyncOption value. Option: TransferMode,PreserveDeletedFiles, Value: ALL,REMOVE. ....	333
My task keeps failing with an EniNotFound error .....	333
My task failed with an NFS permissions denied error .....	334
My task failed with an NFS mount error .....	334
My task failed with an Amazon EFS mount error .....	335
File ownership isn't maintained with NFS transfer .....	335
My task failed with a Cannot allocate memory error .....	335
My task failed with an input/output error .....	336
My task execution has a launching status but nothing seems to be happening .....	337
My task execution seems stuck in the preparing status .....	337
How long does it take DataSync to verify a task I've run? .....	337
My task fails when transferring from a Google Cloud Storage bucket .....	338
My task fails with an Unable to list Azure Blobs on the volume root error ...	338
My task's start and end times don't match up with what's in the logs .....	339
Error: SyncTaskDeletedByUser .....	339
Error: NoMem .....	339
Error: FsS3UnableToConnectToEndpoint .....	339
Error: FsS3HeadBucketFailed .....	340
Error: FsAzureBlobVolRootListBlobsFailed .....	340
Error: SicLocHitAccess .....	340
Error: SyncTaskErrorLocationNotAdded .....	340
Task report errors .....	340
Task with Amazon S3 fails with HeadObject or GetObjectTagging error .....	341
Costs .....	341
<b>Tutorials .....</b>	<b>343</b>
Transferring from on-premises to S3 across accounts .....	343
Overview .....	343

Prerequisite: Required source account permissions .....	344
Prerequisite: Required destination account permissions .....	346
Step 1: In your source account, create a DataSync agent .....	347
Step 2: In your source account, create a DataSync IAM role for destination bucket access ..	347
Step 3: In your destination account, update your S3 bucket policy .....	349
Step 4: In your destination account, disable ACLs for your S3 bucket .....	350
Step 5: In your source account, create a DataSync source location for your on-premises storage .....	351
Step 6: In your source account, create a DataSync destination location for your S3 bucket .....	351
Step 7: In your source account, create and start your DataSync task .....	352
Related resources .....	352
<b>Transferring between S3 buckets across accounts .....</b>	<b>354</b>
Overview .....	354
Prerequisite: Required source account permissions .....	356
Prerequisite: Required destination account permissions .....	358
Step 1: In your source account, create a DataSync IAM role for destination bucket access ..	358
Step 2: In your destination account, update your S3 bucket policy .....	360
Step 3: In your destination account, disable ACLs for your S3 bucket .....	361
Step 4: In your source account, create your DataSync locations .....	362
Step 5: In your source account, create and start your DataSync task .....	364
Troubleshooting .....	365
Related: Cross-account transfers with S3 buckets using server-side encryption .....	364
<b>Additional resources .....</b>	<b>366</b>
Other use cases .....	366
Transferring data in opposite directions .....	366
Using multiple DataSync tasks to transfer to the same S3 bucket .....	367
Allowing DataSync to access a restricted S3 bucket .....	367
Open-source components .....	368
<b>Using the AWS CLI .....</b>	<b>370</b>
Creating an agent .....	371
Creating locations .....	375
Creating an SMB location .....	375
Creating an HDFS location .....	376
Creating an object storage location .....	377
Creating an Amazon EFS location .....	379

Creating an Amazon FSx for Windows File Server location .....	381
Creating an Amazon FSx for Lustre location .....	382
Creating an Amazon FSx for OpenZFS location .....	383
Creating an Amazon FSx for NetApp ONTAP location .....	384
Creating a task .....	385
Starting a task .....	387
Filtering resources .....	388
Parameters for filtering .....	388
Filtering by location .....	389
Filtering by task .....	390
<b>DataSync API .....</b>	<b>392</b>
Actions .....	392
AddStorageSystem .....	395
CancelTaskExecution .....	401
CreateAgent .....	403
CreateLocationAzureBlob .....	408
CreateLocationEfs .....	413
CreateLocationFsxLustre .....	419
CreateLocationFsxOntap .....	423
CreateLocationFsxOpenZfs .....	428
CreateLocationFsxWindows .....	432
CreateLocationHdfs .....	437
CreateLocationNfs .....	444
CreateLocationObjectStorage .....	449
CreateLocationS3 .....	455
CreateLocationSmb .....	462
CreateTask .....	468
DeleteAgent .....	476
DeleteLocation .....	478
DeleteTask .....	480
DescribeAgent .....	482
DescribeDiscoveryJob .....	487
DescribeLocationAzureBlob .....	491
DescribeLocationEfs .....	495
DescribeLocationFsxLustre .....	500
DescribeLocationFsxOntap .....	503

DescribeLocationFsxOpenZfs .....	507
DescribeLocationFsxWindows .....	511
DescribeLocationHdfs .....	515
DescribeLocationNfs .....	520
DescribeLocationObjectStorage .....	524
DescribeLocationS3 .....	528
DescribeLocationSmb .....	533
DescribeStorageSystem .....	538
DescribeStorageSystemResourceMetrics .....	543
DescribeStorageSystemResources .....	549
DescribeTask .....	559
DescribeTaskExecution .....	568
GenerateRecommendations .....	578
ListAgents .....	581
ListDiscoveryJobs .....	584
ListLocations .....	587
ListStorageSystems .....	590
ListTagsForResource .....	593
ListTaskExecutions .....	596
ListTasks .....	599
RemoveStorageSystem .....	602
StartDiscoveryJob .....	604
StartTaskExecution .....	608
StopDiscoveryJob .....	615
TagResource .....	617
UntagResource .....	620
UpdateAgent .....	622
UpdateDiscoveryJob .....	624
UpdateLocationAzureBlob .....	626
UpdateLocationHdfs .....	630
UpdateLocationNfs .....	636
UpdateLocationObjectStorage .....	639
UpdateLocationSmb .....	644
UpdateStorageSystem .....	648
UpdateTask .....	652
UpdateTaskExecution .....	658



Data Types .....	660
AgentListEntry .....	663
AzureBlobSasConfiguration .....	665
Capacity .....	666
Credentials .....	668
DiscoveryJobListEntry .....	670
DiscoveryServerConfiguration .....	671
Ec2Config .....	672
FilterRule .....	674
FsxProtocol .....	676
FsxProtocolNfs .....	677
FsxProtocolSmb .....	678
HdfsNameNode .....	680
IOPS .....	682
Latency .....	684
LocationFilter .....	686
LocationListEntry .....	688
ManifestConfig .....	690
MaxP95Performance .....	692
NetAppONTAPCluster .....	695
NetAppONTAPSVMS .....	699
NetAppONTAPVolume .....	703
NfsMountOptions .....	707
OnPremConfig .....	709
Options .....	710
P95Metrics .....	718
Platform .....	720
PrivateLinkConfig .....	721
QopConfiguration .....	723
Recommendation .....	724
ReportDestination .....	726
ReportDestinationS3 .....	727
ReportOverride .....	729
ReportOverrides .....	730
ReportResult .....	732
ResourceDetails .....	733

---

ResourceMetrics .....	735
S3Config .....	737
S3ManifestConfig .....	738
SmbMountOptions .....	740
SourceManifestConfig .....	742
StorageSystemListEntry .....	743
TagListEntry .....	745
TaskExecutionListEntry .....	746
TaskExecutionResultDetail .....	747
TaskFilter .....	750
TaskListEntry .....	752
TaskReportConfig .....	754
TaskSchedule .....	756
TaskScheduleDetails .....	758
Throughput .....	760
Common Errors .....	761
Common Parameters .....	762
<b>Document history .....</b>	<b>766</b>
<b>AWS Glossary .....</b>	<b>777</b>

# What is AWS DataSync?

AWS DataSync is an online data movement and discovery service that simplifies data migration and helps you quickly, easily, and securely transfer your file or object data to, from, and between AWS storage services.

## On-premises storage transfers

DataSync works with the following on-premises storage systems:

- [Network File System \(NFS\)](#)
- [Server Message Block \(SMB\)](#)
- [Hadoop Distributed File Systems \(HDFS\)](#)
- [Object storage](#)

## AWS storage transfers

DataSync works with the following AWS storage services:

- [Amazon S3](#)
- [Amazon EFS](#)
- [Amazon FSx for Windows File Server](#)
- [Amazon FSx for Lustre](#)
- [Amazon FSx for OpenZFS](#)
- [Amazon FSx for NetApp ONTAP](#)

## Other cloud storage transfers

DataSync works with the following other cloud storage services:

- [Google Cloud Storage](#)
- [Microsoft Azure Blob Storage](#)
- [Microsoft Azure Files](#)
- [Wasabi Cloud Storage](#)
- [DigitalOcean Spaces](#)

- [Oracle Cloud Infrastructure Object Storage](#)
- [Cloudflare R2 Storage](#)
- [Backblaze B2 Cloud Storage](#)
- [NAVER Cloud Object Storage](#)
- [Alibaba Cloud Object Storage Service](#)
- [IBM Cloud Object Storage](#)
- [Seagate Lyve Cloud](#)

## Edge storage transfers

DataSync works with the following edge storage services and devices:

- [Amazon S3 compatible storage on AWS Snowball Edge](#)
- [AWS Snowcone](#)

## Use cases

These are some of the main use cases for DataSync:

- **Discover data** – Get visibility into your on-premises storage performance and utilization. AWS DataSync Discovery can also provide recommendations for migrating your data to AWS storage services.
- **Migrate data** – Move active datasets rapidly over the network into AWS storage services. DataSync includes automatic encryption and data integrity validation to help make sure that your data arrives securely, intact, and ready to use.
- **Archive cold data** – Move cold data stored in on-premises storage directly to durable and secure long-term storage classes such as S3 Glacier Flexible Retrieval or S3 Glacier Deep Archive. Doing so can free up on-premises storage capacity and shut down legacy systems.
- **Replicate data** – Copy data into any Amazon S3 storage class, choosing the most cost-effective storage class for your needs. You can also send data to Amazon EFS, FSx for Windows File Server, FSx for Lustre, or FSx for OpenZFS for a standby file system.
- **Move data for timely in-cloud processing** – Move data in or out of AWS for processing. This approach can speed up critical hybrid cloud workflows across many industries. These include machine learning in the life-sciences industry, video production in media and entertainment, big-data analytics in financial services, and seismic research in the oil and gas industry.

## Benefits

By using DataSync, you can get the following benefits:

- **Simplify migration planning** – With automated data collection and recommendations, DataSync Discovery can minimize the time, effort, and costs associated with planning your data migrations to AWS. You can use recommendations to inform your budget planning and re-run discovery jobs to validate your assumptions as you approach your migration.
- **Automate data movement** – DataSync makes it easier to move data over the network between storage systems and services. DataSync automates both the management of data-transfer processes and the infrastructure required for high performance and secure data transfer.
- **Transfer data securely** – DataSync provides end-to-end security, including encryption and integrity validation, to help ensure that your data arrives securely, intact, and ready to use. DataSync accesses your AWS storage through built-in AWS security mechanisms, such as AWS Identity and Access Management (IAM) roles. It also supports virtual private cloud (VPC) endpoints, giving you the option to transfer data without traversing the public internet and further increasing the security of data copied online.
- **Move data faster** – DataSync uses a purpose-built network protocol and a parallel, multi-threaded architecture to accelerate your transfers. This approach speeds up migrations, recurring data-processing workflows for analytics and machine learning, and data-protection processes.
- **Reduce operational costs** – Move data cost-effectively with the flat, per-gigabyte pricing of DataSync. Avoid having to write and maintain custom scripts or use costly commercial transfer tools.

## Additional resources

We recommend that you read the following:

- [DataSync resources](#) – Includes blogs, videos, and other training materials
- [AWS re:Post](#) – See the latest discussion around DataSync
- [AWS DataSync pricing](#)

# How AWS DataSync works

Learn the key concepts and terminology related to AWS DataSync transfers.

## DataSync transfer architecture

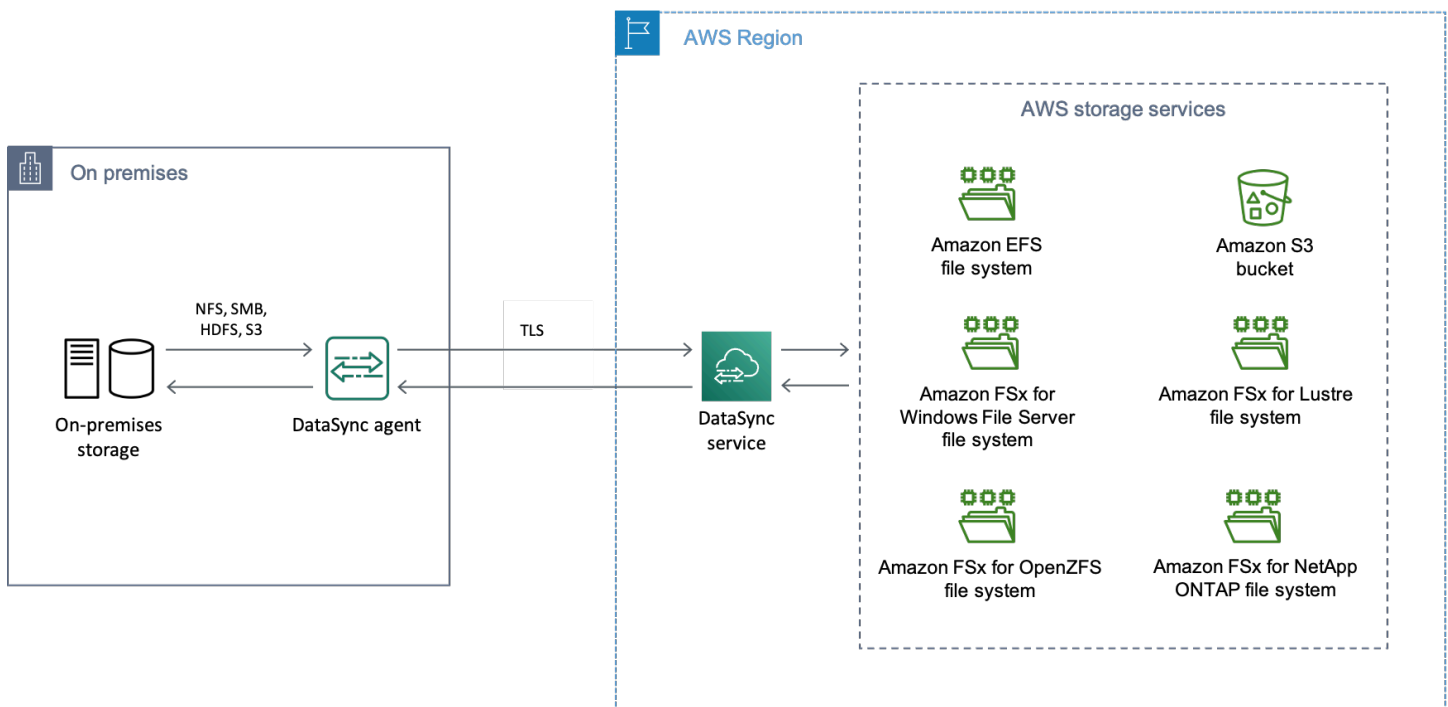
The following diagrams show how and where DataSync commonly transfers storage data. For a full list of DataSync supported storage systems and services, see [Where can I transfer my data with AWS DataSync?](#)

### Topics

- [Transferring between on-premises storage and AWS](#)
- [Transferring between AWS storage services](#)
- [Transferring between cloud storage systems and AWS storage services](#)

## Transferring between on-premises storage and AWS

The following diagram shows a high-level overview of DataSync transferring files between self-managed, on-premises storage systems and AWS services.

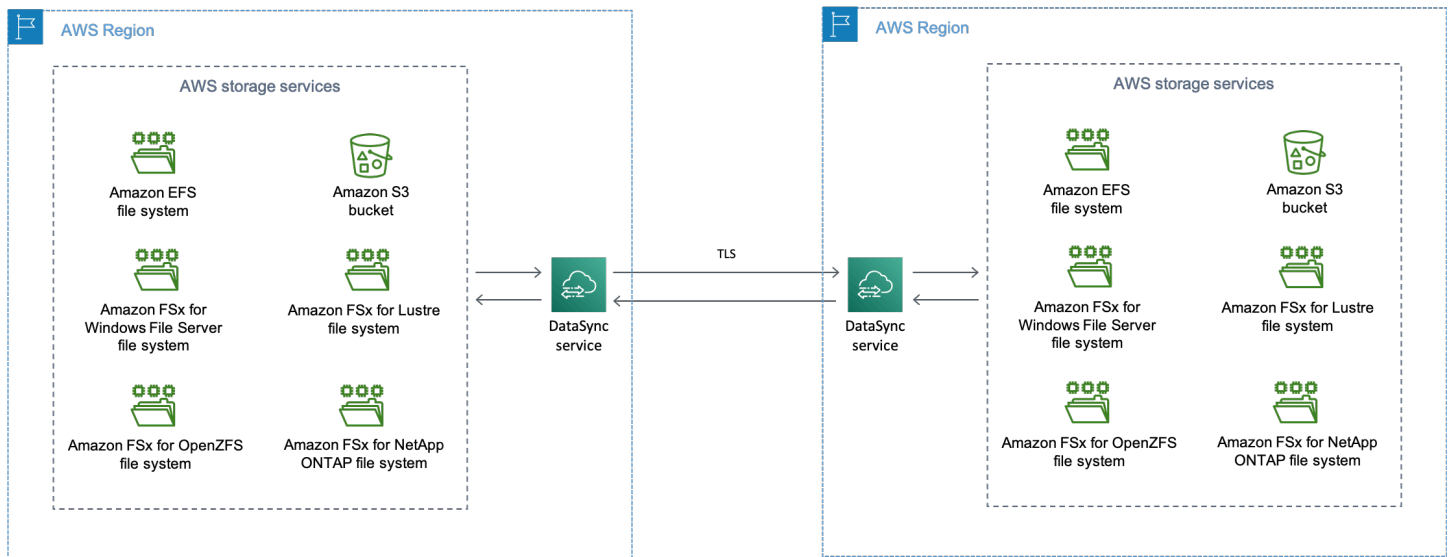


The diagram illustrates a common DataSync use case:

- A DataSync agent copying data from an on-premises storage system.
- Data moving into AWS via Transport Layer Security (TLS).
- DataSync copying data to a supported AWS storage service.

## Transferring between AWS storage services

The following diagram shows a high-level overview of DataSync transferring files between AWS services in the same AWS account.



The diagram illustrates a common DataSync use case:

- DataSync copying data from a supported AWS storage service.
- Data moving across AWS Regions via TLS.
- DataSync copying data to a supported AWS storage service.

When transferring between AWS storage services (whether in the same AWS Region or across AWS Regions), your data remains in the AWS network and doesn't traverse the public internet.

### **⚠ Important**

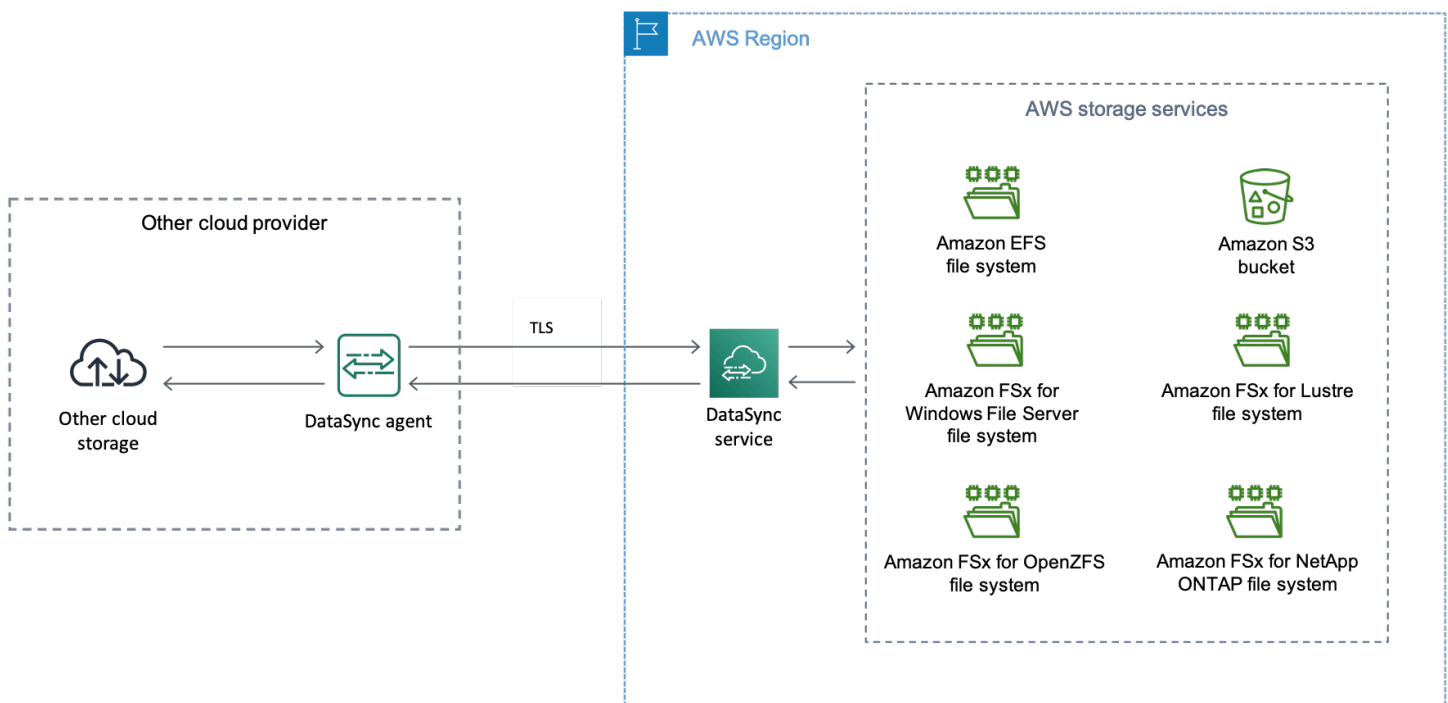
You pay for data transferred between AWS Regions. This is billed as data transfer OUT from your source Region to your destination Region. For more information, see [Data transfer pricing](#).

## Transferring between cloud storage systems and AWS storage services

With DataSync, you can transfer data between other cloud storage systems and AWS services. In this context, cloud storage systems can include:

- Self-managed storage systems, such as an NFS file server in your virtual private cloud (VPC) within AWS.
- Storage systems or services hosted by another cloud provider. For more information, see [Transferring to or from other cloud storage with AWS DataSync](#).

The following diagram shows a high-level overview of DataSync transferring data between AWS storage services and another cloud provider.



## Concepts and terminology

Familiarize yourself with DataSync transfer features.

### Topics

- [Agent](#)
- [Location](#)
- [Task](#)



- [Task execution](#)

## Agent

An *agent* is a virtual machine (VM) appliance that DataSync uses to read from and write to storage during a transfer.

You can deploy an agent in your storage environment on VMware ESXi, Linux Kernel-based Virtual Machine (KVM), or Microsoft Hyper-V hypervisors. For storage in a virtual private cloud (VPC) in AWS, you can deploy an agent as an Amazon EC2 instance.

A DataSync transfer agent is no different than an agent that you can use for DataSync Discovery, but we don't recommend using the same agent for these scenarios.

To get started, see [Do I need an AWS DataSync agent?](#)

## Location

A *location* describes where you're copying data from or to. Each DataSync transfer (also known as a *task*) has a source and destination location. For more information, see [Where can I transfer my data with AWS DataSync?](#)

## Task

A *task* describes a DataSync transfer. It identifies a source and destination location along with details about how to copy data between those locations. You also can specify how a task treats metadata, deleted files, and permissions.

## Task execution

A *task execution* is an individual run of a DataSync transfer task. There are several phases involved in a task execution. For more information, see [Task execution statuses](#).

## How DataSync transfers files, objects, and directories

When you start a task, DataSync prepares your transfer by examining your source and destination locations to determine what to transfer. This is done by recursively scanning the contents and metadata of both locations to identify differences between the two. This process can take just

minutes or a few hours depending on the number of files, objects, or directories in both locations and the performance of your storage systems or services.

During preparation, the number of files, objects, or directories that DataSync takes inventory of in your source and destination counts towards your [task quotas](#). The quotas aren't based on the number of items that DataSync transfers during each task execution.

Once DataSync is done preparing your transfer, it moves your data (including metadata) from the source to the destination based on your [task settings](#). For example, you can specify what metadata gets copied, exclude certain files, limit how much bandwidth DataSync uses, among other options.

At the end of the transfer, DataSync can [verify the integrity](#) of your data.

For information on the specific steps that take place during a task execution, see [DataSync task statuses](#).

## Topics

- [Open and locked files](#)
- [Data integrity](#)
- [Recurring transfers](#)

## Open and locked files


Keep in mind the following when trying to transfer files that are open (in use) or locked:

- In general, DataSync can transfer open files without any limitations.
- If a file is open and being written to during a transfer, DataSync can detect this kind of inconsistency during the transfer task's verification phase. To get the latest version of the file, you must run the task again.
- If a file is locked and the server prevents DataSync from opening it, DataSync skips the file during the transfer and logs an error.
- DataSync can't lock or unlock files.

## Data integrity

DataSync always performs data-integrity checks during a transfer. When your transfer's complete, DataSync can also verify just the data copied or the entire dataset in the source and destination

locations. Depending on how you [configure data verification](#), this can take a significant amount of time on large datasets.

** Tip**

In most cases, we recommend verifying only the data that gets transferred.

DataSync checks data integrity by calculating and comparing the checksum and metadata of every file or object in both locations. If DataSync notices differences between locations, task verification fails with an error that specifies what failed. For example, you might see errors such as `Checksum failure`, `Metadata failure`, `Files were added`, `Files were removed`, and so on.

## Recurring transfers

In addition to one-time transfers, DataSync can move data on a recurring basis. Some of the options for these situations include:

- [Scheduling](#) when your task executes.
- Transferring [only the data that's changed](#) since the previous task execution.
- [Deleting data in the destination location](#) that's no longer present in the source.

# Setting up with AWS DataSync

Before you get started with AWS DataSync, you need to sign up for an AWS account if you don't have one. We also recommend learning where DataSync can be used and how much it might cost to move your data.

## Sign up for an AWS account

If you do not have an AWS account, complete the following steps to create one.

### To sign up for an AWS account

1. Open <https://portal.aws.amazon.com/billing/signup>.
2. Follow the online instructions.

Part of the sign-up procedure involves receiving a phone call and entering a verification code on the phone keypad.

When you sign up for an AWS account, an *AWS account root user* is created. The root user has access to all AWS services and resources in the account. As a security best practice, assign administrative access to a user, and use only the root user to perform [tasks that require root user access](#).

AWS sends you a confirmation email after the sign-up process is complete. At any time, you can view your current account activity and manage your account by going to <https://aws.amazon.com/> and choosing **My Account**.

## Create a user with administrative access

After you sign up for an AWS account, secure your AWS account root user, enable AWS IAM Identity Center, and create an administrative user so that you don't use the root user for everyday tasks.

### Secure your AWS account root user

1. Sign in to the [AWS Management Console](#) as the account owner by choosing **Root user** and entering your AWS account email address. On the next page, enter your password.

For help signing in by using root user, see [Signing in as the root user](#) in the *AWS Sign-In User Guide*.

2. Turn on multi-factor authentication (MFA) for your root user.

For instructions, see [Enable a virtual MFA device for your AWS account root user \(console\)](#) in the *IAM User Guide*.

## Create a user with administrative access

1. Enable IAM Identity Center.

For instructions, see [Enabling AWS IAM Identity Center](#) in the *AWS IAM Identity Center User Guide*.

2. In IAM Identity Center, grant administrative access to a user.

For a tutorial about using the IAM Identity Center directory as your identity source, see [Configure user access with the default IAM Identity Center directory](#) in the *AWS IAM Identity Center User Guide*.

## Sign in as the user with administrative access

- To sign in with your IAM Identity Center user, use the sign-in URL that was sent to your email address when you created the IAM Identity Center user.

For help signing in using an IAM Identity Center user, see [Signing in to the AWS access portal](#) in the *AWS Sign-In User Guide*.

## Assign access to additional users

1. In IAM Identity Center, create a permission set that follows the best practice of applying least-privilege permissions.

For instructions, see [Create a permission set](#) in the *AWS IAM Identity Center User Guide*.

2. Assign users to a group, and then assign single sign-on access to the group.

For instructions, see [Add groups](#) in the *AWS IAM Identity Center User Guide*.

# Required IAM permissions for using DataSync

DataSync can move your data to an Amazon S3 bucket, Amazon EFS file system, or a [number of other AWS storage services](#). To get your data where you want it to go, you need the right IAM permissions granted to your identity. For example, the IAM role that you use with DataSync needs permission to use the Amazon S3 operations required to transfer data to an S3 bucket.

You can grant these permissions with IAM policies provided by AWS or by creating your own policies.

## Contents

- [AWS managed policies](#)
- [Customer managed policies](#)

## AWS managed policies

AWS provides the following managed policies for common DataSync use cases:

- `AWSDataSyncReadOnlyAccess` – Provides read-only access to DataSync.
- `AWSDataSyncFullAccess` – Provides full access to DataSync and minimal access to its dependencies.

For more information, see [AWS managed policies for AWS DataSync](#).

## Customer managed policies

You can create custom IAM policies to use with DataSync. For more information, see [IAM customer managed policies for AWS DataSync](#).

## Where can I use DataSync?

For a list of AWS Regions and endpoints that DataSync supports, see [AWS DataSync endpoints and quotas](#) in the *AWS General Reference*.

## How can I use DataSync?

There are several ways to use DataSync:

- [DataSync console](#), which is part of the AWS Management Console.
- [DataSync API](#) or the [AWS CLI](#) to programmatically configure and manage DataSync.
- [AWS CloudFormation](#) or [Terraform](#) to provision your DataSync resources.
- [AWS SDKs](#) to build applications that use DataSync.

## How much will DataSync cost?

On the [DataSync pricing](#) page, create a custom estimate using the amount of data that you plan to transfer.

## Do I need an AWS DataSync agent?

To use AWS DataSync, you might need an agent. An *agent* is a virtual machine (VM) appliance that you deploy in your storage environment for data transfers or storage discovery.

Whether you need an agent depends on several factors, including the type of storage you're transferring to or from, if you're transferring across AWS accounts, and which AWS Regions you're transferring between. Before reading further, [check that DataSync supports the transfer you're interested in](#).

Once you know that DataSync supports your transfer, review the following information to help you understand if you need an agent.

## Situations when you need a DataSync agent

Most situations that require a DataSync agent involve storage that's managed by you or another cloud provider.

- Transferring to or from on-premises storage
- Transferring to or from other cloud storage
- Transferring to or from edge storage
- Transferring between some AWS storage services across AWS accounts (when neither storage service is Amazon S3)

For more information, see [Supported transfers across AWS accounts](#).

- Transferring between a commercial AWS Region and AWS GovCloud (US) Region
- Transferring between AWS storage services when at least one storage location is in an opt-in Region
- Using AWS DataSync Discovery

## Situations when you don't need a DataSync agent

The situations that don't require an agent apply whether you're transferring in the [same AWS Region](#) or [across Regions](#).

- Transferring between AWS storage services in the same AWS account
- Transferring between an S3 bucket and a different AWS storage service across AWS accounts

## Using multiple DataSync agents

You can use more than one DataSync agent with your data transfers. While most transfers only need one agent, using multiple agents can speed up transfers of large datasets with millions of files.

In these situations, we recommend running transfer tasks in parallel. This approach spreads out the transfer workload across multiple tasks (each of which uses its own agent). It also helps reduce the time it takes DataSync to prepare and transfer your data.

Another option—especially if you have millions of small files—is using multiple agents with a transfer location. For example, you can connect up to four agents to your on-premises Network File System (NFS) file service. This option can speed up your transfer, though the time it takes DataSync to prepare the transfer doesn't change.

With either approach, be mindful that these can increase the I/O operations on your storage system and affect your network bandwidth. For more information on using multiple agents for your DataSync transfers, see the [AWS Storage Blog](#).

If you're thinking of using multiple agents, remember the following:

- Using multiple agents with a location doesn't provide high availability. All the agents associated with a location must be online before you can start your transfer task. If one of the agents is [offline](#), you can't run your task.



- If you're [using a virtual private cloud \(VPC\) service endpoint](#) to communicate with the DataSync service, all the agents must use the same endpoint and subnet.
- With DataSync Discovery, you can only use [one agent per storage system](#).

## Next steps

- If you need an agent, review the [agent requirements](#) to understand what makes sense for your storage environment and what you need DataSync for.
- If you don't need an agent for your transfer, you can start [configuring your transfer](#).

## Requirements for AWS DataSync agents

Before you [deploy](#) an AWS DataSync agent in your storage environment, make sure that you understand the agent hypervisor and resource requirements.

### Hypervisor requirements

You can run a DataSync agent on the following hypervisors:

- **VMware ESXi (version 6.5, 6.7, 7.0, or 8.0):** VMware ESXi is available on the [VMware website](#). You also need a VMware vSphere client to connect to the host.

For information about VMware's supported hypervisor versions, see [VMware lifecycle policy](#) on the VMware website.

- **Microsoft Hyper-V (version 2012 R2, 2016, or 2019):** For this setup, you need a Microsoft Hyper-V Manager on a Microsoft Windows client computer to connect to the host.

The DataSync agent virtual machine (VM) is a generation 1 virtual machine. For more information about the differences between generation 1 and generation 2 VMs, see [Should I create a generation 1 or 2 virtual machine in Hyper-V?](#)

- **Linux Kernel-based Virtual Machine (KVM):** A free, open-source virtualization technology. KVM is included in Linux versions 2.6.20 and newer. DataSync is tested and supported for the CentOS/RHEL 7 and 8, Ubuntu 16.04 LTS, and Ubuntu 18.04 LTS distributions. Other modern Linux distribution might work, but function or performance is not guaranteed. We recommend this option if you already have a KVM environment up and running and you're already familiar with how KVM works.

Running KVM on Amazon EC2 isn't supported, and cannot be used for DataSync agents. To run the agent on Amazon EC2, deploy an agent Amazon Machine Image (AMI). For more information about deploying an agent AMI on Amazon EC2, see [Deploying your Amazon EC2 agent](#).

- **Amazon EC2:** DataSync provides an Amazon Machine Image (AMI) that contains the DataSync image. For the recommended instance types, see [Amazon EC2 instance requirements](#).

## Agent requirements for DataSync transfers

For DataSync transfers, your agent must meet the following resource requirements.

### Important

Keep in mind that the agent requirements for working with up to 20 million files, objects, or directories are general guidelines. Your agent may need more resources because of other factors, such as how many directories you have and object metadata size. For example, the m5.2xlarge instance for an Amazon EC2 agent still might not be enough for a transfer of less than 20 million files.

### Contents

- [Virtual machine requirements](#)
- [Amazon EC2 instance requirements](#)

## Virtual machine requirements

When deploying a DataSync agent that isn't on an Amazon EC2 instance, the agent VM requires the following resources:

- **Virtual processors:** Four virtual processors assigned to the VM.
- **Disk space:** 80 GB of disk space for installing the VM image and system data.
- **RAM:** Depending on your transfer scenario, you need the following amount of memory:
  - 32 GB of RAM assigned to the VM for task executions working with up to 20 million files, objects, or directories.
  - 64 GB of RAM assigned to the VM for task executions working with more than 20 million files, objects, or directories.

## Amazon EC2 instance requirements

When deploying a DataSync agent on an Amazon EC2 instance, the instance size must be at least 2xlarge.

We recommend using one of the following instance sizes:

- **m5.2xlarge**: For task executions working with up to 20 million files, objects, or directories.
- **m5.4xlarge**: For task executions working with more than 20 million files, objects, or directories.
- **snc1.medium**: For agents running on an AWS Snowcone device. This instance provides two CPU cores and 4 GiB of memory.

## Agent requirements for DataSync Discovery

Whether it's a VM or Amazon EC2 instance, the agent that you use with DataSync Discovery must have 80 GB of disk space and 16 GB of RAM.

## Agent requirements for partitions

DataSync agent images are associated with specific [partitions](#). For example, by default you can't download an agent in a commercial AWS Region and then activate it in an AWS GovCloud (US) Region.

## Agent management requirements

Once you [activate](#) your DataSync agent, AWS manages the agent for you. For more information, see [Managing your AWS DataSync agent](#).

## Deploying your AWS DataSync agent

The first step in creating your AWS DataSync agent is deploying the agent in your storage environment. You can deploy an agent as a virtual machine (VM) on VMware ESXi, Linux Kernel-based Virtual Machine (KVM), and Microsoft Hyper-V hypervisors. You also can deploy an agent as an Amazon EC2 instance in a virtual private cloud (VPC) within AWS.

### Tip

We recommend using separate agents for DataSync Discovery and DataSync transfers.

## Deploying your agent on VMware

You can download an agent from the DataSync console and deploy it in your VMware environment.

**Before you begin:** Make sure that your storage environment can support a DataSync agent. For more information, see [Virtual machine requirements](#).

### To deploy an agent on VMware

1. Open the AWS DataSync console at <https://console.aws.amazon.com/datasync/>.
2. In the left navigation pane, choose **Agents**, and then choose **Create agent**.
3. For **Hypervisor**, choose **VMWare ESXi**, and then choose **Download the image**.

The agent downloads in a .zip file that contains an .ova image file.

4. To minimize network latency, deploy the agent as close as possible to the storage system that DataSync needs to access (the same local network if possible). For more information, see [Network requirements for on-premises, self-managed, other cloud, and edge storage](#).

If needed, see your hypervisor's documentation on how to deploy an .ova file in a VMware host.

5. Power on your hypervisor, log in to the agent VM, and get the agent's IP address. You need this IP address to activate the agent.

The agent VM's default credentials are login **admin** and password **password**. If needed, change the password through the [VM's local console](#).

**Next step:** [Choosing a service endpoint for your AWS DataSync agent](#)

## Deploying your agent on KVM

You can download an agent from the DataSync console and deploy it in your KVM environment.

**Before you begin:** Make sure that your storage environment can support a DataSync agent. For more information, see [Virtual machine requirements](#).

### To deploy an agent on KVM

1. Open the AWS DataSync console at <https://console.aws.amazon.com/datasync/>.

2. In the left navigation pane, choose **Agents**, and then choose **Create agent**.
3. For **Hypervisor**, choose **Kernel-based Virtual Machine (KVM)**, and then choose **Download the image**.

The agent downloads in a .zip file that contains a .qcow2 image file.

4. To minimize network latency, deploy the agent as close as possible to the storage system that DataSync needs to access (the same local network if possible). For more information, see [Network requirements for on-premises, self-managed, other cloud, and edge storage](#).
5. Run the following command to install your .qcow2 image.

```
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
```

For information about how to manage this VM and your KVM host, see your hypervisor's documentation.

6. Power on your hypervisor, log in to your VM, and get the IP address of the agent. You need this IP address to activate the agent.

The agent VM's default credentials are login **admin** and password **password**. If needed, change the password through the [VM's local console](#).

**Next step:** [Choosing a service endpoint for your AWS DataSync agent](#)

## Deploying your agent on Microsoft Hyper-V

You can download an agent from the DataSync console and deploy it in your Microsoft Hyper-V environment.

**Before you begin:** Make sure that your storage environment can support a DataSync agent. For more information, see [Virtual machine requirements](#).

## To deploy an agent on Hyper-V

1. Open the AWS DataSync console at <https://console.aws.amazon.com/datasync/>.
2. In the left navigation pane, choose **Agents**, and then choose **Create agent**.
3. For **Hypervisor**, choose **Microsoft Hyper-V**, and then choose **Download the image**.

The agent downloads in a .zip file that contains a .vhdx image file.

4. To minimize network latency, deploy the agent as close as possible to the storage system that DataSync needs to access (the same local network if possible). For more information, see [Network requirements for on-premises, self-managed, other cloud, and edge storage](#).

If needed, see your hypervisor's documentation on how to deploy a .vhdx file in a Hyper-V host.

### **Warning**

You may notice poor network performance if you enable virtual machine queue (VMQ) on a Hyper-V host that's using a Broadcom network adapter. For information about a workaround, see the [Microsoft documentation](#).

5. Power on your hypervisor, log in to your VM, and get the IP address of the agent. You need this IP address to activate the agent.

The agent VM's default credentials are login **admin** and password **password**. If needed, change the password through the [VM's local console](#).

**Next step:** [Choosing a service endpoint for your AWS DataSync agent](#)

## Deploying your Amazon EC2 agent

You might deploy a DataSync agent as an Amazon EC2 instance when transferring data between:

- A self-managed cloud storage system (for example, an NFS file server in AWS) and an AWS storage service.
- A cloud storage provider (such as Microsoft Azure Blob Storage or Google Cloud Storage) and an AWS storage service.
- An S3 bucket in a commercial AWS Region and an S3 bucket in an AWS GovCloud (US) Region.

- [Amazon S3 on AWS Outposts](#) and an AWS storage service.

### Warning

We don't recommend using an Amazon EC2 agent with on-premises storage because of increased network latency. Instead, deploy the agent as a VMware, KVM, or Hyper-V virtual machine in your data center as close to your on-premises storage as possible.

## Deploying your EC2 agent

### To choose the agent AMI for your AWS Region

1. Open a terminal and copy the following AWS CLI command to get the latest DataSync Amazon Machine Image (AMI) ID for the Region where you want to deploy your Amazon EC2 agent.

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

2. Run the command. In the output, take note of the "Value" property with the DataSync AMI ID.

### Example Example command and output

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

{
  "Parameter": {
    "Name": "/aws/service/datasync/ami",
    "Type": "String",
    "Value": "ami-1234567890abcdef0",
    "Version": 6,
    "LastModifiedDate": 1569946277.996,
    "ARN": "arn:aws:ssm:us-east-1::parameter/aws/service/datasync/ami"
  }
}
```

## To deploy your Amazon EC2 agent

### Tip

To avoid charges for transferring across Availability Zones, deploy your agent in a way that it doesn't require network traffic between Availability Zones. (To learn more about data transfer prices for all AWS Regions, see [Amazon EC2 Data Transfer pricing](#).) For example, deploy your agent in the Availability Zone where your self-managed cloud storage system is located.

1. Copy the following URL:

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

- Replace *agent-region* with the Region where you want to deploy your agent.
  - Replace *ami-id* with the DataSync AMI ID that you obtained.
2. Paste the URL into a browser.

The Amazon EC2 instance launch page in the AWS Management Console displays.
  3. For **Instance type**, choose one of the [recommended Amazon EC2 instances](#) for DataSync.
  4. For **Key pair**, choose an existing key pair, or create a new one.
  5. For **Network settings**, choose **Edit** and then do the following:
    - a. For **VPC**, choose a VPC where you want to deploy your agent.
    - b. For **Auto-assign public IP**, choose whether you want your agent to be accessible from the public internet.

You use the instance's public or private IP address later to activate your agent.

- c. For **Firewall (security groups)**, create or select a security group that does the following:
  - Allows inbound traffic from your browser to the Amazon EC2 instance on port 80 (HTTP).
  - Allows inbound and outbound traffic between the Amazon EC2 instance the storage system that you're transferring data to or from. For more information, see [Network requirements for on-premises, self-managed, other cloud, and edge storage](#).



**Note**

There are additional ports to configure depending on the type of [service endpoint](#) that your agent uses.

6. (Recommended) To increase performance when transferring from a cloud-based file system, expand **Advanced details** and choose a **Placement group** value where your storage is located.
7. Choose **Launch instance** to launch your Amazon EC2 instance.
8. Once your instance status is **Running**, choose the instance.
9. If you configured your instance to be accessible from the public internet, make note of the instance's public IP address. If you didn't, make note of the private IP address.

You need this IP address when [activating your agent](#).

## Examples: Deploying your EC2 agent in an AWS Region

The following guidance can help with common scenarios if you deploy an DataSync agent in an AWS Region.

### Topics

- [Deploying your agent for transfers between cloud file systems and Amazon S3](#)
- [Deploying your agent for transfers between Amazon S3 to AWS file systems](#)

## Deploying your agent for transfers between cloud file systems and Amazon S3

To transfer data between AWS accounts, or from a cloud file system, the DataSync agent must be located in the same AWS Region and AWS account where the source file system resides. This type of transfer includes the following:

- Transfers between Amazon EFS or FSx for Windows File Server file systems to AWS storage in a different AWS account.
- Transfers from self-managed file systems to AWS storage services.

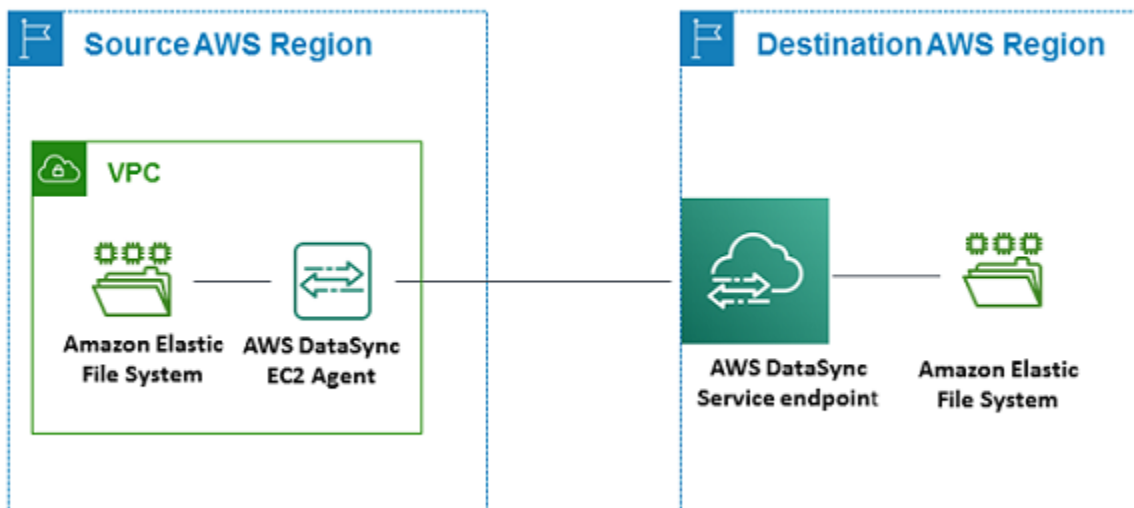
### ⚠ Important

Deploy your agent such that it doesn't require network traffic between Availability Zones (to avoid charges for such traffic).

- To access your Amazon EFS or FSx for Windows File Server file system, deploy the agent in an Availability Zone that has a mount target to your file system.
- For self-managed file systems, deploy the agent in the Availability Zone where your file system resides.

To learn more about data transfer prices for all AWS Regions, see [Amazon EC2 On-Demand pricing](#).

For example, the following diagram shows a high-level view of the DataSync architecture for transferring data from in-cloud Network File System (NFS) to in-cloud NFS or Amazon S3.

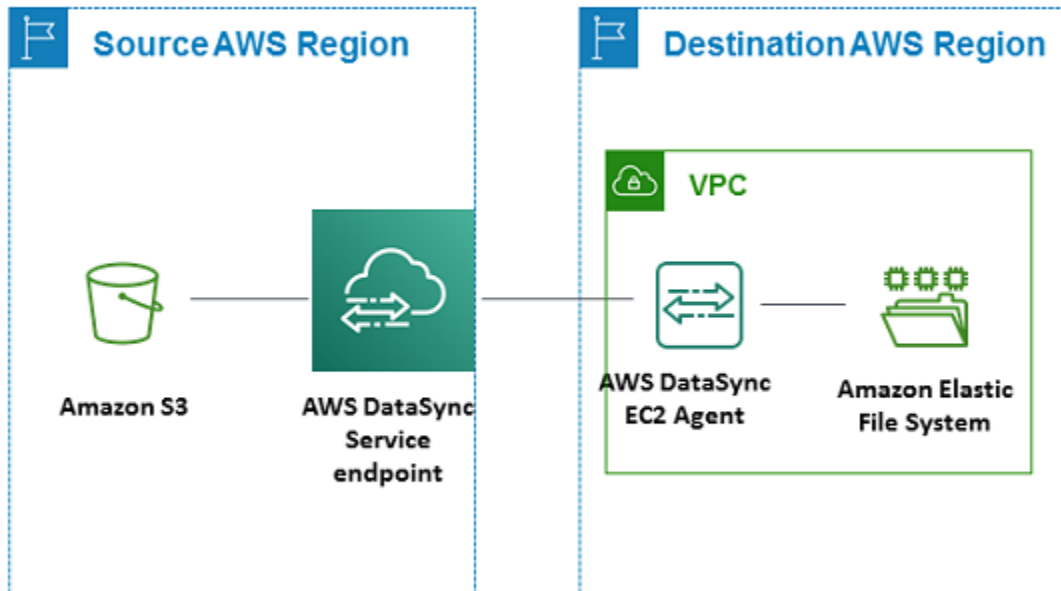


Remember the following when transferring between AWS storage services across AWS accounts:

- When transferring between Amazon EFS file systems, configure your source file system as an [NFS location](#).
- When transferring between Amazon FSx for Windows File Server file systems, configure your source file system as an [SMB location](#).

## Deploying your agent for transfers between Amazon S3 to AWS file systems

The following diagram provides a high-level view of the DataSync architecture for transferring data from Amazon S3 to an AWS file system, such as Amazon EFS or Amazon FSx. You can use this architecture to transfer data from one AWS account to another, or to transfer data from Amazon S3 to a self-managed in-cloud file system.



## Deploying your agent on AWS Snowball Edge

For more information and instructions, see [Creating a DataSync agent in your on-premises storage environment](#) for Amazon S3 compatible storage.

## Deploying your agent on AWS Snowcone

The DataSync agent AMI is pre-installed on your Snowcone device. Launch the agent with one of the following tools:

- [AWS OpsHub](#)
- [Snowball Edge client](#)

Once your agent is deployed and activated, you can [configure your transfer location](#).

## Deploying your agent on AWS Outposts

You can launch a DataSync Amazon EC2 instance on your Outpost. To learn more about launching an AMI on AWS Outposts, see [Launch an instance on your Outpost](#) in the *AWS Outposts User Guide*.

When using DataSync to access Amazon S3 on Outposts, you must launch the agent in a VPC that's allowed to access your Amazon S3 access point, and activate the agent in the parent Region of the Outpost. The agent must also be able to route to the Amazon S3 on Outposts endpoint for the bucket. To learn more about working with Amazon S3 on Outposts endpoints, see [Working with Amazon S3 on Outposts](#) in the *Amazon S3 User Guide*.

## Choosing a service endpoint for your AWS DataSync agent

The second step in creating your AWS DataSync agent is choosing a service endpoint for the agent. A *service endpoint* is how your [agent communicates with the DataSync service](#). An agent can use the following types of service endpoints:

- **Virtual private cloud (VPC) endpoint** – Data is sent through your VPC instead of over the public internet, increasing the security of the transferred data.
- **Public endpoint** – Data is sent over the public internet.
- **Federal Information Processing Standard (FIPS) endpoint** – Data is sent over the public internet by using processes that comply with FIPS.

Remember the following when choosing a service endpoint:

- An agent can only use one type of endpoint. If you need to transfer data with different endpoint types, create an agent for each type.
- How you [connect your storage network to AWS](#) determines what service endpoints you can use.
- For DataSync Discovery, you can only use a public endpoint.

For more information, see [AWS service endpoints](#) in the *AWS General Reference*.

## Using a VPC service endpoint with DataSync

With a VPC service endpoint, you don't have to transfer your data across the public internet. DataSync can transfer data to AWS through a VPC that's based on the Amazon VPC service.

## Contents

- [How DataSync agents work with VPC service endpoints](#)
- [DataSync limitations with VPCs](#)
- [Configuring your DataSync agent to use a VPC service endpoint](#)

## How DataSync agents work with VPC service endpoints

VPC service endpoints are provided by AWS PrivateLink. These types of endpoints let you privately connect supported AWS services to your VPC. When you use a VPC service endpoint with DataSync, all communication between your DataSync agent and AWS remains in your VPC. DataSync also creates [network interfaces](#) in your VPC for data transfer traffic.

These interfaces and the VPC service endpoint are private IP addresses that are only accessible from inside your VPC. For more information, see [Networking with AWS DataSync](#).

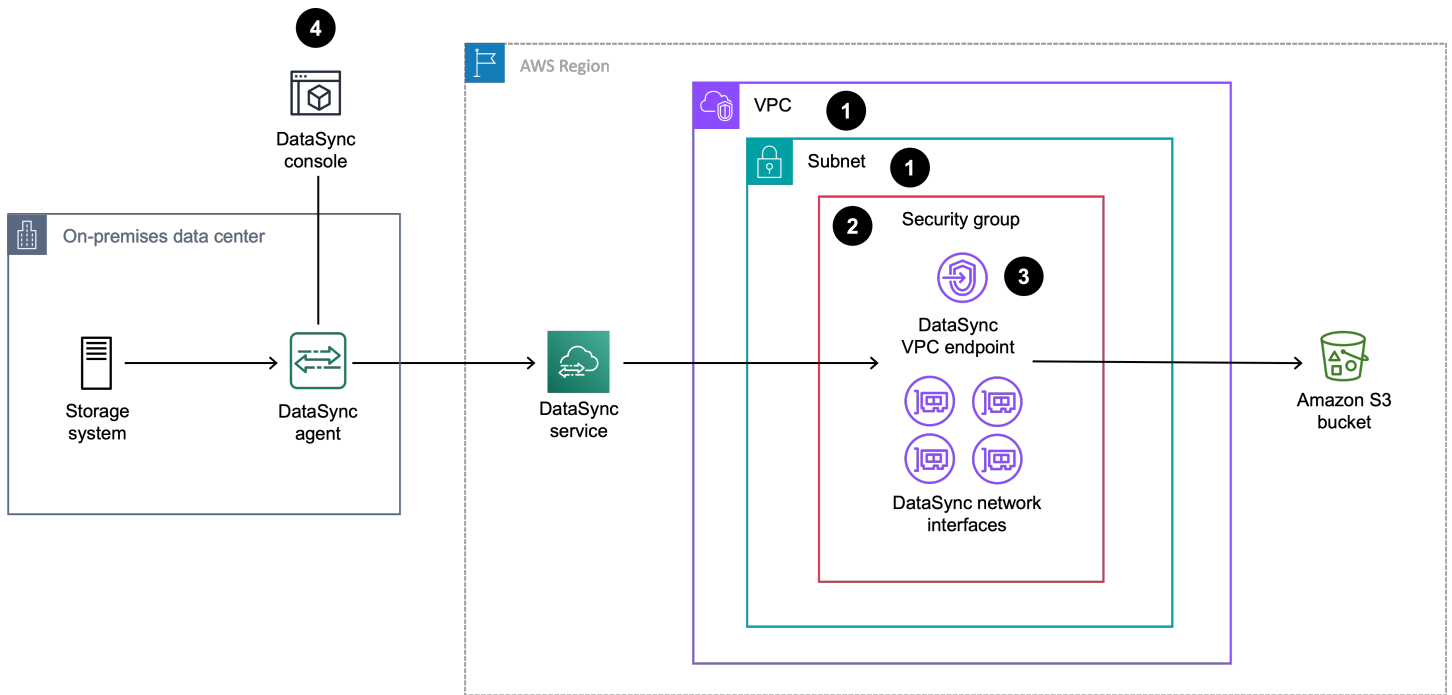
## DataSync limitations with VPCs

- VPCs that you use with DataSync must have default tenancy. VPCs with dedicated tenancy are not supported. For more information, see [Work with VPCs](#).
- DataSync doesn't support [shared VPCs](#).
- DataSync VPC service endpoints only support IPv4. [IPv6 and dualstack options](#) aren't supported.

## Configuring your DataSync agent to use a VPC service endpoint

After [deploying your DataSync agent](#), configure the agent to use a VPC service endpoint.

The following diagram shows an example DataSync transfer from an on-premises storage system to an Amazon S3 bucket. The numbered callouts correspond to the steps to configure your DataSync agent to use a VPC service endpoint.



## To configure your DataSync agent to use a VPC service endpoint

1. [Create](#) or determine a VPC and subnet where you want to create your VPC service endpoint.

If you're transferring to or from storage that's outside AWS, the VPC should extend to that storage environment (for example, the data center where your on-premises NFS file server is located). You can do this by using routing rules over [AWS Direct Connect](#) or VPN.

2. In your VPC, configure a [security group](#) that allows the traffic required for using DataSync VPC service endpoints.

The security group must allow your agent to connect with the private IP addresses of the VPC service endpoint and your task's [network interfaces](#). For specific ports and protocols, see [Network requirements for VPC service endpoints](#).

3. Create a DataSync VPC service endpoint by doing the following:
  - a. Open the Amazon VPC console at <https://console.aws.amazon.com/vpc/>.
  - b. In the left navigation pane, choose **Endpoints**, then choose **Create endpoint**.
  - c. For **Service category**, choose **AWS services**.
  - d. For **Services**, search for **datasync** and choose the endpoint for the Region you're in (for example, `com.amazonaws.us-east-1.datasync`).
  - e. For **VPC**, choose the VPC where you want to create the VPC service endpoint.

- f. Expand **Additional settings** and clear the **Enable Private DNS Name** check box.
- g. For **Subnet**, choose the subnet where you want to create the VPC service endpoint.
- h. Choose **Create endpoint**.

For more information on creating VPC service endpoints, see the [Amazon VPC User Guide](#).

4. Open the AWS DataSync console at <https://console.aws.amazon.com/datasync/>.

In the left navigation pane, choose **Agents**, and then choose **Create agent**. In the **Service endpoint** section, do the following:

- a. For **Endpoint type**, choose **VPC endpoints using AWS PrivateLink**.
- b. For **VPC endpoint**, choose the VPC service endpoint that your agent will use.
- c. For **Subnet**, choose the subnet where your VPC service endpoint is located.
- d. For **Security group**, choose the security group that you configured for allowing your agent to use your VPC service endpoint.

Next step: [Activating your AWS DataSync agent](#)

## Using a public service endpoint with DataSync

If you use a public service endpoint, all communication between your DataSync agent and DataSync service occurs over the public internet.

### To configure your DataSync agent to use a public service endpoint

1. After [deploying your DataSync agent](#), [configure your network](#) to allow the traffic required for using DataSync public service endpoints.
2. Open the AWS DataSync console at <https://console.aws.amazon.com/datasync/>.
3. In the left navigation pane, choose **Agents**, and then choose **Create agent**.
4. In the **Service endpoint** section, choose **Public service endpoints in *AWS Region name***.

For a list of endpoints that you can use, see [DataSync service endpoints](#) in the *AWS General Reference*.

Next step: [Activating your AWS DataSync agent](#)

## Using a FIPS service endpoint with DataSync

DataSync offers some service endpoints that comply with FIPS. For more information, see [FIPS endpoints](#) in the *AWS General Reference*.

### To configure your DataSync agent to use a FIPS service endpoint

1. After [deploying your DataSync agent](#), [configure your network](#) to allow the traffic required for using DataSync FIPS service endpoints.
2. Open the AWS DataSync console at <https://console.aws.amazon.com/datasync/>.
3. In the left navigation pane, choose **Agents**, and then choose **Create agent**.
4. In the **Service endpoint** section, choose **FIPS service endpoints in *AWS Region name***.

For a list of endpoints that you can use, see [DataSync service endpoints](#) in the *AWS General Reference*.

Next step: [Activating your AWS DataSync agent](#)

## Activating your AWS DataSync agent

To finish creating your AWS DataSync agent (assuming you have already deployed the agent and chose its service endpoint), you must activate it. This step associates the agent with your AWS account.

### Note

You can't activate an agent in more than one AWS account and AWS Region at a time.


### To activate your agent by using the DataSync console

1. On the same **Create agent** page, go to the **Activation key** section.
2. Choose one of the following options to activate your agent:
  - **Automatically get the activation key from your agent** – This option requires that your browser access the agent by using port 80. Once activated, the agent closes the port.
  - For **Agent address**, enter the agent's IP address or domain name and choose **Get key**.



Your browser connects to the IP address and gets a unique activation key from your agent. If the activation fails, [check your network configuration](#).

- **Manually enter your agent's activation key** – Use this option if you don't want a connection between your browser and agent.
  - Get the key from the [agent's local console](#).
  - Back in the DataSync console, enter the key in the **Activation key** field.


 **Note**

Agent activation keys expire in 30 minutes if unused.

3. (Optional) For **Agent name**, enter a name for your agent.
4. (Optional) For **Tags**, enter values for the **Key** and **Value** fields to tag your agent.

Tags help you manage, filter, and search for your AWS resources.

5. Choose **Create agent**.
6. On the **Agents** page, verify that your service endpoint is correct.

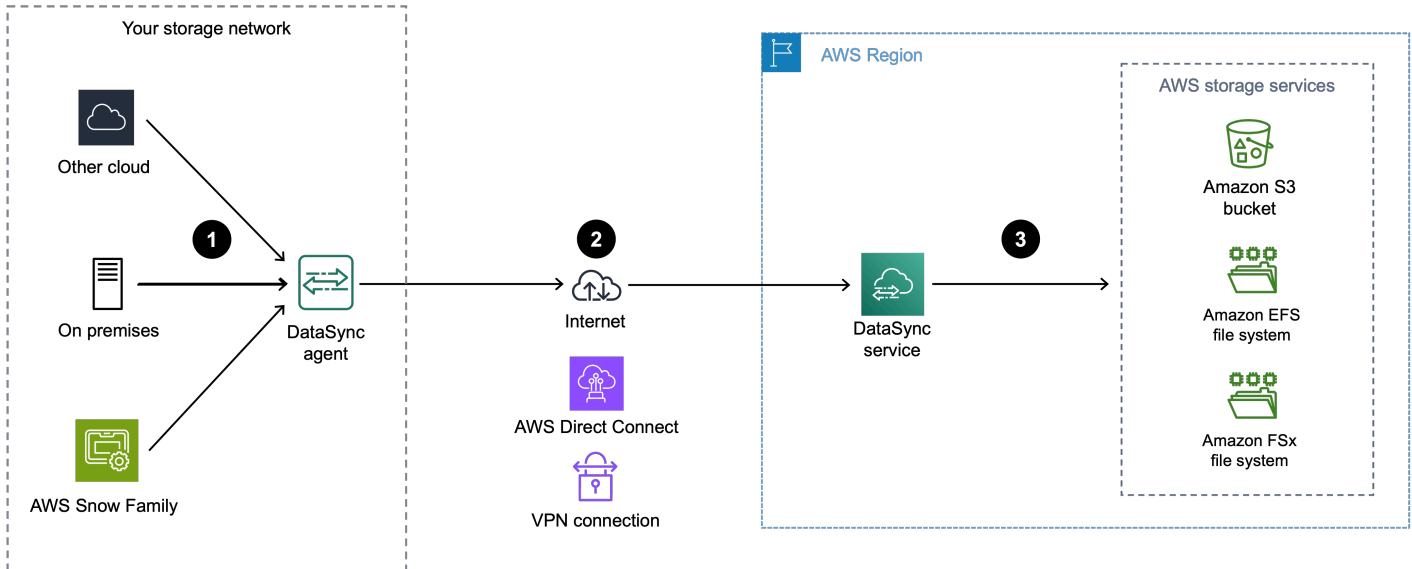
 **Note**

At this point, you might notice your agent is offline. This happens briefly after agent activation.

You're done creating your agent, which [AWS manages for you](#) going forward.

# Networking with AWS DataSync

If [you need an AWS DataSync agent](#), you must establish several network connections for a data transfer or storage discovery. The following diagram shows the three network connections in a DataSync transfer from a storage system (which could be on premises, in another cloud, or at the edge) to an AWS storage service.



## 1. Network connection between your storage system and agent

Your DataSync agent connects to your on-premises, other cloud, or edge storage system. For more information, see [Network requirements for on-premises, self-managed, other cloud, and edge storage](#).

## 2. Network connection between your agent and DataSync service

There are a few aspects to connecting your agent to the DataSync service. First, you must connect your storage network to AWS. Second, your agent needs a service endpoint to communicate with DataSync.

### Contents

- [Connecting your storage network to AWS](#)

- [Choosing a service endpoint](#)

## Connecting your storage network to AWS

When using DataSync, consider the following options for connecting your storage network to AWS:

- **AWS Direct Connect** - With [Direct Connect](#), you can create a dedicated connection between your storage network and AWS. From a DataSync perspective, this lets you:
  - Transfer data over a private path to your virtual private cloud (VPC), which avoids routing over the public internet.
  - Get a more predictable connection than using a virtual private network (VPN) to connect your storage network to AWS (particularly if your agent is an Amazon EC2 instance).
  - Use any type of DataSync service endpoint, including [VPC](#), [public](#), or [Federal Information Processing Standard \(FIPS\)](#) endpoints.

For more information, see [AWS DataSync architecture and routing examples with AWS Direct Connect](#).

- **VPN** - You can connect your storage network to AWS by using a VPN (such as [AWS Site-to-Site VPN](#)).
- **Public internet** - You can connect your storage network directly to DataSync over the internet by using a [public](#) or [FIPS](#) service endpoint.

## Choosing a service endpoint

Your agent uses a service endpoint to communicate with DataSync. For more information, see [Choosing a service endpoint for your AWS DataSync agent](#).

## 3. Network connection between DataSync service and AWS storage service

To connect DataSync to an AWS storage service, you just have to make sure that the DataSync service can access your S3 bucket or file system. For more information, see [Network requirements for AWS storage services](#).

## Networking when you don't need a DataSync agent

For transfers that [don't require a DataSync agent](#), you just have to make sure that the DataSync service can access the AWS storage services you're transferring between. For more information, see [Network requirements for AWS storage services](#).

## How and where DataSync traffic flows through the network

DataSync has *data plane* and *control plane* traffic. Knowing how each of these flows through the network is important if you want to separate your DataSync traffic.

- **Data plane traffic** – Includes the file or object data moving between your storage locations. In most cases, data plane traffic routes through [network interfaces](#) that DataSync automatically generates and manages when you create a task. Where these network interfaces get created depends on the type of AWS storage service you're transferring to or from and the service endpoint that your DataSync agent uses.
- **Control plane traffic** – Includes management activities for your DataSync resources. This traffic routes through the service endpoint that your agent uses.

## Network security for DataSync

For information about how your storage data (including metadata) is secured during a transfer, see [AWS DataSync encryption in transit](#).

## AWS DataSync network requirements

Configuring your network is an important step in setting up AWS DataSync. Your network configuration depends on several factors, such as whether you want information about your storage or are ready to transfer data. It's also based on what kind of service endpoint you plan to use for sending data to AWS.

## Network requirements for on-premises, self-managed, other cloud, and edge storage

The following network requirements can apply to on-premises, self-managed, other cloud, and edge storage systems. These are typically storage systems that you manage or might be managed by another cloud provider.

**Note**

Depending on your network, you might need to allow traffic on ports other than what's listed here for your DataSync agent to connect with your storage.

From	To	Protocol	Port	How it's used by DataSync
DataSync agent	NFS file server	TCP	2049	Mounts the NFS file server.  DataSync supports NFS versions 3.x, 4.0, and 4.1.
DataSync agent	SMB file server	TCP	139 or 445	Mounts the SMB file server.  DataSync supports SMB versions 1.0 and later.
DataSync agent	Object storage	TCP	443 (HTTPS) or 80 (HTTP)	Accesses your object storage.
DataSync agent	Hadoop cluster	TCP	NameNode port (default is 8020)  In most clusters, you can find this port number in the core-site	Accesses the NameNodes in your Hadoop cluster. Specify the port used when creating an HDFS location.

From	To	Protocol	Port	How it's used by DataSync
			<p>.xml file under the fs.default or fs.default.name property (depending on the Hadoop distribution).</p>	
DataSync agent	Hadoop cluster	TCP	<p>DataNode port (default is 50010)</p> <p>In most clusters, you can find this port number in the hdfs-site.xml file under the dfs.datanode.address property.</p>	<p>Accesses the DataNodes in your Hadoop cluster. The DataSync agent automatically determines the port to use.</p>

From	To	Protocol	Port	How it's used by DataSync
DataSync agent	Hadoop Key Management Server (KMS)	TCP	KMS port (default is 9600)	Accesses the KMS for your Hadoop cluster.
DataSync agent	Kerberos Key Distribution Center (KDC) server	TCP	KDC port (default is 88)	Authenticates with the Kerberos realm. This port is used only with HDFS.
DataSync agent	Storage system's management interface	TCP	Depends on your network	Connects to your storage system. DataSync Discovery uses this connection to collect information about your system.

## Network requirements for AWS storage services

The network ports required for DataSync to connect to an AWS storage service during a transfer vary.


From	To	Protocol	Port
DataSync service	Amazon EFS	TCP	2049
DataSync service	FSx for Windows File Server	See <a href="#">file system access control for FSx for Windows File Server</a> .	
DataSync service	FSx for Lustre	See <a href="#">file system access control for FSx for Lustre</a> .	
DataSync service	FSx for OpenZFS	See <a href="#">file system access control for FSx for OpenZFS</a> .	
DataSync service	FSx for ONTAP	TCP	111, 635, and 2049 (NFS)

From	To	Protocol	Port
			445 (SMB)
DataSync service	Amazon S3	TCP	443 (HTTPS)

## Network requirements for VPC service endpoints

A virtual private cloud (VPC) endpoint provides a private connection between your agent and AWS that doesn't cross the internet or use public IP addresses. This also helps prevent packets from entering or exiting the network. For more information, see [Using a VPC service endpoint with DataSync](#).

DataSync requires the following ports for your agent to use a VPC service endpoint.

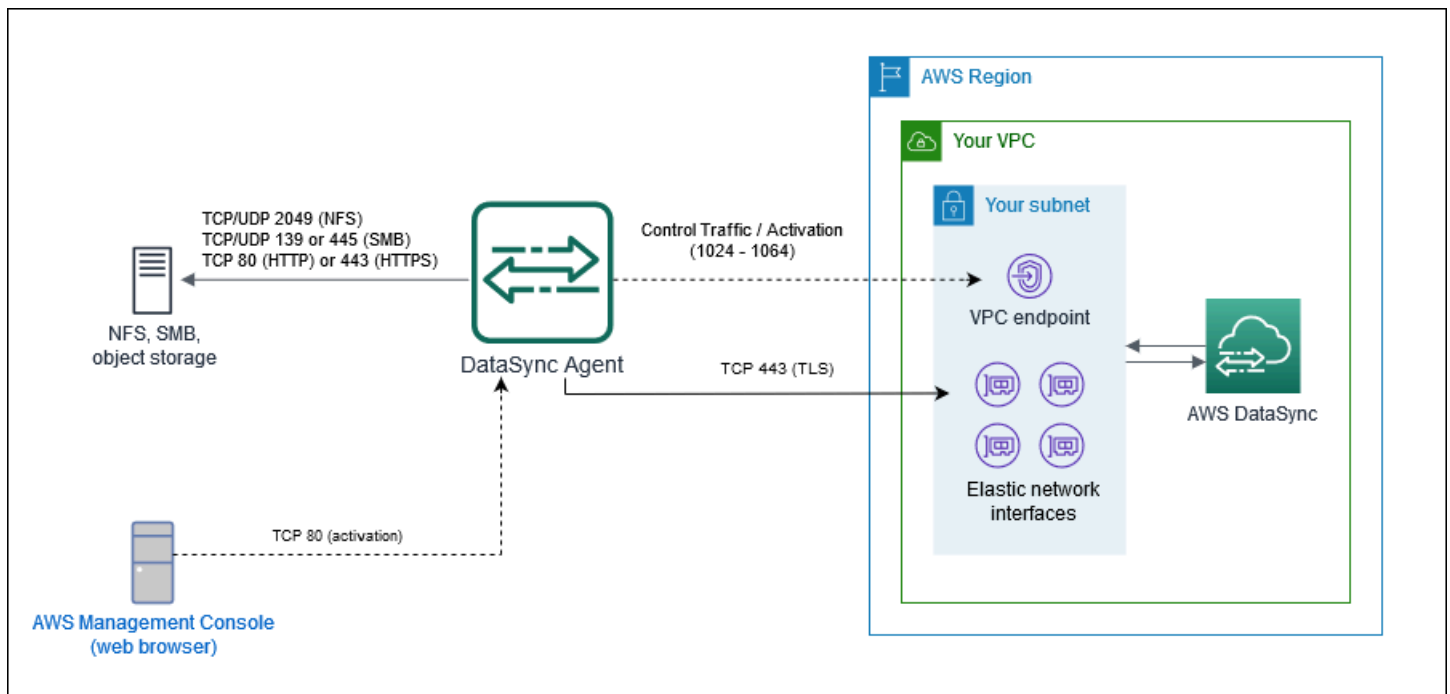
From	To	Protocol	Port	How it's used
Your web browser	Your DataSync agent	TCP	80 (HTTP)	<p>Allows your browser to obtain the agent activation key. Once activated, DataSync closes the agent's port 80.</p> <p>The DataSync agent doesn't require port 80 to be publicly accessible. The required level of access to port 80 depends on your network configuration.</p> <div data-bbox="1156 1640 1198 1677" style="float: left; margin-right: 5px;">  </div> <p><b>Note</b></p> <p>Alternatively, you can obtain the activation key from the agent's</p>



From	To	Protocol	Port	How it's used
				<p>local console. This doesn't require connectivity between your browser and agent. For more information, see <a href="#">Getting an agent activation key</a>.</p>
DataSync agent	<p>Your DataSync VPC service endpoint</p> <p>To find the endpoint's IP address, open the <a href="#">Amazon VPC console</a>, and choose <b>Endpoints</b> from the left navigation pane. Choose the DataSync endpoint, and check the <b>Subnets</b> list to find the private IP address that corresponds to the subnet that you chose for your <a href="#">VPC service endpoint setup</a>. This is the endpoint's IP address.</p>	TCP	1024-1064	For <a href="#">control plane traffic</a> .


From	To	Protocol	Port	How it's used
DataSync agent	Your DataSync task's <a href="#">network interfaces</a>  To find the IP addresses of these interfaces, see <a href="#">Viewing your network interfaces</a> .	TCP	443 (HTTPS)	For <a href="#">data plane traffic</a> .
DataSync agent	Your DataSync VPC endpoint	TCP	22 (Support channel)	To allow AWS Support to access your DataSync agent for troubleshooting.  You don't need this port open for normal operation.

The following diagram shows the ports required by DataSync when using VPC endpoints.



## Network requirements for public or FIPS service endpoints

Your DataSync agent requires the following network access when using public or FIPS service endpoints. If you use a firewall or router to filter or limit network traffic, configure your firewall or router to allow these endpoints.

From	To	Protocol	Port	How it's used	Endpoints accessed by the agent
Your web browser	DataSync agent	TCP	80 (HTTP)	<p>Allows your browser to obtain the DataSync agent's activation key. Once activated, DataSync closes the agent's port 80.</p> <p>The agent doesn't require port 80 to be publicly accessible. The required level of access to port 80 depends on your network configuration.</p> <div data-bbox="732 1591 967 1871" style="border: 1px solid #ccc; border-radius: 10px; padding: 10px; background-color: #e6f2ff;"> <p> <b>Note</b></p> <p>Alternatively, you can obtain</p> </div>	N/A

From	To	Protocol	Port	How it's used	Endpoints accessed by the agent
				<p>the activation key from the agent's local console. This doesn't require connectivity between your browser and agent. For more information, see <a href="#">Getting an agent activation key</a>.</p>	

From	To	Protocol	Port	How it's used	Endpoints accessed by the agent
DataSync agent	Amazon CloudFront	TCP	443 (HTTPS)	Helps bootstrap your DataSync agent prior to activation.	<p><b>AWS Regions:</b></p> <ul style="list-style-type: none"> <li>d3dvvaliwoko8h.cloudfront.net</li> </ul> <p><b>AWS GovCloud (US) Regions:</b></p> <ul style="list-style-type: none"> <li>s3.us-gov-west-1.amazonaws.com/fmrse-ndpoints-endpoints-bucket-go4p5gpna6sk</li> </ul>
DataSync agent	AWS	TCP	443 (HTTPS)	Activates your DataSync agent and associates it with your AWS account. You can block the public endpoints after activation.	<p>The <i>activation-region</i> is the AWS Region where you activate your DataSync agent.</p> <p><b>Public endpoint activation:</b></p> <ul style="list-style-type: none"> <li>activation.datasyncc.<i>activation-region</i>.amazonaws.com</li> </ul> <p><b>FIPS endpoint activation:</b></p> <ul style="list-style-type: none"> <li>activation.datasyncc-fips.<i>activation-region</i>.amazonaws.com</li> </ul>

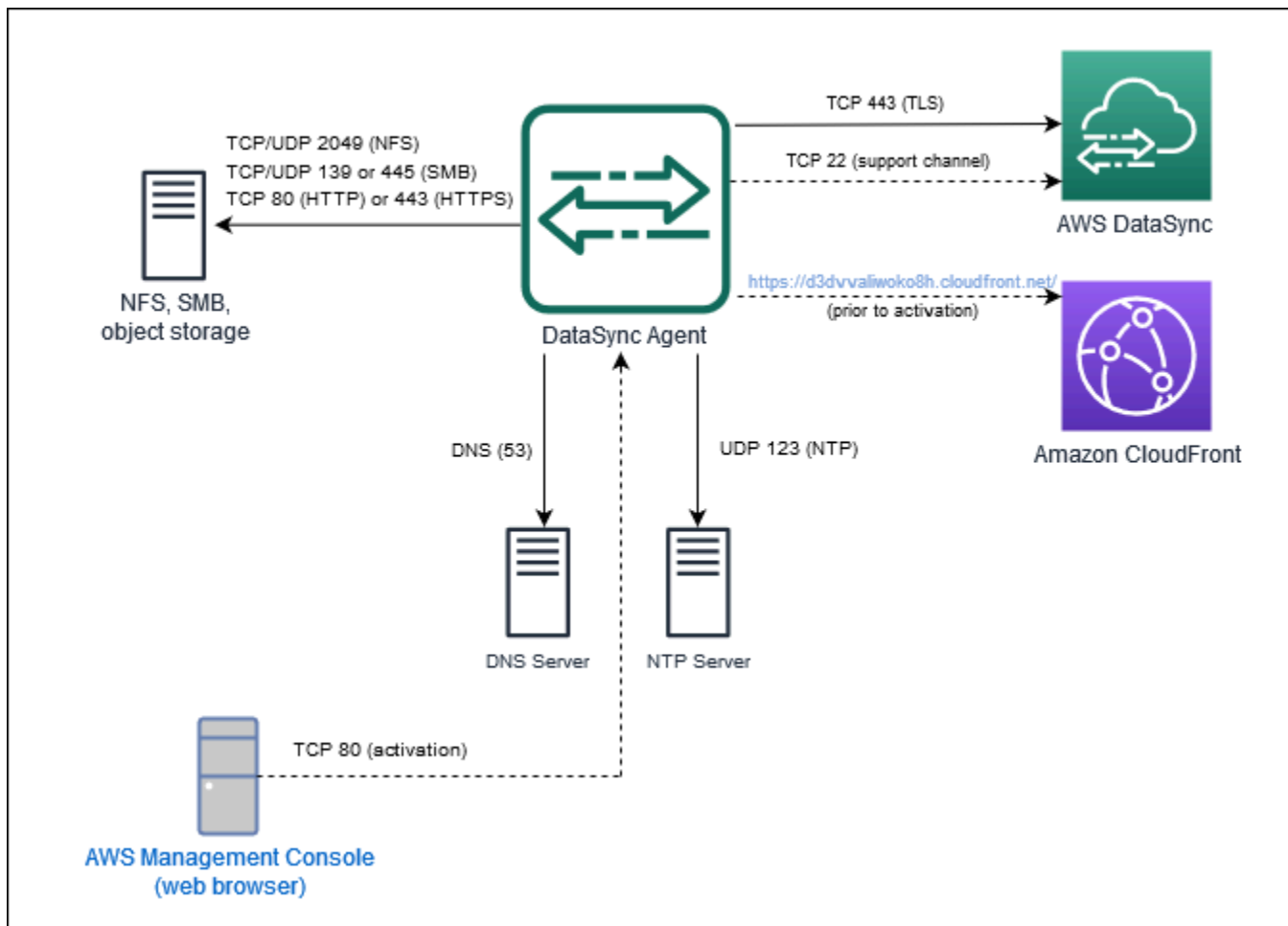
From	To	Protocol	Port	How it's used	Endpoints accessed by the agent
DataSync agent	AWS	TCP	443 (HTTPS)	<p>Allows communication between the DataSync agent and AWS service endpoint.</p> <p>For information, see <a href="#">Choosing a service endpoint for your AWS DataSync agent</a>.</p>	<p>The <i>activation-region</i> is the AWS Region where you activated your DataSync agent. Depending on what you're using DataSync for, you might not need to allow access to every endpoint listed here.</p> <p><b>DataSync API endpoint:</b></p> <ul style="list-style-type: none"> <li>• <code>datasync. <i>activation-region</i>.amazonaws.com</code></li> </ul> <p><b>DataSync Discovery endpoint (for discovery jobs only):</b></p> <ul style="list-style-type: none"> <li>• <code>discovery-datasync. <i>activation-region</i>.amazonaws.com</code></li> </ul> <p><b>DataSync control plane endpoints:</b></p> <ul style="list-style-type: none"> <li>• <b>Public endpoint:</b> <code>cp.datasync. <i>activation-region</i>.amazonaws.com</code></li> <li>• <b>FIPS endpoint:</b> <code>cp.datasync-fips. <i>activation-region</i>.amazonaws.com</code></li> </ul> <p><b>DataSync data plane endpoint (for transfer tasks only):</b></p>

From	To	Protocol	Port	How it's used	Endpoints accessed by the agent
					<ul style="list-style-type: none"> <li><code>your-task-id .datasync-dp. activation-region.amazonaws.com</code></li> </ul>
DataSync agent	AWS	TCP	443 (HTTPS)	Allows the DataSync agent to get updates from AWS. For more information, see <a href="#">Managing your AWS DataSync agent</a> .	<p>The <i>activation-region</i> is the AWS Region where you activated your DataSync agent.</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- activation-region .s3.dualstack. activation-region.amazonaws.com</code></li> <li><code>amazonlinux-2-repos- activation-region.s3. activation-region.amazonaws.com</code></li> <li><code>*.s3. activation-region.amazonaws.com</code></li> </ul>
DataSync agent	Domain Name Service (DNS) server	TCP/UDP	53 (DNS)	Allows communication between the DataSync agent and DNS server.	N/A

From	To	Protocol	Port	How it's used	Endpoints accessed by the agent
DataSync agent	AWS	TCP	22 (Support channel)	Allows AWS Support to access your DataSync agent to help you troubleshoot issues. You don't need this port open for normal operation.	<b>AWS Support channel:</b> <ul style="list-style-type: none"> <li>54.201.223.107</li> </ul>
DataSync agent	Network Time Protocol (NTP) server	UDP	123 (NTP)	Allows local systems to synchronize the VM time to the host time.	<b>NTP:</b> <ul style="list-style-type: none"> <li>0.amazon.pool.ntp.org</li> <li>1.amazon.pool.ntp.org</li> <li>2.amazon.pool.ntp.org</li> <li>3.amazon.pool.ntp.org</li> </ul> <div style="border: 1px solid #add8e6; border-radius: 10px; padding: 10px; margin-top: 10px;"> <p><b>Note</b></p> <p>To change the default NTP configuration of your VM agent to use a different NTP server using the local console, see <a href="#">Synchronizing the time on your VMware agent</a>.</p> </div>

The following diagram shows the ports required by DataSync when using public or FIPS endpoints.





## Network interfaces for AWS DataSync transfers

For every task you create, AWS DataSync automatically generates and manages [network interfaces](#) for data transfer traffic. How many network interfaces DataSync creates and where they're created depends on the following details about your transfer task:

- Whether your task [requires a DataSync agent](#).
- Your source and destination locations (where you're copying data from and to).
- The type of service endpoint that your agent uses.

Each network interface uses a single IP address in your subnet (the more network interfaces there are, the more IP addresses you need). Use the following tables to make sure your subnet has enough IP addresses for your task.

## Network interfaces for transfers with agents

In general, you need a DataSync agent when copying data between an AWS storage service and storage system that isn't AWS.

Location	Network interfaces created by default	Where network interfaces are created when using a public or FIPS endpoint	Where network interfaces are created when using a private (VPC) endpoint
Amazon S3	4	N/A <sup>1</sup>	The subnet you specify when activating your DataSync agent.
Amazon EFS	4	The subnet you specify when creating the Amazon EFS location.	
Amazon FSx for Windows File Server	4	The same subnet as the file system's preferred file server.	
Amazon FSx for Lustre	4	The same subnet as the file system.	
Amazon FSx for OpenZFS	4	The same subnet as the file system.	
Amazon FSx for NetApp ONTAP	4	The same subnet as the file system.	

<sup>1</sup> Network interfaces aren't needed because the DataSync service communicates directly with the S3 bucket.

## Network interfaces for transfers without agents

You don't need a DataSync agent when copying data between AWS services.

The total number of network interfaces depends on the DataSync locations in your transfer. For example, transferring between Amazon EFS and FSx for Lustre file systems requires four network interfaces. Meanwhile, transferring between FSx for Windows File Server and an S3 bucket requires two network interfaces.

Location	Network interfaces created by default	Where network interfaces are created
Amazon S3	N/A <sup>1</sup>	N/A <sup>1</sup>
Amazon EFS	2	The subnet you specify when creating the Amazon EFS location.
FSx for Windows File Server	2	The same subnet as the preferred file server for the file system.
FSx for Lustre	2	The same subnet as the file system.
FSx for OpenZFS	2	The same subnet as the file system.
FSx for ONTAP	2	The same subnet as the file system.

<sup>1</sup> Network interfaces aren't needed because the DataSync service communicates directly with the S3 bucket.

## Viewing your network interfaces

To see the network interfaces allocated to your DataSync transfer task, do one of the following:

- Use the [DescribeTask](#) operation. The operation returns `SourceNetworkInterfaceArns` and `DestinationNetworkInterfaceArns` with responses that look like this:

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

In this example, the network interface ID is `eni-f012345678abcdef0`.

- In the Amazon EC2 console, search for your task ID (such as `task-f012345678abcdef0`) to find its network interfaces.

## AWS DataSync architecture and routing examples with AWS Direct Connect

Consider the following network architectures when using [AWS Direct Connect](#) with your AWS DataSync transfers.

### Tip

If your network uses a transit gateway, we recommend separating your DataSync transfer's logical path to optimize costs (particularly if you're migrating a large amount of data).

For example, if you use [AWS Transit Gateway](#) for normal traffic between your on-premises networks and virtual private clouds (VPCs), you can configure your network so that DataSync traffic bypasses the transit gateway and its data processing charges.

## Using Direct Connect with a DataSync VPC service endpoint

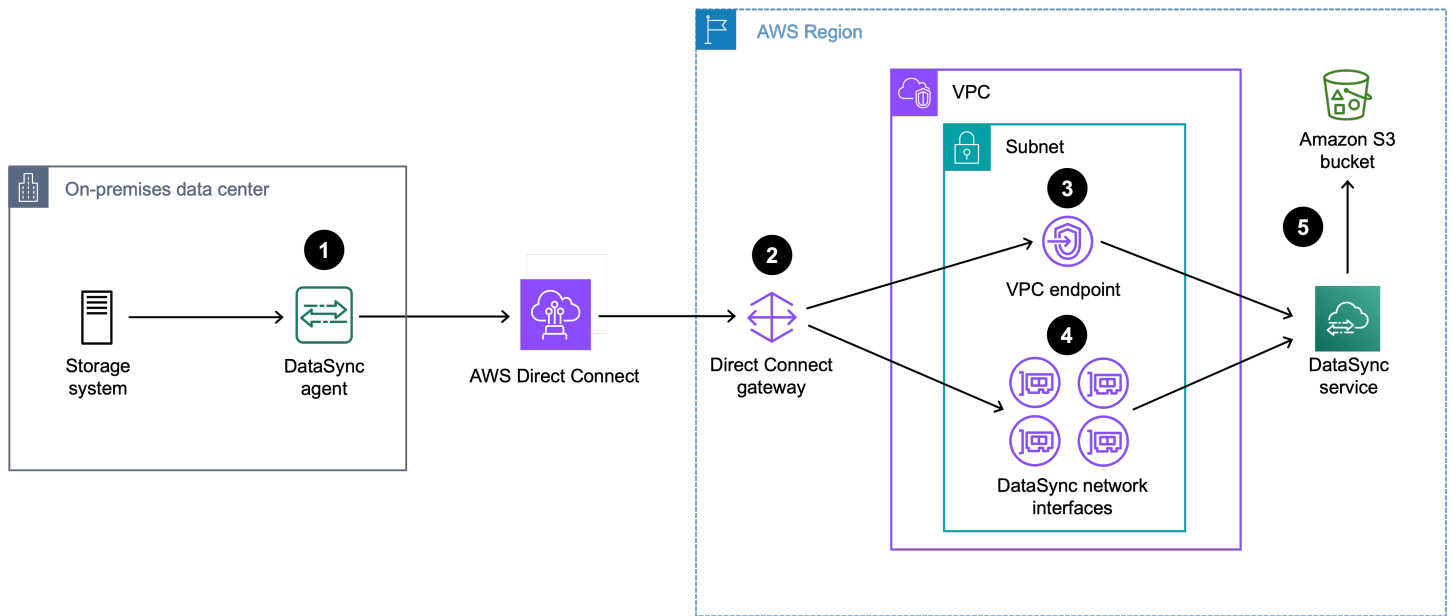
If your DataSync agent uses a [VPC service endpoint](#), you need a [Direct Connect gateway](#) to connect to your VPC.

### Contents

- [Direct Connect architecture with VPC endpoint and S3 destination](#)
- [Direct Connect architecture with VPC endpoint and file system destination in same subnet](#)
- [Direct Connect architecture with VPC endpoint and file system destination in different subnets](#)

## Direct Connect architecture with VPC endpoint and S3 destination

The following Direct Connect architecture shows a DataSync transfer from an on-premises storage system to an S3 bucket.

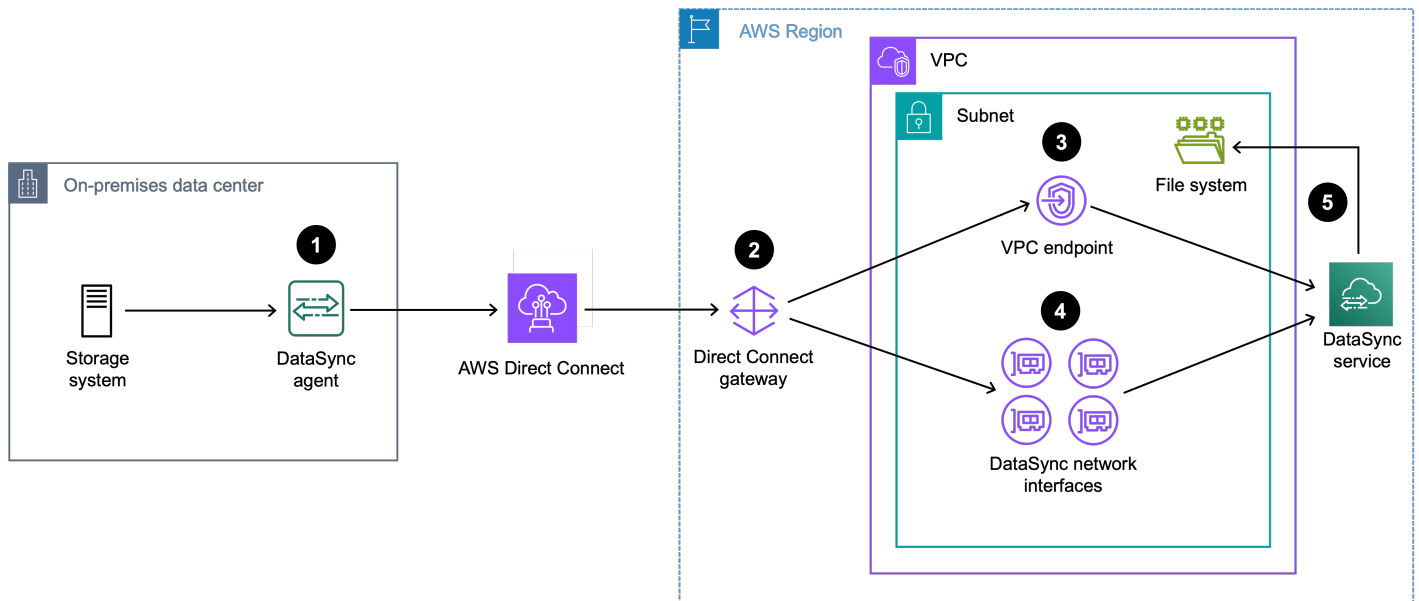


1. The DataSync agent routes DataSync traffic from the on-premises storage system (source location) to the Direct Connect connection.
2. DataSync traffic routes to a Direct Connect gateway that's used for your transfer. To set this up, you must:
  - a. Associate the Direct Connect gateway with a [virtual private gateway](#) for the VPC. This is the VPC where the DataSync VPC endpoint is located and where the DataSync task creates [network interfaces](#).
  - b. Create a [private virtual interface](#) that connects this VPC to the Direct Connect gateway.
3. DataSync traffic (control plane) routes through the DataSync VPC endpoint.
4. DataSync traffic (data plane) routes through the DataSync network interfaces in the subnet that you specify when [creating the DataSync agent](#).
5. DataSync traffic routes through the DataSync service to the S3 bucket (destination location).

## Direct Connect architecture with VPC endpoint and file system destination in same subnet

When transferring to or from an Amazon EFS or Amazon FSx file system, your file system and DataSync VPC endpoint can be in the same subnet.

The following Direct Connect architecture shows a DataSync transfer from an on-premises storage system to an Amazon EFS or Amazon FSx file system.

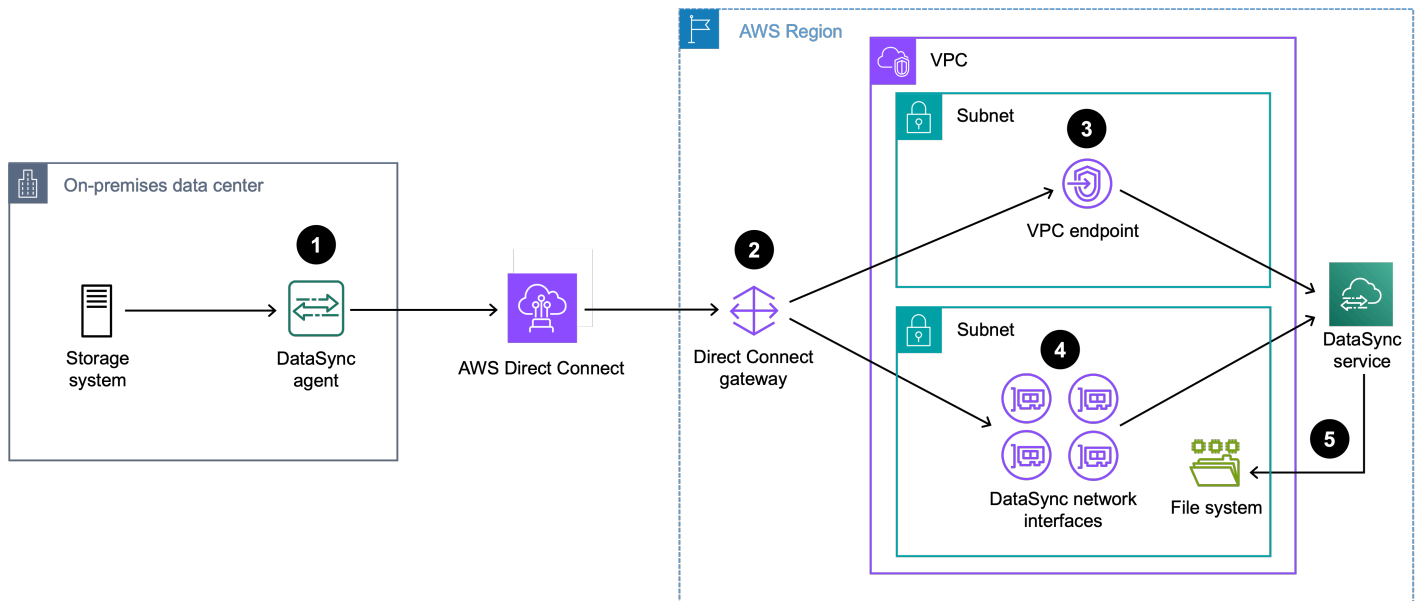


1. The DataSync agent routes DataSync traffic from the on-premises storage system (source location) to the Direct Connect connection.
2. DataSync traffic routes to a Direct Connect gateway that's used for your transfer. To set this up, you must:
  - a. Associate the Direct Connect gateway with a [virtual private gateway](#) for the VPC. This is the VPC where the DataSync VPC endpoint is located and where the DataSync task creates [network interfaces](#) for the file system (destination location).
  - b. Create a [private virtual interface](#) that connects this VPC to the Direct Connect gateway.
3. DataSync traffic (control plane) routes through the DataSync VPC endpoint.
4. DataSync traffic (data plane) routes through the DataSync network interfaces in the file system's subnet. This is the same subnet where the DataSync VPC endpoint is located.
5. DataSync traffic routes through the DataSync service to the file system (destination location).

## Direct Connect architecture with VPC endpoint and file system destination in different subnets

When transferring to or from an Amazon EFS or Amazon FSx file system, your file system and DataSync VPC endpoint can be in different subnets.

The following Direct Connect architecture shows a DataSync transfer from an on-premises storage system to an Amazon EFS or Amazon FSx file system.



1. The DataSync agent routes DataSync traffic from the on-premises storage system (source location) to the Direct Connect connection.
2. DataSync traffic routes to a Direct Connect gateway that's used for your transfer. To set this up, you must:
  - a. Associate the Direct Connect gateway with a [virtual private gateway](#) for the VPC. This is the VPC where the DataSync VPC endpoint is located and where the DataSync task creates [network interfaces](#) for the file system (destination location).
  - b. Create a [private virtual interface](#) that connects these VPCs to the Direct Connect gateway.
3. DataSync traffic (control plane) routes through the DataSync VPC endpoint.
4. DataSync traffic (data plane) routes through the DataSync network interfaces in the file system's subnet. This is a different subnet than where the DataSync VPC endpoint is located.
5. DataSync traffic routes through the DataSync service to the file system (destination location).

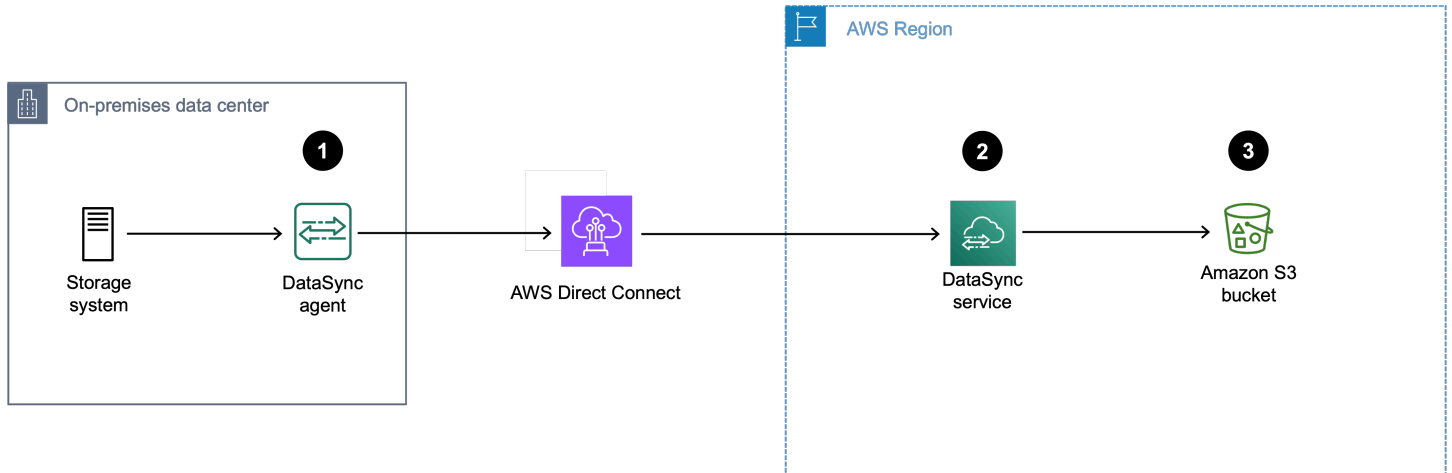
## Using Direct Connect with a DataSync public or FIPS service endpoint

If your DataSync agent uses a [public](#) or [Federal Information Processing Standard \(FIPS\)](#) service endpoint, you can route your data transfer traffic through a Direct Connect connection by using a [public virtual interface](#).

While Direct Connect advertises all local and remote AWS Region prefixes by default, you can use [BGP community tags](#) to control the scope (Regional or global) and route preference of traffic on

the public virtual interface. You must advertise at least one public prefix to create your DataSync agent.

The following Direct Connect architecture shows a DataSync transfer from an on-premises storage system through a public or FIPS endpoint to an S3 bucket.



1. The DataSync agent routes DataSync traffic from the on-premises storage system (source location) to the Direct Connect connection.
2. DataSync traffic routes to the DataSync service through a public virtual interface.
3. DataSync traffic to the S3 bucket (destination location).

## Next steps

If [you need a DataSync agent](#) and haven't created one yet, [deploy](#) the agent, [choose a service endpoint](#) for the agent, and then [activate](#) the agent.

Once you create the agent, you can [configure your network](#) for DataSync.

## Configuring your AWS DataSync agent for multiple NICs

If you configure your AWS DataSync agent to use multiple network adapters (NICs), the agent can be accessed by more than one IP address. You might want to do this in the following situations:

- **Maximizing throughput** – You might want to maximize throughput to an agent when network adapters are a bottleneck.



- **Network isolation** – Your Network File System (NFS), Server Message Block (SMB), Hadoop Distributed File System (HDFS), or object storage server might reside on a virtual LAN (VLAN) that lacks internet connectivity for security reasons.

In a typical multiple-adapter use case, one adapter is configured as the route by which the agent communicates with AWS (as the default agent). Except for this one adapter, NFS, SMB, HDFS, or self-managed object storage locations must be in the same subnet as the adapter that connects to them.

Otherwise, communication with the intended NFS, SMB, HDFS, or object storage locations might not be possible. In some cases, you might configure an NFS, SMB, HDFS, or object storage location on the same adapter that's used for communication with AWS. In these cases, NFS, SMB, HDFS, or object storage traffic for that server and AWS traffic flows through the same adapter.

In some cases, you might configure one adapter to connect to the AWS DataSync console and then add a second adapter. In such a case, DataSync automatically configures the route table to use the second adapter as the preferred route.

# Transferring your data with AWS DataSync

With AWS DataSync, you can move data from storage that's on-premises, in AWS, in other clouds, and on the edge.

## Topics

- [Where can I transfer my data with AWS DataSync?](#)
- [Transferring to or from on-premises storage with AWS DataSync](#)
- [Transferring to or from AWS storage with AWS DataSync](#)
- [Transferring to or from other cloud storage with AWS DataSync](#)
- [Transferring to or from edge storage with AWS DataSync](#)
- [Configuring what AWS DataSync transfers](#)
- [Working with AWS DataSync transfer tasks](#)

## Where can I transfer my data with AWS DataSync?

Where you can transfer your data with AWS DataSync depends on the following factors:

- Your transfer's source and destination [locations](#)
- If your locations are in different AWS accounts
- If your locations are in different AWS Regions

## DataSync transfer requirements

Depending on your transfer scenario, you might need a DataSync agent. For more information, see [Do I need an AWS DataSync agent?](#)

## Supported transfers in the same AWS account

DataSync supports transfers between the following storage resources that are associated with the same AWS account.

Source (from)	Destination (to)
<ul style="list-style-type: none"><li>• NFS</li></ul>	<ul style="list-style-type: none"><li>• Amazon S3 (in AWS Regions)</li></ul>

Source (from)	Destination (to)
<ul style="list-style-type: none"> <li>• SMB</li> <li>• HDFS</li> <li>• Object storage</li> </ul>	<ul style="list-style-type: none"> <li>• Amazon EFS</li> <li>• Amazon 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 (in AWS Regions)</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>• Object storage</li> </ul>
<ul style="list-style-type: none"> <li>• Google Cloud Storage</li> <li>• Microsoft Azure Blob Storage</li> <li>• Microsoft Azure Files</li> <li>• Wasabi Cloud Storage</li> <li>• DigitalOcean Spaces</li> <li>• Oracle Cloud Infrastructure Object Storage</li> <li>• Cloudflare R2 Storage</li> <li>• Backblaze B2 Cloud Storage</li> <li>• NAVER Cloud Object Storage</li> <li>• Alibaba Cloud Object Storage Service</li> <li>• IBM Cloud Object Storage</li> <li>• Seagate Lyve Cloud</li> </ul>	<ul style="list-style-type: none"> <li>• Amazon S3 (in AWS Regions)</li> <li>• Amazon EFS</li> <li>• Amazon FSx for Windows File Server</li> <li>• FSx for Lustre</li> <li>• FSx for OpenZFS</li> <li>• FSx for ONTAP</li> </ul>

Source (from)	Destination (to)
<ul style="list-style-type: none"> <li>• Amazon S3 (in AWS Regions)</li> <li>• Amazon EFS</li> <li>• Amazon 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>• Google Cloud Storage</li> <li>• Microsoft Azure Blob Storage</li> <li>• Microsoft Azure Files</li> <li>• Wasabi Cloud Storage</li> <li>• DigitalOcean Spaces</li> <li>• Oracle Cloud Infrastructure Object Storage</li> <li>• Cloudflare R2 Storage</li> <li>• Backblaze B2 Cloud Storage</li> <li>• NAVER Cloud Object Storage</li> <li>• Alibaba Cloud Object Storage Service</li> <li>• IBM Cloud Object Storage</li> <li>• Seagate Lyve Cloud</li> </ul>
<ul style="list-style-type: none"> <li>• Amazon S3 compatible storage on AWS Snowball Edge</li> <li>• AWS Snowcone</li> </ul>	<ul style="list-style-type: none"> <li>• Amazon S3 (in AWS Regions)</li> <li>• Amazon EFS</li> <li>• Amazon 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 (in AWS Regions)</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 compatible storage on Snowball Edge</li> <li>• AWS Snowcone</li> </ul>

Source (from)	Destination (to)
<ul style="list-style-type: none"> <li>• Amazon S3 (in AWS Regions)</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 (in AWS Regions)</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 (in AWS Regions)</li> </ul>	<ul style="list-style-type: none"> <li>• Amazon S3 on AWS Outposts</li> </ul>
<ul style="list-style-type: none"> <li>• Amazon S3 on AWS Outposts</li> </ul>	<ul style="list-style-type: none"> <li>• Amazon S3 (in AWS Regions)</li> </ul>

## Supported transfers across AWS accounts

DataSync supports some transfers between storage resources that are associated with different AWS accounts.

Source (from)	Destination (to)
<ul style="list-style-type: none"> <li>• Amazon EFS<sup>1</sup></li> <li>• FSx for Windows File Server<sup>2</sup></li> </ul>	<ul style="list-style-type: none"> <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 (in AWS Regions)</li> </ul>	<ul style="list-style-type: none"> <li>• Amazon S3 (in AWS Regions)</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>

Source (from)	Destination (to)
<ul style="list-style-type: none"> <li>• Amazon S3 (in AWS Regions)</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 (in AWS Regions)</li> </ul>
<ul style="list-style-type: none"> <li>• NFS</li> <li>• SMB</li> <li>• HDFS</li> <li>• Object storage</li> </ul>	<ul style="list-style-type: none"> <li>• Amazon S3 (in AWS Regions)</li> </ul>

<sup>1</sup> Configured as an [NFS location](#).

<sup>2</sup> Configured as an [SMB location](#).

## Supported transfers in the same AWS Region

There are no restrictions when transferring data within the same AWS Region (including [opt-in Regions](#)). For more information, see [AWS Regions supported by DataSync](#).

## Supported transfers between AWS Regions

Note the following when transferring data between [AWS Regions supported by DataSync](#):

- When you transfer data between AWS storage services in different AWS Regions, one of the two locations must be in the Region where you're using DataSync.
- With AWS GovCloud (US) Regions, you can:
  - Transfer between the AWS GovCloud (US-East) and AWS GovCloud (US-West) Regions.
  - Transfer between an AWS GovCloud (US) Region and commercial AWS Region, such as US East (N. Virginia).

- You can't transfer to or from an [opt-in Region](#) (even if you enable the opt-in Region) unless you use a DataSync agent.

### Important

You pay for data transferred between AWS Regions. This transfer is billed as data transfer out from the source to destination Region. For more information, see [AWS DataSync Pricing](#).

## Transferring to or from on-premises storage with AWS DataSync

With AWS DataSync, you can transfer files and objects to or from a number of on-premises and self-managed storage systems.

### Topics

- [Configuring AWS DataSync transfers with an NFS file server](#)
- [Configuring AWS DataSync transfers with an SMB file server](#)
- [Configuring AWS DataSync transfers with HDFS](#)
- [Configuring AWS DataSync transfers with an object storage system](#)

## Configuring AWS DataSync transfers with an NFS file server

With AWS DataSync, you can transfer data between your Network File System (NFS) file server and one of the following AWS storage services:

- [Amazon S3](#)
- [Amazon EFS](#)
- [Amazon FSx for Windows File Server](#)
- [Amazon FSx for Lustre](#)
- [Amazon FSx for OpenZFS](#)
- [Amazon FSx for NetApp ONTAP](#)

To set up this kind of transfer, you create a [location](#) for your NFS file server. You can use this location as a transfer source or destination.

## Providing DataSync access to NFS file servers

For DataSync to access your NFS file server, you need a DataSync [agent](#). The agent mounts an export on your file server by using the NFS protocol.

### Topics

- [Configuring your NFS export](#)
- [Supported NFS versions](#)

### Configuring your NFS export

The export that DataSync needs for your transfer depends on if your NFS file server is a source or destination location and how your file server's permissions are configured.

If your file server is a source location, DataSync just has to read and traverse your files and folders. If it's a destination location, DataSync needs root access to write to the location and set ownership, permissions, and other metadata on the files and folders that you're copying. You can use the `no_root_squash` option to allow root access for your export.

The following examples describe how to configure an NFS export that provides access to DataSync.

#### When your NFS file server is a source location (root access)

Configure your export by using the following command, which provides DataSync read-only permissions (`ro`) and root access (`no_root_squash`):

```
export-path datasync-agent-ip-address(ro,no_root_squash)
```

#### When your NFS file server is a destination location

Configure your export by using the following command, which provides DataSync write permissions (`rw`) and root access (`no_root_squash`):

```
export-path datasync-agent-ip-address(rw,no_root_squash)
```

#### When your NFS file server is a source location (no root access)



Configure your export by using the following command, which specifies the POSIX user ID (UID) and group ID (GID) that you know would provide DataSync read-only permissions on the export:

```
export-path datasync-agent-ip-address(ro,all_squash,anonuid=uid,anongid=gid)
```

## Supported NFS versions

By default, DataSync uses NFS version 4.1. DataSync also supports NFS 3.x and 4.0.

## Configuring your network for NFS transfers

For your DataSync transfer, you must configure traffic for a few network connections:

1. Allow traffic on TCP port 2049 from your DataSync agent to your NFS file server.  
  
Other NFS clients in your network should be able to mount the NFS export that you're using to transfer data. The export must also be accessible without Kerberos authentication.
2. Configure traffic for your [service endpoint connection](#) (such as a VPC, public, or FIPS endpoint).
3. Allow traffic from the DataSync service to the [AWS storage service](#) you're transferring to or from.

## Creating your NFS transfer location

Before you begin, note the following:

- You need an NFS file server that you want to transfer data from.
- You need a DataSync agent that can [access your file server](#).
- DataSync doesn't support copying NFS version 4 access control lists (ACLs).

## Using the DataSync console

1. Open the AWS DataSync console at <https://console.aws.amazon.com/datasync/>.
2. In the left navigation pane, expand **Data transfer**, then choose **Locations** and **Create location**.
3. For **Location type**, choose **Network File System (NFS)**.
4. For **Agents**, choose the DataSync agent that you want to connect to your NFS file server.

You can choose more than one agent. For more information, see [Using multiple DataSync agents](#).

5. For **NFS server**, enter the Domain Name System (DNS) name or IP address of the NFS file server that your DataSync agent connects to.
6. For **Mount path**, enter the NFS export path that you want DataSync to mount.

This path (or a subdirectory of the path) is where DataSync transfers data to or from. For more information, see [Configuring your NFS export](#).

7. (Optional) Expand **Additional settings** and choose a specific **NFS version** for DataSync to use when accessing your file server.

For more information, see [Supported NFS versions](#).

8. (Optional) Choose **Add tag** to tag your NFS location.

*Tags* are key-value pairs that help you manage, filter, and search for your locations. We recommend creating at least a name tag for your location.

9. Choose **Create location**.

## Using the AWS CLI

- Use the following command to create an NFS location.

```
$ aws datasync create-location-nfs \  
  --server-hostname nfs-server-address \  
  --on-prem-config AgentArns=datasync-agent-arns \  
  --subdirectory nfs-export-path
```

For more information on creating the location, see [Providing DataSync access to NFS file servers](#).

DataSync automatically chooses the NFS version that it uses to read from an NFS location. To specify an NFS version, use the optional `Version` parameter in the [NfsMountOptions](#) API operation.

This command returns the Amazon Resource Name (ARN) of the NFS location, similar to the ARN shown following.

```
{  
  "LocationArn": "arn:aws:datasync:us-east-1:111222333444:location/  
loc-0f01451b140b2af49"
```

```
}
```

To make sure that the directory can be mounted, you can connect to any computer that has the same network configuration as your agent and run the following command.

```
mount -t nfs -o nfsvers=<nfs-server-version <nfs-server-address:<nfs-export-path <test-  
folder
```

The following is an example of the command.

```
mount -t nfs -o nfsvers=3 198.51.100.123:/path_for_sync_to_read_from /  
temp_folder_to_test_mount_on_local_machine
```

## Configuring AWS DataSync transfers with an SMB file server

With AWS DataSync, you can transfer data between your Server Message Block (SMB) file server and one of the following AWS storage services:

- [Amazon S3](#)
- [Amazon EFS](#)
- [Amazon FSx for Windows File Server](#)
- [Amazon FSx for Lustre](#)
- [Amazon FSx for OpenZFS](#)
- [Amazon FSx for NetApp ONTAP](#)

To set up this kind of transfer, you create a [location](#) for your SMB file server. You can use this as a transfer source or destination.

### Providing DataSync access to SMB file servers

DataSync connects to your file server using the SMB protocol and authenticates with credentials that you provide it.

#### Topics

- [Supported SMB versions](#)
- [Required permissions](#)

- [Required authentication protocols](#)
- [DFS Namespaces](#)

## Supported SMB versions

By default, DataSync automatically chooses a version of the SMB protocol based on negotiation with your SMB file server.

You also can configure DataSync to use a specific SMB version, but we recommend doing this only if DataSync has trouble negotiating with the SMB file server automatically. (DataSync supports SMB versions 1.0 and later.)

See the following table for a list of options in the DataSync console and API:

Console option	API option	Description
Automatic	AUTOMATIC	DataSync and the SMB file server negotiate the highest version of SMB that they mutually support between 2.1 and 3.1.1.  This is the default and recommended option. If you instead choose a specific version that your file server doesn't support, you may get an <code>Operation Not Supported</code> error.
SMB 3.0.2	SMB3	Restricts the protocol negotiation to only SMB version 3.0.2.
SMB 2.1	SMB2	Restricts the protocol negotiation to only SMB version 2.1.
SMB 2.0	SMB2_0	Restricts the protocol negotiation to only SMB version 2.0.
SMB 1.0	SMB1	Restricts the protocol negotiation to only SMB version 1.0.

## Required permissions

You must provide DataSync a user with the necessary rights to mount and access your SMB file server's files, folders, and file metadata. This can be a local user on your file server or a domain user in your Microsoft Active Directory.

If you provide a user in your Active Directory, the user must be a member of an Active Directory group with one or both of the following user rights (depending the [metadata that you want DataSync to copy](#)):

User right	Description
<b>Restore files and directories</b> (SE_RESTORE_NAME )	<p>Allows DataSync to copy object ownership, permissions, file metadata, and NTFS discretionary access lists (DACLS).</p> <p>This user right is usually granted to members of the <b>Domain Admins</b> and <b>Backup Operators</b> groups (both of which are default Active Directory groups).</p>
<b>Manage auditing and security log</b> (SE_SECURITY_NAME )	<p>Allows DataSync to copy NTFS system access control lists (SACLs).</p> <p>This user right is usually granted to members of the <b>Domain Admins</b> group.</p>

If you want to copy Windows ACLs and are transferring between an SMB file server and another storage system that uses SMB (such as Amazon FSx for Windows File Server or FSx for ONTAP), the users that you provide DataSync must belong to the same Active Directory domain or have an Active Directory trust relationship between their domains.

## Required authentication protocols

Your SMB file server must use NTLM authentication for DataSync to access it. DataSync can't access an SMB file server that uses Kerberos authentication.

## DFS Namespaces

DataSync doesn't support Microsoft Distributed File System (DFS) Namespaces. We recommend specifying an underlying file server or share instead when creating your DataSync location.

## Creating your SMB transfer location

Before you begin, you need an SMB file server that you want to transfer data from.

### To create an SMB location by using the console

1. Open the AWS DataSync console at <https://console.aws.amazon.com/datasync/>.
2. In the left navigation pane, expand **Data transfer**, then choose **Locations** and **Create location**.
3. For **Location type**, choose **Server Message Block (SMB)**.

You configure this location as a source or destination later.

4. For **Agents**, choose one or more DataSync agents that you want to connect to your SMB file server.

If you choose more than one agent, make sure you understand using [multiple agents for a location](#).

5. For **SMB Server**, enter the Domain Name System (DNS) name or IP address of the SMB file server that your DataSync agent will mount.

#### Note

You can't specify an IP version 6 (IPv6) address.

6. For **Share name**, enter the name of the share exported by your SMB file server where DataSync will read or write data.

You can include a subdirectory in the share path (for example, `/path/to/subdirectory`). Make sure that other SMB clients in your network can also mount this path.

To copy all the data in the subdirectory, DataSync must be able to mount the SMB share and access all of its data. For more information, see [Required permissions](#).

7. (Optional) Expand **Additional settings** and choose an **SMB Version** for DataSync to use when accessing your file server.

By default, DataSync automatically chooses a version based on negotiation with the SMB file server. For information, see [Supported SMB versions](#).

8. For **User**, enter a user name that can mount your SMB file server and has permission to access the files and folders involved in your transfer.

For more information, see [Required permissions](#).

9. For **Password**, enter the password of the user who can mount your SMB file server and has permission to access the files and folders involved in your transfer.
10. (Optional) For **Domain**, enter the Windows domain name that your SMB file server belongs to.

If you have multiple domains in your environment, configuring this setting makes sure that DataSync connects to the right SMB file server.

11. (Optional) Choose **Add tag** to tag your SMB location.

*Tags* are key-value pairs that help you manage, filter, and search for your locations. We recommend creating at least a name tag for your location.

12. Choose **Create location**.

## Configuring AWS DataSync transfers with HDFS

With AWS DataSync, you can transfer data between your Hadoop Distributed File System (HDFS) cluster and one of the following AWS storage services:

- [Amazon S3](#)
- [Amazon EFS](#)
- [Amazon FSx for Windows File Server](#)
- [Amazon FSx for Lustre](#)
- [Amazon FSx for OpenZFS](#)
- [Amazon FSx for NetApp ONTAP](#)

To set up this kind of transfer, you create a [location](#) for your HDFS cluster. You can use this location as a transfer source or destination.

## Providing DataSync access to HDFS clusters

To connect to your HDFS cluster, DataSync uses an agent that you deploy near your HDFS cluster. For more information, see [Deploying your AWS DataSync agent](#). The DataSync agent acts as an HDFS client and communicates with the NameNodes and DataNodes in your clusters.

When you start a task, DataSync queries the NameNode for locations of files and folders on the cluster. If the HDFS location is configured as a source, then DataSync reads files and folder data from the DataNodes in the cluster and copies the data to the destination. If the HDFS location is configured as a destination, then DataSync writes files and folders from the destination to the DataNodes in the cluster. Before running your DataSync task, verify agent connectivity to the HDFS cluster. For more information, see [Testing your agent's connection to your storage](#).

### Authentication

When connecting to an HDFS cluster, DataSync supports simple authentication or Kerberos authentication. To use simple authentication, provide the user name of a user with rights to read and write to the HDFS cluster. To use Kerberos authentication, provide a Kerberos configuration file, a Kerberos key table (keytab) file, and a Kerberos principal name. The credentials of the Kerberos principal must be in the provided keytab file.

### Encryption

When using Kerberos authentication, DataSync supports encryption of data as it's transmitted between the DataSync agent and your HDFS cluster. Encrypt your data by using the Quality of Protection (QOP) configuration settings on your HDFS cluster and by specifying the QOP settings when creating your HDFS location. The QOP configuration includes settings for data transfer protection and Remote Procedure Call (RPC) protection.

### DataSync supports the following Kerberos encryption types:

- `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

You can also configure HDFS clusters for encryption at rest using Transparent Data Encryption (TDE). When using simple authentication, DataSync reads and writes to TDE-enabled clusters. If you're using DataSync to copy data to a TDE-enabled cluster, first configure the encryption zones on the HDFS cluster. DataSync doesn't create encryption zones.

## Creating your HDFS transfer location

Configure a location that you can use a source for your DataSync transfer.

**Before you begin:** Verify network connectivity between your agent and Hadoop cluster by doing the following:

- Test access to the TCP ports listed in [Network requirements for on-premises, self-managed, other cloud, and edge storage](#).
- Test access between your local agent and your Hadoop cluster. For instructions, see [Testing your agent's connection to your storage](#).

### To create an HDFS location by using the DataSync console

1. Open the AWS DataSync console at <https://console.aws.amazon.com/datasync/>.
2. In the left navigation pane, expand **Data transfer**, then choose **Locations** and **Create location**.
3. For **Location type**, choose **Hadoop Distributed File System (HDFS)**. You can configure this location as a source or destination later.
4. For **Agents**, choose one or more agents that you want to use from the list of available agents. The agent connects to your HDFS cluster to securely transfer data between the HDFS cluster and DataSync.
5. For **NameNode**, provide the domain name or IP address of the HDFS cluster's primary NameNode.

6. For **Folder**, enter a folder on your HDFS cluster that DataSync will use for the data transfer. When the location is used as a source for a task, DataSync copies files in the provided folder. When your location is used as a destination for a task, DataSync writes all files to the provided folder.
7. To set the **Block size** or **Replication factor**, choose **Additional settings**. The default block size is 128 MiB, and any provided block sizes must be a multiple of 512 bytes. The default replication factor is three DataNodes when transferring data to the HDFS cluster.
8. In the **Security** section, choose the **Authentication type** used on your HDFS cluster.
  - **Simple** – For **User**, specify the user name with the following permissions on the HDFS cluster (depending on your use case):
    - If you plan to use this location as a source location, specify a user that only has read permissions.
    - If you plan to use this location as a destination location, specify a user that has read and write permissions.

Optionally, specify the URI of the Key Management Server (KMS) of the HDFS cluster.

- **Kerberos** – Specify the Kerberos **Principal** with access to your HDFS cluster. Next, provide the **KeyTab file** that contains the provided Kerberos principal. Then, provide the **Kerberos configuration file**. Finally, specify the type of encryption in transit protection in the **RPC protection** and **Data transfer protection** dropdown lists.
9. (Optional) Choose **Add tag** to tag your HDFS location.

*Tags* are key-value pairs that help you manage, filter, and search for your locations. We recommend creating at least a name tag for your location.
  10. Choose **Create location**.

## Unsupported HDFS features

The following capabilities of HDFS aren't currently supported by DataSync:

- Transparent Data Encryption (TDE) when using Kerberos authentication
- Configuring multiple NameNodes
- Hadoop HDFS over HTTP (HttpFS)
- POSIX access control lists (ACLs)
- HDFS extended attributes (xattrs)

## Configuring AWS DataSync transfers with an object storage system

With AWS DataSync, you can transfer data between your object storage system and one of the following AWS storage services:

- [Amazon S3](#)
- [Amazon EFS](#)
- [Amazon FSx for Windows File Server](#)
- [Amazon FSx for Lustre](#)
- [Amazon FSx for OpenZFS](#)
- [Amazon FSx for NetApp ONTAP](#)

To set up this kind of transfer, you create a [location](#) for your object storage system. You can use this location as a transfer source or destination.

### Prerequisites

Your object storage system must be compatible with the following [Amazon S3 API operations](#) for DataSync to connect to it:

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

- PutObjectTagging
- UploadPart

## Creating your object storage transfer location

Before you begin, you need an object storage system that you plan to transfer data from.

### To create an object storage location by using the console

1. Open the AWS DataSync console at <https://console.aws.amazon.com/datasync/>.
2. In the left navigation pane, expand **Data transfer**, then choose **Locations** and **Create location**.
3. For **Location type**, choose **Object storage**.

You configure this location as a source or destination later.

4. For **Agents**, choose one or more DataSync agents.

During the transfer, the agents securely connect to your object storage server.

5. For **Server**, provide the domain name or IP address of the object storage server.
6. For **Bucket name**, enter the name of the object storage bucket involved in the transfer.
7. For **Folder**, enter an object prefix.

DataSync only copies objects with this prefix.

8. To configure the connection to the object storage server, expand **Additional settings** and do the following:
  - a. For **Server protocol**, choose **HTTP** or **HTTPS**.
  - b. For **Server port**, use a default port (**80** for HTTP or **443** for HTTPS) or specify a custom port if needed.
  - c. For **Certificate**, if your object storage system uses a private or self-signed certificate authority (CA), select **Choose file** and specify a single .pem file with a full certificate chain.

The certificate chain might include:

- The object storage system's certificate
- All intermediate certificates (if there are any)
- The root certificate of the signing CA

You can concatenate your certificates into a `.pem` file (which can be up to 32768 bytes before base64 encoding). The following example `cat` command creates an `object_storage_certificates.pem` file that includes three certificates:

```
cat object_server_certificate.pem intermediate_certificate.pem ca_root_certificate.pem  
> object_storage_certificates.pem
```

9. If credentials are required to access the object storage server, select **Requires credentials** and enter the **Access key** and **Secret key** for accessing the bucket.

The access key and secret key can be a user name and password, respectively.

10. (Optional) Choose **Add tag** to tag your object storage location.

*Tags* are key-value pairs that help you manage, filter, and search for your locations. We recommend creating at least a name tag for your location.

11. Choose **Create location**.

## Transferring to or from AWS storage with AWS DataSync

With AWS DataSync, you can transfer data to or from a number of AWS storage services. For more information, see [Where can I transfer my data with DataSync?](#).

### Topics

- [Configuring AWS DataSync transfers with Amazon S3](#)
- [Configuring AWS DataSync transfers with Amazon EFS](#)
- [Configuring AWS DataSync transfers with Amazon FSx for Windows File Server](#)
- [Configuring AWS DataSync transfers with Amazon FSx for Lustre](#)
- [Configuring AWS DataSync transfers with Amazon FSx for OpenZFS](#)
- [Configuring AWS DataSync transfers with Amazon FSx for NetApp ONTAP](#)

## Configuring AWS DataSync transfers with Amazon S3

To transfer data to or from your Amazon S3 bucket, you create an AWS DataSync transfer *location*. DataSync can use this location as a source or destination for transferring data.

### **⚠ Important**

Before you create your location, make sure that you read the following sections:

- [Storage class considerations with Amazon S3 transfers](#)
- [Evaluating S3 request costs when using DataSync](#)

## Providing DataSync access to S3 buckets

DataSync needs access to the S3 bucket that you're transferring to or from. To do this, you must create an AWS Identity and Access Management (IAM) role that DataSync assumes with the permissions required to access the bucket. You then specify this role when [creating your Amazon S3 location for DataSync](#).

### Contents

- [Required permissions](#)
- [Creating an IAM role for DataSync to access your Amazon S3 location](#)
- [Accessing S3 buckets using server-side encryption](#)
- [Accessing S3 buckets with restricted VPC access](#)

### Required permissions

The permissions that your IAM role needs can depend on whether bucket is a DataSync source or destination location. Amazon S3 on Outposts requires a different set of permissions.

#### Amazon S3 (source location)

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Action": [
        "s3:GetBucketLocation",
        "s3:ListBucket",
        "s3:ListBucketMultipartUploads"
      ],
      "Effect": "Allow",
      "Resource": "arn:aws:s3:::bucket-name"
    }
  ]
}
```

```

    },
    {
      "Action": [
        "s3:AbortMultipartUpload",
        "s3:GetObject",
        "s3:GetObjectTagging",
        "s3:GetObjectVersion",
        "s3:GetObjectVersionTagging",
        "s3:ListMultipartUploadParts"
      ],
      "Effect": "Allow",
      "Resource": "arn:aws:s3:::bucket-name/*"
    }
  ]
}

```

### Amazon S3 (destination location)

```

{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Action": [
        "s3:GetBucketLocation",
        "s3:ListBucket",
        "s3:ListBucketMultipartUploads"
      ],
      "Effect": "Allow",
      "Resource": "arn:aws:s3:::bucket-name"
    },
    {
      "Action": [
        "s3:AbortMultipartUpload",
        "s3:DeleteObject",
        "s3:GetObject",
        "s3:GetObjectTagging",
        "s3:GetObjectVersion",
        "s3:GetObjectVersionTagging",
        "s3:ListMultipartUploadParts",
        "s3:PutObject",
        "s3:PutObjectTagging"
      ],
      "Effect": "Allow",

```

```

    "Resource": "arn:aws:s3::bucket-name/*"
  }
]
}

```

## Amazon S3 on Outposts

```

{
  "Version": "2012-10-17",
  "Statement": [{
    "Action": [
      "s3-outposts:ListBucket",
      "s3-outposts:ListBucketMultipartUploads"
    ],
    "Effect": "Allow",
    "Resource": [
      "arn:aws:s3-outposts:region:account-id:outpost/outpost-id/bucket-name",
      "arn:aws:s3-outposts:region:account-id:outpost/outpost-id/accesspoint/bucket-access-point-name"
    ]
  },
  {
    "Action": [
      "s3-outposts:AbortMultipartUpload",
      "s3-outposts>DeleteObject",
      "s3-outposts:GetObject",
      "s3-outposts:GetObjectTagging",
      "s3-outposts:GetObjectVersion",
      "s3-outposts:GetObjectVersionTagging",
      "s3-outposts:ListMultipartUploadParts",
      "s3-outposts:PutObject",
      "s3-outposts:PutObjectTagging"
    ],
    "Effect": "Allow",
    "Resource": [
      "arn:aws:s3-outposts:region:account-id:outpost/outpost-id/bucket-name/*",
      "arn:aws:s3-outposts:region:account-id:outpost/outpost-id/accesspoint/bucket-access-point-name/*"
    ]
  },
  {

```



```
        "Action": "s3-outposts:GetAccessPoint",
        "Effect": "Allow",
        "Resource": "arn:aws:s3-outposts:region:account-
id:outpost/outpost-id/accesspoint/bucket-access-point-name"
    }
  ]
}
```

## Creating an IAM role for DataSync to access your Amazon S3 location

When [creating your Amazon S3 location](#) in the console, DataSync can automatically create and assume an IAM role that normally has the right permissions to access your S3 bucket.

In some situations, you might need to create this role manually (for example, accessing buckets with extra layers of security or transferring to or from a bucket in a different AWS accounts).

### Manually creating an IAM role for DataSync

1. Open the IAM console at <https://console.aws.amazon.com/iam/>.
2. In the left navigation pane, under **Access management**, choose **Roles**, and then choose **Create role**.
3. On the **Select trusted entity** page, for **Trusted entity type**, choose **AWS service**.
4. For **Use case**, choose **DataSync** in the dropdown list and select **DataSync**. Choose **Next**.
5. On the **Add permissions** page, choose **Next**. Give your role a name and choose **Create role**.
6. On the **Roles** page, search for the role that you just created and choose its name.
7. On the role's details page, choose the **Permissions** tab. Choose **Add permissions** then **Create inline policy**.
8. Choose the **JSON** tab and [add the permissions required](#) to access your bucket into the policy editor.
9. Choose **Next**. Give your policy a name and choose **Create policy**.
10. (Recommended) To prevent the [cross-service confused deputy problem](#), do the following:
  - a. On the role's details page, choose the **Trust relationships** tab. Choose **Edit trust policy**.
  - b. Update the trust policy by using the following example, which includes the `aws:SourceArn` and `aws:SourceAccount` global condition context keys:

```
{
```

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

- c. Choose **Update policy**.

You can specify this role when creating your Amazon S3 location.

## Accessing S3 buckets using server-side encryption

DataSync can transfer data to or from [S3 buckets that use server-side encryption](#). The type of encryption key a bucket uses can determine if you need a custom policy allowing DataSync to access the bucket.

When using DataSync with S3 buckets that use server-side encryption, remember the following:

- **If your S3 bucket is encrypted with an AWS managed key** – DataSync can access the bucket's objects by default if all your resources are in the same AWS account.
- **If your S3 bucket is encrypted with a customer-managed AWS Key Management Service (AWS KMS) key (SSE-KMS)** – The [key's policy](#) must include the IAM role that DataSync uses to access the bucket.
- **If your S3 bucket is encrypted with a customer-managed SSE-KMS key and in a different AWS account** – DataSync needs permission to access the bucket in the other AWS account. You can set up this up by doing the following:
  - In the IAM role that DataSync uses, you must specify the cross-account bucket's SSE-KMS key by using the key's fully qualified Amazon Resource Name (ARN). This is the same key ARN that

you use to configure the bucket's [default encryption](#). You can't specify a key ID, alias name, or alias ARN in this situation.

Here's an example key ARN:

```
arn:aws:kms:us-  
west-2:111122223333:key/1234abcd-12ab-34cd-56ef-1234567890ab
```

For more information on specifying KMS keys in IAM policy statements, see the [AWS Key Management Service Developer Guide](#).

- In the SSE-KMS key policy, [specify the IAM role used by DataSync](#).
- **If your S3 bucket is encrypted with a customer-provided encryption key (SSE-C)** – DataSync can't access this bucket.

### Example: SSE-KMS key policy for DataSync

The following example is a [key policy](#) for a customer-managed SSE-KMS key. The policy is associated with an S3 bucket that uses server-side encryption.

If you want to use this example, replace the following values with your own:

- *account-id* – Your AWS account.
- *admin-role-name* – The name of the IAM role that can administer the key.
- *datasync-role-name* – The name of the IAM role that allows DataSync to use the key when accessing the bucket.

```
{  
  "Id": "key-consolepolicy-3",  
  "Version": "2012-10-17",  
  "Statement": [  
    {  
      "Sid": "Enable IAM Permissions",  
      "Effect": "Allow",  
      "Principal": {  
        "AWS": "arn:aws:iam::account-id:root"  
      },  
      "Action": "kms:*",  
      "Resource": "*"   
    },  
  ],  
}
```

```

{
  "Sid": "Allow access for Key Administrators",
  "Effect": "Allow",
  "Principal": {
    "AWS": "arn:aws:iam::account-id:role/admin-role-name"
  },
  "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::account-id:role/datasync-role-name"
  },
  "Action": [
    "kms:Encrypt",
    "kms:Decrypt",
    "kms:ReEncrypt*",
    "kms:GenerateDataKey*"
  ],
  "Resource": "*"
},
{
  "Sid": "Allow attachment of persistent resources",
  "Effect": "Allow",
  "Principal": {
    "AWS": "arn:aws:iam::account-id:role/datasync-role-name"
  },

```

```

    "Action": [
      "kms:CreateGrant",
      "kms:ListGrants",
      "kms:RevokeGrant"
    ],
    "Resource": "*",
    "Condition": {
      "Bool": {
        "kms:GrantIsForAWSResource": "true"
      }
    }
  }
}

```

## Accessing S3 buckets with restricted VPC access

An Amazon S3 bucket that [limits access to specific virtual private cloud \(VPC\) endpoints or VPCs](#) will deny DataSync from transferring to or from that bucket. To enable transfers in these situations, you can update the bucket's policy to include the IAM role that you [specify with your DataSync location](#).

### Option 1: Allowing access based on DataSync location role ARN

In the S3 bucket policy, you can specify the Amazon Resource Name (ARN) of your DataSync location IAM role.

The following example is an S3 bucket policy that denies access from all but two VPCs (`vpc-1234567890abcdef0` and `vpc-abcdef01234567890`). However, the policy also includes the [ArnNotLikeExists](#) condition and [aws:PrincipalArn](#) condition key, which allow the ARN of a DataSync location role to access the bucket.

```

{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "Access-to-specific-VPCs-only",
      "Effect": "Deny",
      "Principal": "*",
      "Action": "s3:*",
      "Resource": "arn:aws:s3:::bucket-name/*",
      "Condition": {

```

```

        "StringNotEqualsIfExists": {
            "aws:SourceVpc": [
                "vpc-1234567890abcdef0",
                "vpc-abcdef01234567890"
            ]
        },
        "ArnNotLikeIfExists": {
            "aws:PrincipalArn": [
                "arn:aws:iam::account-id:role/datasync-location-role-name"
            ]
        }
    }
}
]
}

```

## Option 2: Allowing access based on DataSync location role tag

In the S3 bucket policy, you can specify a tag attached to your DataSync location IAM role.

The following example is an S3 bucket policy that denies access from all but two VPCs (vpc-1234567890abcdef0 and vpc-abcdef01234567890). However, the policy also includes the [StringNotEqualsIfExists](#) condition and [aws:PrincipalTag](#) condition key, which allow a principal with the tag key `exclude-from-vpc-restriction` and value `true`. You can try a similar approach in your bucket policy by specifying a tag attached to your DataSync location role.

```

{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "Access-to-specific-VPCs-only",
      "Effect": "Deny",
      "Principal": "*",
      "Action": "s3:*",
      "Resource": "arn:aws:s3:::bucket-name/*",
      "Condition": {
        "StringNotEqualsIfExists": {
          "aws:SourceVpc": [
            "vpc-1234567890abcdef0",
            "vpc-abcdef01234567890"
          ],
          "aws:PrincipalTag/exclude-from-vpc-restriction": "true"
        }
      }
    }
  ]
}

```

```

    }
  }
]
}

```

## Storage class considerations with Amazon S3 transfers

When Amazon S3 is your transfer destination, DataSync can transfer data directly into a specific [Amazon S3 storage class](#). Some storage classes have behaviors that can affect your Amazon S3 storage costs. For more information, see [Amazon S3 pricing](#).

### Important

New objects copied to an S3 bucket are stored using the storage class that you specify when creating your Amazon S3 transfer location. DataSync won't change the storage class of existing objects in the bucket (even if that object was modified in the source location).

Amazon S3 storage class	Considerations
S3 Standard	Choose S3 Standard to store your frequently accessed files redundantly in multiple Availability Zones that are geographically separated. This is the default if you don't specify a storage class.
S3 Intelligent-Tiering	Choose S3 Intelligent-Tiering to optimize storage costs by automatically moving data to the most cost-effective storage access tier.  You pay a monthly charge per object stored in the S3 Intelligent-Tiering storage class. This Amazon S3 charge includes monitoring data access patterns and moving objects between tiers.
S3 Standard-IA	Choose S3 Standard-IA to store your infrequently accessed objects redundantly in multiple Availability Zones that are geographically separated.

<b>Amazon S3 storage class</b>	<b>Considerations</b>
	<p>Objects stored in the S3 Standard-IA storage class can incur additional charges for overwriting, deleting, or retrieving. Consider how often these objects change, how long you plan to keep these objects, and how often you need to access them. Changes to object data or metadata are equivalent to deleting an object and creating a new one to replace it. This results in additional charges for objects stored in the S3 Standard-IA storage class.</p> <p>Objects less than 128 KB are smaller than the minimum capacity charge per object in the S3 Standard-IA storage class. These objects are stored in the S3 Standard storage class.</p>
<b>S3 One Zone-IA</b>	<p>Choose S3 One Zone-IA to store your infrequently accessed objects in a single Availability Zone.</p> <p>Objects stored in the S3 One Zone-IA storage class can incur additional charges for overwriting, deleting, or retrieving. Consider how often these objects change, how long you plan to keep these objects, and how often you need to access them. Changes to object data or metadata are equivalent to deleting an object and creating a new one to replace it. This results in additional charges for objects stored in the S3 One Zone-IA storage class.</p> <p>Objects less than 128 KB are smaller than the minimum capacity charge per object in the S3 One Zone-IA storage class. These objects are stored in the S3 Standard storage class.</p>



<b>Amazon S3 storage class</b>	<b>Considerations</b>
S3 Glacier Instant Retrieval	<p>Choose S3 Glacier Instant Retrieval to archive objects that are rarely accessed but require retrieval in milliseconds.</p> <p>Data stored in the S3 Glacier Instant Retrieval storage class offers cost savings compared to the S3 Standard-IA storage class with the same latency and throughput performance. S3 Glacier Instant Retrieval has higher data access costs than S3 Standard-IA, though.</p> <p>Objects stored in S3 Glacier Instant Retrieval can incur additional charges for overwriting, deleting, or retrieving. Consider how often these objects change, how long you plan to keep these objects, and how often you need to access them. Changes to object data or metadata are equivalent to deleting an object and creating a new one to replace it. This results in additional charges for objects stored in the S3 Glacier Instant Retrieval storage class.</p> <p>Objects less than 128 KB are smaller than the minimum capacity charge per object in the S3 Glacier Instant Retrieval storage class. These objects are stored in the S3 Standard storage class.</p>

Amazon S3 storage class	Considerations
S3 Glacier Flexible Retrieval	<p>Choose S3 Glacier Flexible Retrieval for more active archives.</p> <p>Objects stored in S3 Glacier Flexible Retrieval can incur additional charges for overwriting, deleting, or retrieving. Consider how often these objects change, how long you plan to keep these objects, and how often you need to access them. Changes to object data or metadata are equivalent to deleting an object and creating a new one to replace it. This results in additional charges for objects stored in the S3 Glacier Flexible Retrieval storage class.</p> <p>Objects less than 40 KB are smaller than the minimum capacity charge per object in the S3 Glacier Flexible Retrieval storage class. These objects are stored in the S3 Standard storage class.</p> <p>You must restore objects archived in this storage class before DataSync can read them. For information, see <a href="#">Working with archived objects</a> in the <i>Amazon S3 User Guide</i>.</p> <p>When using S3 Glacier Flexible Retrieval, choose the <b>Verify only the data transferred</b> task option to compare data and metadata checksums at the end of the transfer. You can't use the <b>Verify all data in the destination</b> option for this storage class because it requires retrieving all existing objects from the destination.</p>

Amazon S3 storage class	Considerations
S3 Glacier Deep Archive	<p>Choose S3 Glacier Deep Archive to archive your objects for long-term data retention and digital preservation where data is accessed once or twice a year.</p> <p>Objects stored in S3 Glacier Deep Archive can incur additional charges for overwriting, deleting, or retrieving. Consider how often these objects change, how long you plan to keep these objects, and how often you need to access them. Changes to object data or metadata are equivalent to deleting an object and creating a new one to replace it. This results in additional charges for objects stored in the S3 Glacier Deep Archive storage class.</p> <p>Objects less than 40 KB are smaller than the minimum capacity charge per object in the S3 Glacier Deep Archive storage class. These objects are stored in the S3 Standard storage class.</p> <p>You must restore objects archived in this storage class before DataSync can read them. For information, see <a href="#">Working with archived objects</a> in the <i>Amazon S3 User Guide</i>.</p> <p>When using S3 Glacier Deep Archive, choose the <b>Verify only the data transferred</b> task option to compare data and metadata checksums at the end of the transfer. You can't use the <b>Verify all data in the destination</b> option for this storage class because it requires retrieving all existing objects from the destination.</p>
S3 Outposts	The storage class for Amazon S3 on Outposts.

## Evaluating S3 request costs when using DataSync

With Amazon S3 locations, you incur costs related to S3 API requests made by DataSync. This section can help you understand how DataSync uses these requests and how they might affect your [Amazon S3 costs](#).

### Topics

- [S3 requests made by DataSync](#)
- [Cost considerations](#)

## S3 requests made by DataSync

The following table describes the S3 requests that DataSync can make when you're copying data to or from an Amazon S3 location.

S3 request	How DataSync uses it
<a href="#">ListObjectV2</a>	DataSync makes at least one LIST request for every object ending in a forward slash (/) to list the objects that start with that prefix. This request is called during a task's <a href="#">preparing</a> phase.
<a href="#">HeadObject</a>	DataSync makes HEAD requests to retrieve object metadata during a task's <a href="#">preparing</a> and <a href="#">verifying</a> phases. There can be multiple HEAD requests per object depending on how you want DataSync to <a href="#">verify the integrity of the data it transfers</a> .
<a href="#">GetObject</a>	DataSync makes GET requests to read data from an object during a task's <a href="#">transferring</a> phase. There can be multiple GET requests for large objects.
<a href="#">GetObjectTagging</a>	If you configure your task to <a href="#">copy object tags</a> , DataSync makes these GET requests to check for object tags during the task's <a href="#">preparing</a> and <a href="#">transferring</a> phases.
<a href="#">PutObject</a>	DataSync makes PUT requests to create objects and prefixes in a destination S3 bucket during a task's <a href="#">transferring</a> phase. Since DataSync uses the <a href="#">Amazon S3 multipart</a>

S3 request	How DataSync uses it
	<a href="#">upload feature</a> , there can be multiple PUT requests for large objects.
<a href="#">PutObjectTagging</a>	If your source objects have tags and you configure your task to <a href="#">copy object tags</a> , DataSync makes these PUT requests when <a href="#">transferring</a> those tags.
<a href="#">CopyObject</a>	DataSync makes a COPY request to create a copy of an object only if that object's metadata changes. This can happen if you originally copied data to the S3 bucket using another service or tool that didn't carry over its metadata.

## Cost considerations

DataSync makes S3 requests on S3 buckets every time you run your task. This can lead to charges adding up in certain situations. For example:

- You're frequently transferring objects to or from an S3 bucket.
- You may not be transferring much data, but your S3 bucket has lots of objects in it. You can still see high charges in this scenario because DataSync makes S3 requests on each of the bucket's objects.
- You're transferring between S3 buckets, so DataSync is making S3 requests on the source and destination.

To help minimize S3 request costs related to DataSync, consider the following:

## Topics

- [What S3 storage classes am I using?](#)
- [How often do I need to transfer my data?](#)

## What S3 storage classes am I using?

S3 request charges can vary based on the Amazon S3 storage class your objects are using, particularly for classes that archive objects (such as S3 Glacier Instant Retrieval, S3 Glacier Flexible Retrieval, and S3 Glacier Deep Archive).

Here are some scenarios in which storage classes can affect your S3 request charges when using DataSync:

- Each time you run a task, DataSync makes HEAD requests to retrieve object metadata. These requests result in charges even if you aren't moving any objects. How much these requests affect your bill depends on the storage class your objects are using along with the number of objects that DataSync scans.
- If you moved objects into the S3 Glacier Instant Retrieval storage class (either directly or through a bucket lifecycle configuration), requests on objects in this class are more expensive than objects in other storage classes.
- If you configure your DataSync task to [verify that your source and destination locations are fully synchronized](#), there will be GET requests for each object in all storage classes (except S3 Glacier Flexible Retrieval and S3 Glacier Deep Archive).
- In addition to GET requests, you incur data retrieval costs for objects in the S3 Standard-IA, S3 One Zone-IA, or S3 Glacier Instant Retrieval storage class.

For more information, see [Amazon S3 pricing](#).

## How often do I need to transfer my data?

If you need to move data on a recurring basis, think about a [schedule](#) that doesn't run more tasks than you need.

You may also consider limiting the scope of your transfers. For example, you can configure DataSync to focus on objects in certain prefixes or [filter what data gets transferred](#). These options can help reduce the number of S3 requests made each time you run your DataSync task.

## Other considerations with Amazon S3 transfers

When using Amazon S3 with DataSync, remember the following:

- Changes to object data or metadata are equivalent to deleting and replacing an object. These changes result in additional charges in the following scenarios:

- **When using object versioning** – Changes to object data or metadata create a new version of the object.
- **When using storage classes that can incur additional charges for overwriting, deleting, or retrieving objects** – Changes to object data or metadata result in such charges. For more information, see [Storage class considerations with Amazon S3 transfers](#).
- When using object versioning in Amazon S3, running a DataSync transfer task once might create more than one version of an Amazon S3 object.
- DataSync might not transfer an object if it has nonstandard characters in its name. For more information, see the [object key naming guidelines](#) in the *Amazon S3 User Guide*.
- To help minimize your Amazon S3 storage costs, we recommend using a [lifecycle configuration](#) to stop incomplete multipart uploads.
- After initially transferring data from an S3 bucket to a file system (for example, NFS or Amazon FSx), subsequent runs of the same DataSync task won't include objects that have been modified but are the same size they were during the first transfer.
- If you're transferring from an S3 bucket, use [Amazon S3 Storage Lens](#) to figure out how many objects you're moving.

#### Tip

When transferring between S3 buckets, DataSync can't work with more than 25 million objects per task execution. If there are more than 25 million objects involved, we recommend a couple options:

- [Organizing your objects](#) using prefixes that don't include more than 25 million objects. You can then create separate DataSync tasks for each prefix.
- [Filtering the data](#) transferred by DataSync.

## Creating your Amazon S3 transfer location

To create an Amazon S3 location for your transfer, you need an existing S3 bucket. If you don't have one, see [Getting started with Amazon S3](#) in the *Amazon S3 User Guide*.

**Tip**

If your S3 bucket has objects with different storage classes, learn how [DataSync works with these storage classes](#) and how it can [affect your AWS bill](#).

**Using the DataSync console**

1. Open the AWS DataSync console at <https://console.aws.amazon.com/datasync/>.
2. In the left navigation pane, expand **Data transfer**, then choose **Locations** and **Create location**.
3. For **Location type**, choose **Amazon S3**.
4. For **S3 bucket**, choose the bucket that you want to use as a location.

If your S3 bucket is located on an AWS Outposts resource, you must specify an Amazon S3 access point. For more information, see [Managing data access with Amazon S3 access points](#) in the *Amazon S3 User Guide*.

5. For **S3 storage class**, choose a storage class that you want your objects to use when Amazon S3 is a transfer destination.

For more information, see [Storage class considerations with Amazon S3 transfers](#).

6. (Amazon S3 on Outposts only) For **Agents**, specify the Amazon Resource Name (ARN) of the DataSync agent on your Outpost.

For more information, see [Deploying your agent on AWS Outposts](#).

7. For **Folder**, enter a prefix in the S3 bucket that DataSync reads from or writes to (depending on whether the bucket is a source or destination location).

**Warning**

DataSync can't transfer objects with a prefix that begins with a slash (/) or includes //, /./, or /../ patterns. For example:

- /photos
- photos//2006/January
- photos/./2006/February
- photos/../2006/March



8. For **IAM role**, do one of the following:

- Choose **Autogenerate** for DataSync to automatically create an IAM role with the permissions required to access the S3 bucket.

If DataSync previously created an IAM role for this S3 bucket, that role is chosen by default.

- Choose a custom IAM role that you created. For more information, see [Creating an IAM role for DataSync to access your Amazon S3 location](#).

9. (Optional) Choose **Add tag** to tag your Amazon S3 location.

Tags can help you manage, filter, and search for your resources. We recommend creating a name tag for your location.

10. Choose **Create location**.

## Using the AWS CLI

1. Copy the following `create-location-s3` command:

```
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
```

2. For `--s3-bucket-arn`, specify the ARN of the S3 bucket that you want to use as a location.

(When creating your DataSync task later, you specify whether this location is a source or destination location.)

If your S3 bucket is located on an AWS Outposts resource, you must specify an Amazon S3 access point. For more information, see [Managing data access with Amazon S3 access points](#) in the *Amazon S3 User Guide*.

3. For `--s3-storage-class`, specify a storage class that you want your objects to use when Amazon S3 is a transfer destination.

For more information, see [Storage class considerations with Amazon S3 transfers](#). DataSync by default uses the S3 Outposts storage class for Amazon S3 on Outposts.

4. For `--s3-config`, specify the ARN of the IAM role that DataSync needs to access your bucket.

For more information, see [Creating an IAM role for DataSync to access your Amazon S3 location](#).

- For `--subdirectory`, specify a prefix in the S3 bucket that DataSync reads from or writes to (depending on whether the bucket is a source or destination location).

#### **Warning**

DataSync can't transfer objects with a prefix that begins with a slash (/) or includes //, /./, or ../ patterns. For example:

- `/photos`
- `photos//2006/January`
- `photos/./2006/February`
- `photos/../2006/March`

- (Amazon S3 on Outposts only) Add the `--agent-arns` option to the command and use it to specify the ARN of the DataSync agent on your Outpost.

For more information, see [Deploying your agent on AWS Outposts](#).

- Run the `create-location-s3` command.

If the command is successful, you get a response that shows you the ARN of the location that you created. For example:

```
{
  "LocationArn": "arn:aws:datsync:us-east-1:111222333444:location/
loc-0b3017fc4ba4a2d8d"
}
```

You can use this location as a source or destination for your DataSync task.

## Amazon S3 transfers across AWS accounts

With DataSync, you can move data to or from S3 buckets in [different AWS accounts](#). For more information, see the following tutorials:

- [Transferring data from on-premises storage to Amazon S3 across AWS accounts](#)

- [Transferring data from Amazon S3 to Amazon S3 across AWS accounts](#)

## Amazon S3 transfers between commercial and AWS GovCloud (US) Regions

By default, DataSync doesn't transfer between S3 buckets in commercial and AWS GovCloud (US) Regions. You can still set up this kind of transfer, though, by creating an object storage location for one of the S3 buckets in your transfer. This type of location requires a DataSync agent.

**Before you begin:** Make sure that you understand the cost implications of transferring between Regions. For more information, see [AWS DataSync Pricing](#).

### Contents

- [Providing DataSync access to your object storage location's bucket](#)
- [Creating your DataSync agent](#)
- [Creating an object storage location for your S3 bucket](#)

### Providing DataSync access to your object storage location's bucket

When creating the object storage location for this transfer, you must provide DataSync the credentials of an IAM user with permission to access the location's S3 bucket. For more information, see [Required permissions](#).

#### Warning

IAM users have long-term credentials, which presents a security risk. To help mitigate this risk, we recommend that you provide these users with only the permissions they require to perform the task and that you remove these users when they are no longer needed.

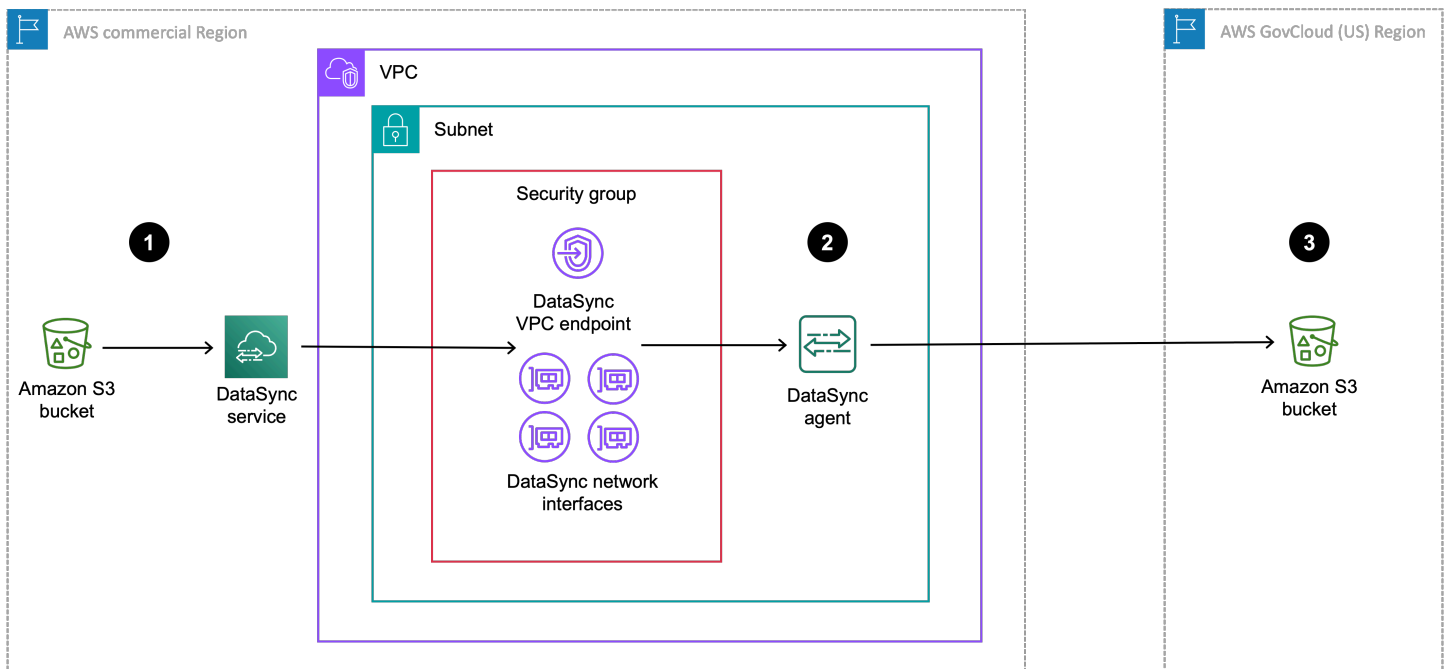
### Creating your DataSync agent

Since you're transferring between a commercial and AWS GovCloud (US) Region, you deploy your DataSync agent as an Amazon EC2 instance in one of the Regions. We recommend that your agent use a VPC service endpoint to avoid data transfer charges out to the public internet. For more information, see [Amazon EC2 Data Transfer pricing](#).

Choose one of the following scenarios that describe how to create an agent based on the Region where you plan to run your DataSync task.

## When running a DataSync task in a commercial Region

The following diagram shows a transfer where your DataSync task and agent are in the commercial Region.



Reference	Description
1	In the commercial Region where you're running a DataSync task, data transfers from the source S3 bucket. The source bucket is configured as an <a href="#">Amazon S3 location</a> in the commercial Region.
2	Data transfers through the DataSync agent, which is in the same VPC and subnet where the VPC service endpoint and <a href="#">network interfaces</a> are located.
3	Data transfers to the destination S3 bucket in the AWS GovCloud (US) Region. The destination bucket is configured as an <a href="#">object storage location</a> in the commercial Region.

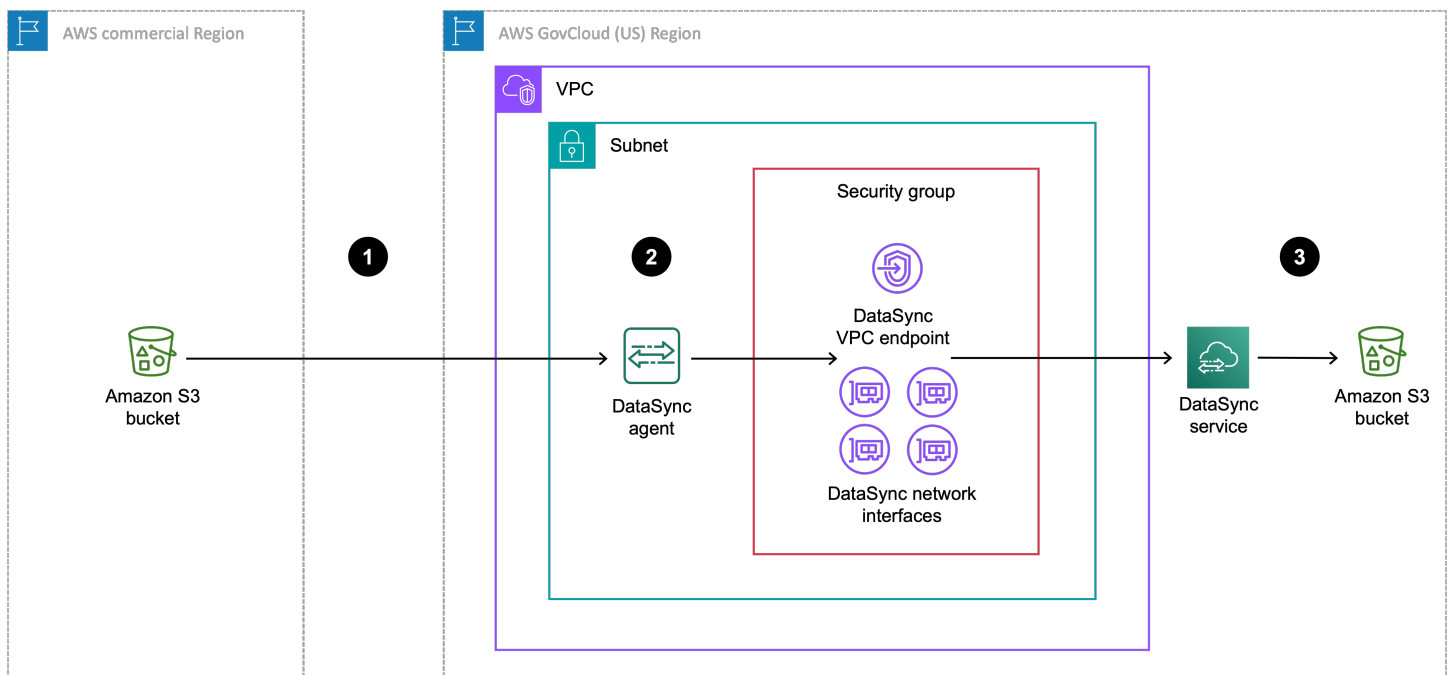
You can use this same setup to transfer the opposite direction, too, from the AWS GovCloud (US) Region to the commercial Region.

## To create your DataSync agent

1. [Deploy an Amazon EC2 agent](#) in your commercial Region.
2. Configure your agent to use a [VPC service endpoint](#).
3. [Activate your agent](#).

## When running a DataSync task in a GovCloud (US) Region

The following diagram shows a transfer where your DataSync task and agent are in the AWS GovCloud (US) Region.



Reference	Description
1	Data transfers from the source S3 bucket in the commercial Region to the AWS GovCloud (US) Region where you're running a DataSync task. The source bucket is configured as an <a href="#">object storage location</a> in the AWS GovCloud (US) Region.
2	In the AWS GovCloud (US) Region, data transfers through the DataSync agent in the same VPC and subnet where the VPC service endpoint and <a href="#">network interfaces</a> are located.

Reference	Description
3	Data transfers to the destination S3 bucket in the AWS GovCloud (US) Region. The destination bucket is configured as an <a href="#">Amazon S3 location</a> in the AWS GovCloud (US) Region.

You can use this same setup to transfer the opposite direction, too, from the AWS GovCloud (US) Region to the commercial Region.

### To create your DataSync agent

1. [Deploy an Amazon EC2 agent](#) in your AWS GovCloud (US) Region.
2. Configure your agent to use a [VPC service endpoint](#).
3. [Activate your agent](#).

If your dataset is highly compressible, you might see reduced costs by instead creating your agent in a commercial Region while running a task in an AWS GovCloud (US) Region. There's more setup than normal for creating this agent, including preparing the agent for use in a commercial Region. For information about creating an agent for this setup, see the [Move data in and out of AWS GovCloud \(US\) with AWS DataSync](#) blog.

### Creating an object storage location for your S3 bucket

You need an object storage location for the S3 bucket that's in the Region where you aren't running your DataSync task.

### Using the DataSync console

1. Open the AWS DataSync console at <https://console.aws.amazon.com/datasync/>.
2. Make sure that you're in the same Region where you plan to run your task.
3. In the left navigation pane, expand **Data transfer**, then choose **Locations** and **Create location**.
4. For **Location type**, choose **Object storage**.
5. For **Agents**, choose the DataSync agent that you created for this transfer.
6. For **Server**, enter an Amazon S3 endpoint for your bucket by using one of the following formats:
  - **Commercial Region bucket:** `s3.your-region.amazonaws.com`

- **AWS GovCloud (US) Region bucket:** `s3.your-gov-region.amazonaws.com`

For a list of Amazon S3 endpoints, see the [AWS General Reference](#).

7. For **Bucket** name, enter the name of the S3 bucket.
8. For **Folder**, enter a prefix in the S3 bucket that DataSync reads from or writes to (depending on whether the bucket is a source or destination location).

### **Warning**

DataSync can't transfer objects with a prefix that begins with a slash (/) or includes //, /./, or /../ patterns. For example:

- `/photos`
- `photos//2006/January`
- `photos/./2006/February`
- `photos/../2006/March`

9. Select **Requires credentials** and do the following:
  - For **Access key**, enter the access key for an [IAM user](#) that can access the bucket.
  - For **Secret key**, enter the same IAM user's secret key.
10. (Optional) Choose **Add tag** to tag your location.

Tags can help you manage, filter, and search for your resources. We recommend creating a name tag for your location.

11. Choose **Create location**.

## Using the AWS CLI

1. Copy the following create-location-object-storage command:

```
aws datasync create-location-object-storage \  
  --server-hostname s3-endpoint \  
  --bucket-name your-bucket \  
  --agent-arns arn:aws:datasync:your-region:123456789012:agent/  
agent-01234567890deadfb
```

2. For the `--server-hostname` parameter, specify an Amazon S3 endpoint for your bucket by using one of the following formats:

- **Commercial Region bucket:** `s3.your-region.amazonaws.com`
- **AWS GovCloud (US) Region bucket:** `s3.your-gov-region.amazonaws.com`

For the Region in the endpoint, make sure that you specify the same Region where you plan to run your task.

For a list of Amazon S3 endpoints, see the [AWS General Reference](#).

3. For the `--bucket-name` parameter, specify the name of the S3 bucket.
4. For the `--agent-arns` parameter, specify the DataSync agent that you created for this transfer.
5. For the `--access-key` parameter, specify the access key for an [IAM user](#) that can access the bucket.
6. For the `--secret-key` parameter, enter the same IAM user's secret key.
7. (Optional) For the `--subdirectory` parameter, specify a prefix in the S3 bucket that DataSync reads from or writes to (depending on whether the bucket is a source or destination location).

#### Warning

DataSync can't transfer objects with a prefix that begins with a slash (/) or includes //, /./, or /../ patterns. For example:

- `/photos`
- `photos//2006/January`
- `photos/./2006/February`
- `photos/../2006/March`

8. (Optional) For the `--tags` parameter, specify key-value pairs that represent tags for the location resource.

Tags can help you manage, filter, and search for your resources. We recommend creating a name tag for your location.

9. Run the `create-location-object-storage` command.



You get a response that shows you the location ARN that you just created.

```
{
  "LocationArn": "arn:aws:datasync:us-east-1:123456789012:location/
loc-01234567890abcdef"
}
```

You can use this location as a source or destination for your DataSync task. For the other S3 bucket in this transfer, [create an Amazon S3 location](#).

## Next steps

Some possible next steps include:

1. If needed, create your other location. For more information, see [Where can I transfer my data with AWS DataSync?](#)
2. [Configure DataSync task settings](#), such as what files to transfer, how to handle metadata, among other options.
3. [Set a schedule](#) for your DataSync task.
4. [Configure monitoring](#) for your DataSync task.
5. [Start](#) your task.

## Configuring AWS DataSync transfers with Amazon EFS

To transfer data to or from your Amazon EFS file system, you must create an AWS DataSync transfer *location*. DataSync can use this location as a source or destination for transferring data.

### Providing DataSync access to Amazon EFS file systems

DataSync mounts your Amazon EFS file system as the root user from your virtual private cloud (VPC) using [network interfaces](#).

When creating your location, you specify the subnet and security groups that DataSync uses to connect to one of your Amazon EFS file system's mount targets or [access points](#) using Network File System (NFS) port 2049.

**Note**

If DataSync accesses a destination file system through an access point that [enforces user identity](#), the POSIX user and group IDs for your source data aren't preserved if you configure your DataSync task to [copy ownership](#). Instead, the transferred files and folders are set to the access point's user and group IDs. When this happens, task verification fails because DataSync detects a mismatch between metadata in the source and destination locations.

DataSync can also mount Amazon EFS file systems configured for restricted access. For example, you can specify an AWS Identity and Access Management (IAM) role that gives DataSync the necessary level of permission to connect to your file system. For more information, see [Using IAM policies to access your Amazon EFS file system](#).

## Network considerations with Amazon EFS transfers

VPCs that you use with DataSync must have default tenancy. VPCs with dedicated tenancy are not supported. For more information, see [Work with VPCs](#).

## Performance considerations with Amazon EFS transfers

Your Amazon EFS file system's throughput mode can affect transfer duration and file system performance during the transfer. Consider the following:

- For best results, we recommend using Elastic throughput mode. If you don't use Elastic throughput mode, your transfer might take longer.
- If you use Bursting throughput mode, the performance of your file system's applications might be affected because DataSync consumes file system burst credits.
- How you [configure DataSync to verify your transferred data](#) can affect file system performance and data access costs.

For more information, see [Amazon EFS performance](#) in the *Amazon Elastic File System User Guide* and the [Amazon EFS Pricing](#) page.

## Creating your Amazon EFS transfer location

To create the transfer location, you need an existing Amazon EFS file system. If you don't have one, see [Getting started with Amazon Elastic File System](#) in the *Amazon Elastic File System User Guide*.

## To create an Amazon EFS location by using the console

1. Open the AWS DataSync console at <https://console.aws.amazon.com/datasync/>.
2. In the left navigation pane, expand **Data transfer**, then choose **Locations** and **Create location**.
3. For **Location type**, choose **Amazon EFS file system**.

You configure this location as a source or destination later.

4. For **File system**, choose the Amazon EFS file system that you want to use as a location.

You configure this location as a source or destination later.

5. For **Mount path**, enter a mount path for your Amazon EFS file system.

This specifies where DataSync reads or writes data (depending on if this is a source or destination location).

By default, DataSync uses the root directory (or access point if you configure one). You can also specify subdirectories using forward slashes (for example, /path/to/directory).

6. For **Subnet** choose a subnet where DataSync creates the network interfaces for managing traffic during your transfer.

The subnet must be located:

- In the same VPC as the Amazon EFS file system.
- In the same Availability Zone as at least one file system mount target.

### Note


You don't need to specify a subnet that includes a file system mount target.

7. For **Security groups**, choose the security groups associated with an Amazon EFS file system's mount target.

### Note

The security groups that you specify must allow inbound traffic on NFS port 2049. For more information, see [Using VPC security groups for Amazon EC2 instances and mount targets](#) in the *Amazon Elastic File System User Guide*.

8. For **In-transit encryption**, choose whether you want DataSync to use Transport Layer Security (TLS) encryption when it copies data to or from your file system.

 **Note**

You must enable this setting if you want to configure an access point, IAM role, or both with your location.

9. (Optional) For **EFS access point**, choose an access point that DataSync can use to mount your Amazon EFS file system.
10. (Optional) For **IAM role**, specify a role that allows DataSync to access your file system.

For information on creating this role, see [Using IAM policies to access your Amazon EFS file system](#)


11. (Optional) Select **Add tag** to tag your file system.

A *tag* is a key-value pair that helps you manage, filter, and search for your locations.

12. Choose **Create location**.

## Using IAM policies to access your Amazon EFS file system

You can configure your Amazon EFS file system with a higher level of security by using IAM policies. In your [file system policy](#), you can specify an IAM role that still allows DataSync to connect with the file system.

 **Note**

To use an IAM role, you must enable TLS for in-transit encryption when creating a DataSync location for your file system.

For more information, see [Using IAM to control file system data access](#) in the *Amazon Elastic File System User Guide*.

## Creating an IAM role for DataSync

Create a role with DataSync as the trusted entity.

## To create the IAM role

1. Open the IAM console at <https://console.aws.amazon.com/iam/>.
2. In the left navigation pane, under **Access management**, choose **Roles**, and then choose **Create role**.
3. On the **Select trusted entity** page, for **Trusted entity type**, choose **Custom trust policy**.
4. Paste the following JSON into the policy editor:

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

5. Choose **Next**. On the **Add permissions** page, choose **Next**.
6. Give your role a name and choose **Create role**.

Specify this role when creating the location for your Amazon EFS file system.

### Example Amazon EFS file system policy

The following sample IAM policy includes elements that help restrict access to an Amazon EFS file system (identified in the policy as fs-*1234567890abcdef0*):

- **Principal**: Specifies an IAM role that gives DataSync permission to connect to the file system.
- **Action**: Gives DataSync root access and allows it to read from and write to the file system.
- **aws:SecureTransport**: Requires NFS clients to use TLS when connecting to the file system.
- **elasticfilesystem:AccessPointArn**: Allows access to the file system only through a specific access point.

```
{
  "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"
    }
  }
}]
}

```

## Configuring AWS DataSync transfers with Amazon FSx for Windows File Server

To transfer data to or from your Amazon FSx for Windows File Server file system, you must create an AWS DataSync transfer *location*. DataSync can use this location as a source or destination for transferring data.

### Accessing FSx for Windows File Server file systems

DataSync connects to your FSx for Windows File Server file system with the Server Message Block (SMB) protocol and mounts it from your virtual private cloud (VPC) using [network interfaces](#).

#### Note

VPCs that you use with DataSync must have default tenancy. VPCs with dedicated tenancy are not supported. For more information, see [Work with VPCs](#).

## Topics

- [Required permissions](#)
- [Required authentication protocols](#)
- [DFS Namespaces](#)

## Required permissions

You must provide DataSync a user with the necessary rights to mount and access your FSx for Windows File Server files, folders, and file metadata.

We recommend that this user belong to a Microsoft Active Directory group for administering your file system. The specifics of this group depends on your Active Directory setup:

- If you're using AWS Directory Service for Microsoft Active Directory with FSx for Windows File Server, the user must be a member of the **AWS Delegated FSx Administrators** group.
- If you're using self-managed Active Directory with FSx for Windows File Server, the user must be a member of one of two groups:
  - The **Domain Admins** group, which is the default delegated administrators group.
  - A custom delegated administrators group with user rights that allow DataSync to copy object ownership permissions and Windows access control lists (ACLs).

### Important

You can't change the delegated administrators group after the file system has been deployed. You must either redeploy the file system or restore it from a backup to use the custom delegated administrator group with the following user rights that DataSync needs to copy metadata.

User right	Description
<b>Restore files and directories</b> (SE_RESTORE_NAME )	Allows DataSync to copy object ownership , permissions, file metadata, and NTFS discretionary access lists (DACLS).

User right	Description
	This user right is usually granted to members of the <b>Domain Admins</b> and <b>Backup Operators</b> groups (both of which are default Active Directory groups).
<b>Manage auditing and security log</b> (SE_SECURITY_NAME )	Allows DataSync to copy NTFS system access control lists (SACLs).  This user right is usually granted to members of the <b>Domain Admins</b> group.

- If you want to copy Windows ACLs and are transferring between an SMB file server and FSx for Windows File Server file system or between FSx for Windows File Server file systems, the users that you provide DataSync must belong to the same Active Directory domain or have an Active Directory trust relationship between their domains.

### **Warning**

Your FSx for Windows File Server file system's SYSTEM user must have **Full control** permissions on all folders in your file system. Do not change the NTFS ACL permissions for this user on your folders. If you do, DataSync can change your file system's permissions in a way that makes your file share inaccessible and prevents file system backups from being usable. For more information on file- and folder-level access, see the [Amazon FSx for Windows File Server User Guide](#).

## Required authentication protocols

Your FSx for Windows File Server must use NTLM authentication for DataSync to access it. DataSync can't access a file server that uses Kerberos authentication.

## DFS Namespaces

DataSync doesn't support Microsoft Distributed File System (DFS) Namespaces. We recommend specifying an underlying file server or share instead when creating your DataSync location.



For more information, see [Grouping multiple file systems with DFS Namespaces](#) in the *Amazon FSx for Windows File Server User Guide*.

## Creating your FSx for Windows File Server transfer location

Before you begin, make sure that you have an existing FSx for Windows File Server in your AWS Region. For more information, see [Getting started with Amazon FSx](#) in the *Amazon FSx for Windows File Server User Guide*.

### To create an FSx for Windows File Server location by using the DataSync console

1. Open the AWS DataSync console at <https://console.aws.amazon.com/datasync/>.
2. In the left navigation pane, expand **Data transfer**, then choose **Locations** and **Create location**.
3. For **Location type**, choose **Amazon FSx**.
4. For **FSx file system**, choose the FSx for Windows File Server file system that you want to use as a location.
5. For **Share name**, enter a mount path for your FSx for Windows File Server using forward slashes.

This specifies the path where DataSync reads or writes data (depending on if this is a source or destination location).

You can also include subdirectories (for example, `/path/to/directory`).

6. For **Security groups**, choose up to five security groups that provide access to your file system's preferred subnet.

#### Note

If you choose a security group that doesn't allow connections from within itself, do one of the following:

- Configure the security group to allow it to communicate within itself.
- Choose a different security group that can communicate with the mount target's security group.

7. For **User**, enter the name of a user that can access your FSx for Windows File Server.

For more information, see [Required permissions](#).

8. For **Password**, enter password of the user name.
9. (Optional) For **Domain**, enter the name of the Windows domain that your FSx for Windows File Server file system belongs to.

If you have multiple Active Directory domains in your environment, configuring this setting makes sure that DataSync connects to the right file system.

10. (Optional) Enter values for the **Key** and **Value** fields to tag the FSx for Windows File Server.

Tags help you manage, filter, and search for your AWS resources. We recommend creating at least a name tag for your location.

11. Choose **Create location**.

## Configuring AWS DataSync transfers with Amazon FSx for Lustre

To transfer data to or from your Amazon FSx for Lustre file system, you must create an AWS DataSync transfer *location*. DataSync can use this location as a source or destination for transferring data.

### Accessing FSx for Lustre file systems

DataSync accesses your FSx for Lustre file system using the Lustre client. DataSync requires access to all data on your FSx for Lustre file system. To have this level of access, DataSync mounts your file system as the root user using a user ID (UID) and group ID (GID) of 0.

DataSync mounts your file system from your virtual private cloud (VPC) using [network interfaces](#). DataSync fully manages the creation, the use, and the deletion of these network interfaces on your behalf.

#### Note

VPCs that you use with DataSync must have default tenancy. VPCs with dedicated tenancy are not supported. For more information, see [Work with VPCs](#).

### Creating your FSx for Lustre transfer location

If you don't have an FSx for Lustre in your AWS Region, create one. For more information, see [Getting started with Amazon FSx for Lustre](#) in the *Amazon FSx for Lustre User Guide*.

## To create an FSx for Lustre location by using the DataSync console

1. Open the AWS DataSync console at <https://console.aws.amazon.com/datasync/>.
2. In the left navigation pane, expand **Data transfer**, then choose **Locations** and **Create location**.
3. For **Location type**, choose **Amazon FSx**.

You configure this location as a source or destination later.

4. For **FSx file system**, choose the FSx for Lustre file system that you want to use as a location.
5. For **Mount path**, enter the mount path for your FSx for Lustre file system.

The path can include a subdirectory. When the location is used as a source, DataSync reads data from the mount path. When the location is used as a destination, DataSync writes all data to the mount path. If a subdirectory isn't provided, DataSync uses the root directory (/).

6. For **Security groups**, choose up to five security groups that provide access to your FSx for Lustre file system.

The security groups must be able to access the file system's ports. Also, the file system must allow access from the security groups.

For more information about security groups, see [File System Access Control with Amazon VPC](#) in the *Amazon FSx for Lustre User Guide*.

7. (Optional) Enter values for the **Key** and **Value** fields to tag the FSx for Lustre file system.

Tags help you manage, filter, and search for your AWS resources. We recommend creating at least a name tag for your location.

8. Choose **Create location**.

## Configuring AWS DataSync transfers with Amazon FSx for OpenZFS

To transfer data to or from your Amazon FSx for OpenZFS file system, you must create an AWS DataSync transfer *location*. DataSync can use this location as a source or destination for transferring data.

### Accessing FSx for OpenZFS file systems

DataSync mounts your FSx for OpenZFS file system from your virtual private cloud (VPC) using [network interfaces](#). DataSync fully manages the creation, the use, and the deletion of these network interfaces on your behalf.

**Note**

VPCs that you use with DataSync must have default tenancy. VPCs with dedicated tenancy are not supported. For more information, see [Work with VPCs](#).

## Configuring FSx for OpenZFS file system authorization

DataSync accesses your FSx for OpenZFS file system as an NFS client, mounting the file system as a root user with a user ID (UID) and group ID (GID) of 0.

For DataSync to copy all of your file metadata, you must configure the NFS export settings on your file system volumes using `no_root_squash`. However, you can limit this level of access to only a specific DataSync task.

For more information, see [Volume properties](#) in the *Amazon FSx for OpenZFS User Guide*.

### Configuring NFS exports specific to DataSync (recommended)

You can configure an NFS export specific to each volume that's accessed only by your DataSync task. Do this for the most recent ancestor volume of the mount path that you specify when creating your FSx for OpenZFS location.

#### To configure an NFS export specific to DataSync

1. Create your [DataSync task](#).

This creates the task's elastic network interfaces that you'll specify in your NFS export settings.

2. Locate the private IP addresses of the task's network interfaces by using the Amazon EC2 console or AWS CLI.
3. For your FSx for OpenZFS file system volume, configure the following NFS export settings for each of the task's network interfaces:
  - **Client address:** Enter the network interface's private IP address (for example, `10.24.34.0`).
  - **NFS options:** Enter `rw,no_root_squash`.

### Configuring NFS exports for all clients

You can specify an NFS export that allows root access to all clients.

## To configure an NFS export for all clients

- For your FSx for OpenZFS file system volume, configure the following NFS export settings:
  - **Client address:** Enter `*`.
  - **NFS options:** Enter `rw,no_root_squash`.

## Creating your FSx for OpenZFS transfer location

To create the location, you need an existing FSx for OpenZFS file system. If you don't have one, see [Getting started with Amazon FSx for OpenZFS](#) in the *Amazon FSx for OpenZFS User Guide*.

### To create an FSx for OpenZFS location by using the DataSync console

1. Open the AWS DataSync console at <https://console.aws.amazon.com/datasync/>.
2. In the left navigation pane, choose **Locations**, and then choose **Create location**.
3. For **Location type**, choose **Amazon FSx**.

You configure this location as a source or destination later.

4. For **FSx file system**, choose the FSx for OpenZFS file system that you want to use as a location.
5. For **Mount path**, enter the mount path for your FSx for OpenZFS file system.

The path must begin with `/fsx` and can be any existing directory path in the file system. When the location is used as a source, DataSync reads data from the mount path. When the location is used as a destination, DataSync writes all data to the mount path. If a subdirectory isn't provided, DataSync uses the root volume directory (for example, `/fsx`).

6. For **Security groups**, choose up to five security groups that provide network access to your FSx for OpenZFS file system.

The security groups must provide access to the network ports that are used by the FSx for OpenZFS file system. The file system must allow network access from the security groups.

For more information about security groups, see [File system access control with Amazon VPC](#) in the *Amazon FSx for OpenZFS User Guide*.

7. (Optional) Expand **Additional settings** and for **NFS version** choose the NFS version that DataSync uses to access your file system.

By default, DataSync uses NFS version 4.1.

8. (Optional) Enter values for the **Key** and **Value** fields to tag the FSx for OpenZFS file system.

Tags help you manage, filter, and search for your location. We recommend creating at least a name tag for your location.

9. Choose **Create location**.

## Configuring AWS DataSync transfers with Amazon FSx for NetApp ONTAP

To transfer data to or from your Amazon FSx for NetApp ONTAP file system, you must create an AWS DataSync transfer *location*. DataSync can use this location as a source or destination for transferring data.

### Accessing FSx for ONTAP file systems

To access an FSx for ONTAP file system, DataSync mounts a storage virtual machine (SVM) on your file system using [network interfaces](#) in your virtual private cloud (VPC). DataSync creates these network interfaces in your file system's preferred subnet only when you create a task that includes your FSx for ONTAP location.

#### Note

VPCs that you use with DataSync must have default tenancy. VPCs with dedicated tenancy are not supported. For more information, see [Work with VPCs](#).

DataSync can connect to an FSx for ONTAP file system's SVM and copy data by using the Network File System (NFS) or Server Message Block (SMB) protocol.

### Topics

- [Using the NFS protocol](#)
- [Using the SMB protocol](#)
- [Unsupported protocols](#)
- [Choosing the right protocol](#)

## Using the NFS protocol

With the NFS protocol, DataSync uses the AUTH\_SYS security mechanism with a user ID (UID) and group ID (GID) of 0 to authenticate with your SVM.

### Note

DataSync currently only supports NFS version 3 with FSx for ONTAP locations.

## Using the SMB protocol

With the SMB protocol, DataSync uses credentials that you provide to authenticate with your SVM.

### Supported SMB versions

By default, DataSync automatically chooses a version of the SMB protocol based on negotiation with your SMB file server. You also can configure DataSync to use a specific version, but we recommend doing this only if DataSync has trouble negotiating with the SMB file server automatically.

See the following table for a list of options in the DataSync console and API for configuring an SMB version with your FSx for ONTAP location:

Console option	API option	Description
Automatic	AUTOMATIC	DataSync and the SMB file server negotiate the highest version of SMB that they mutually support between 2.1 and 3.1.1.  This is the default and recommended option. If you instead choose a specific version that your file server doesn't support, you may get an <code>Operation Not Supported</code> error.
SMB 3.0.2	SMB3	Restricts the protocol negotiation to only SMB version 3.0.2.
SMB 2.1	SMB2	Restricts the protocol negotiation to only SMB version 2.1.

Console option	API option	Description
SMB 2.0	SMB2_0	Restricts the protocol negotiation to only SMB version 2.0.

## Required permissions

You must provide DataSync a local user in your SVM or a domain user in your Microsoft Active Directory with the necessary rights to mount and access your files, folders, and file metadata.

If you provide a user in your Active Directory, note the following:

- If you're using AWS Directory Service for Microsoft Active Directory, the user must be a member of the **AWS Delegated FSx Administrators** group.
- If you're using a self-managed Active Directory, the user must be a member of one of two groups:
  - The **Domain Admins** group, which is the default delegated administrators group.
  - A custom delegated administrators group with user rights that allow DataSync to copy object ownership permissions and Windows access control lists (ACLs).

### Important

You can't change the delegated administrators group after the file system has been deployed. You must either redeploy the file system or restore it from a backup to use the custom delegated administrator group with the following user rights that DataSync needs to copy metadata.

User right	Description
<b>Act as part of the operating system</b> (SE_TCB_NAME )	Allows DataSync to copy object ownership , permissions, file metadata, and NTFS discretionary access lists (DACLS).  This user right is usually granted to members of the <b>Domain Admins</b> and



User right	Description
	<b>Backup Operators</b> groups (both of which are default Active Directory groups).
<b>Manage auditing and security log</b> (SE_SECURITY_NAME )	Allows DataSync to copy NTFS system access control lists (SACLs).  This user right is usually granted to members of the <b>Domain Admins</b> group.

- If you want to copy Windows ACLs and are transferring between FSx for ONTAP file systems using SMB (or other types of file systems using SMB), the users that you provide DataSync must belong to the same Active Directory domain or have an Active Directory trust relationship between their domains.

### Required authentication protocols

For DataSync to access your SMB share, your file system must use NTLM authentication. DataSync can't access the file system if it uses Kerberos authentication.

### DFS Namespaces

DataSync doesn't support Microsoft Distributed File System (DFS) Namespaces. We recommend specifying an underlying file server or share instead when creating your DataSync location.

### Unsupported protocols

DataSync can't access FSx for ONTAP file systems using the iSCSI (Internet Small Computer Systems Interface) protocol.

### Choosing the right protocol

To preserve file metadata in FSx for ONTAP migrations, configure your DataSync source and destination locations to use the same protocol. Between the supported protocols, SMB preserves metadata with the highest fidelity (see [Metadata copied by AWS DataSync](#) for details).

When migrating from a Unix (Linux) server or network-attached storage (NAS) share that serves users through NFS, do the following:

1. [Create an NFS location](#) for the Unix (Linux) server or NAS share. (This will be your source location.)

2. Configure the FSx for ONTAP volume you're transferring data to with the [Unix security style](#).
3. Create a location for your FSx for ONTAP file system that's configured for NFS. (This will be your destination location.)

When migrating from a Windows server or NAS share that serves users through SMB, do the following:

1. [Create an SMB location](#) for the Windows server or NAS share. (This will be your source location.)
2. Configure the FSx for ONTAP volume you're transferring data to with the [NTFS security style](#).
3. Create a location for your FSx for ONTAP file system that's configured for SMB. (This will be your destination location.)

If your FSx for ONTAP environment uses multiple protocols, we recommend working with an AWS storage specialist. To learn about best practices for multiprotocol access, see [Enabling multiprotocol workloads with Amazon FSx for NetApp ONTAP](#).

## Creating your FSx for ONTAP transfer location

To create the location, you need an existing FSx for ONTAP file system. If you don't have one, see [Getting started with Amazon FSx for NetApp ONTAP](#) in the *Amazon FSx for NetApp ONTAP User Guide*.

### To specify an FSx for ONTAP file system by using the DataSync console

1. Open the AWS DataSync console at <https://console.aws.amazon.com/datasync/>.
2. In the left navigation pane, expand **Data transfer**, then choose **Locations** and **Create location**.
3. For **Location type**, choose **Amazon FSx**.

You configure this location as a source or destination later.

4. For **FSx file system**, choose the FSx for ONTAP file system that you want to use as a location.
5. For **Storage virtual machine**, choose a storage virtual machine (SVM) in your file system where you want to copy data to or from.
6. For **Mount path**, specify a path to the file share in that SVM where you'll copy your data.

You can specify a junction path (also known as a mount point), qtree path (for NFS file shares), or share name (for SMB file shares). For example, your mount path might be `/vol1`, `/vol1/tree1`, or `/share1`.

**Tip**

Don't specify a path in the SVM's root volume. For more information, see [Managing FSx for ONTAP storage virtual machines](#) in the *Amazon FSx for NetApp ONTAP User Guide*.

7. For **Security groups**, choose up to five Amazon EC2 security groups that provide access to your file system's preferred subnet.

The security groups must allow outbound traffic on the following ports (depending on the protocol you use):

- **NFS** – TCP ports 111, 635, and 2049
- **SMB** – TCP port 445

Your file system's security groups must also allow inbound traffic on the same ports.

8. For **Protocol**, choose the data transfer protocol that DataSync uses to access your file system's SVM.

For more information, see [Choosing the right protocol](#).

## NFS

DataSync uses NFS version 3.

## SMB

Configure an SMB version, user, password, and Active Directory domain name (if needed) to access the SVM.

- (Optional) Expand **Additional settings** and choose an **SMB version** for DataSync to use when accessing your SVM.

By default, DataSync automatically chooses a version based on negotiation with the SMB file server. For more information, see [Using the SMB protocol](#).

- For **User**, enter a user name that can mount and access the files, folders, and metadata that you want to transfer in the SVM.

For more information, see [Using the SMB protocol](#).

- For **Password**, enter the password of the user that you specified that can access the SVM.

- (Optional) For **Active Directory domain name**, enter the fully qualified domain name (FQDN) of the Active Directory that your SVM belongs to.

If you have multiple domains in your environment, configuring this setting makes sure that DataSync connects to the right SVM.

9. (Optional) Enter values for the **Key** and **Value** fields to tag the FSx for ONTAP file system.

Tags help you manage, filter, and search for your AWS resources. We recommend creating at least a name tag for your location.

10. Choose **Create location**.

## Transferring to or from other cloud storage with AWS DataSync

With AWS DataSync, you can transfer data between some other clouds and AWS storage services. For more information, see [Where can I transfer my data with DataSync?](#).

### Topics

- [Configuring AWS DataSync transfers with Google Cloud Storage](#)
- [Configuring AWS DataSync transfers with Microsoft Azure Blob Storage](#)
- [Configuring AWS DataSync transfers with Microsoft Azure Files SMB shares](#)
- [Configuring AWS DataSync transfers with other cloud object storage](#)

## Configuring AWS DataSync transfers with Google Cloud Storage

The following tutorial shows how you can use AWS DataSync to migrate objects from a Google Cloud Storage bucket to an Amazon S3 bucket.

### Overview

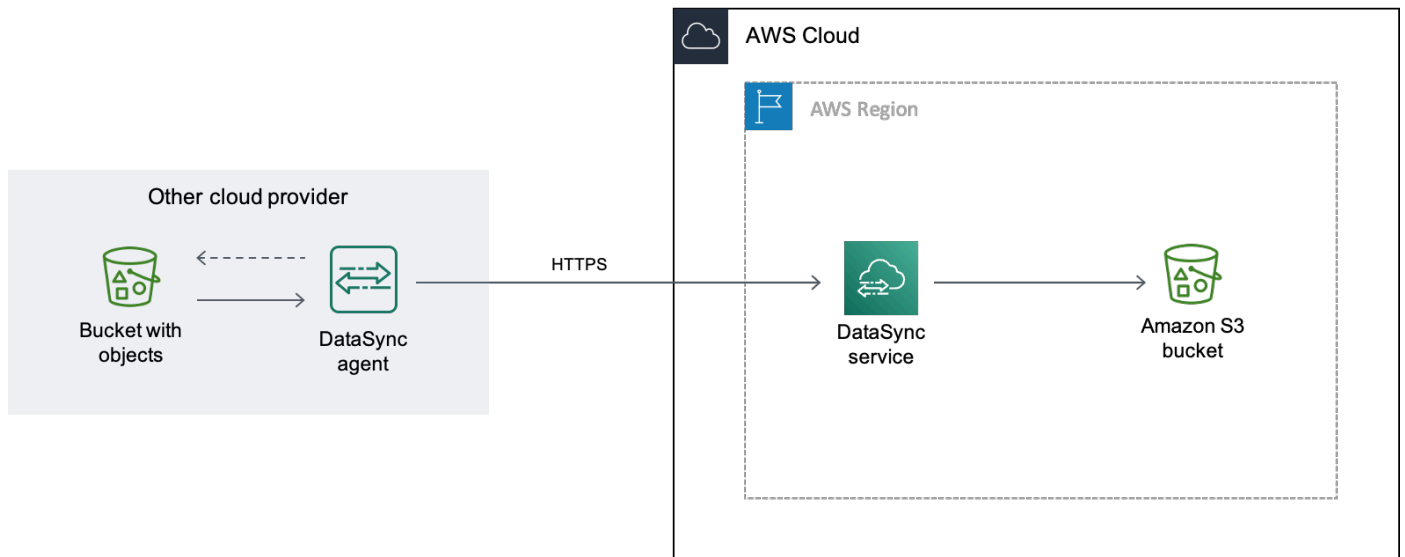
Because DataSync integrates with the [Google Cloud Storage XML API](#), you can copy objects into Amazon S3 without writing code. How this works depends on where you deploy the DataSync agent that facilitates the transfer.

#### Agent in Google Cloud

1. You deploy a DataSync agent in your Google Cloud environment.

2. The agent reads your Google Cloud Storage bucket by using a Hash-based Message Authentication Code (HMAC) key.
3. The objects from your Google Cloud Storage bucket move securely through TLS 1.3 into the AWS Cloud by using a public endpoint.
4. The DataSync service writes the data to your S3 bucket.

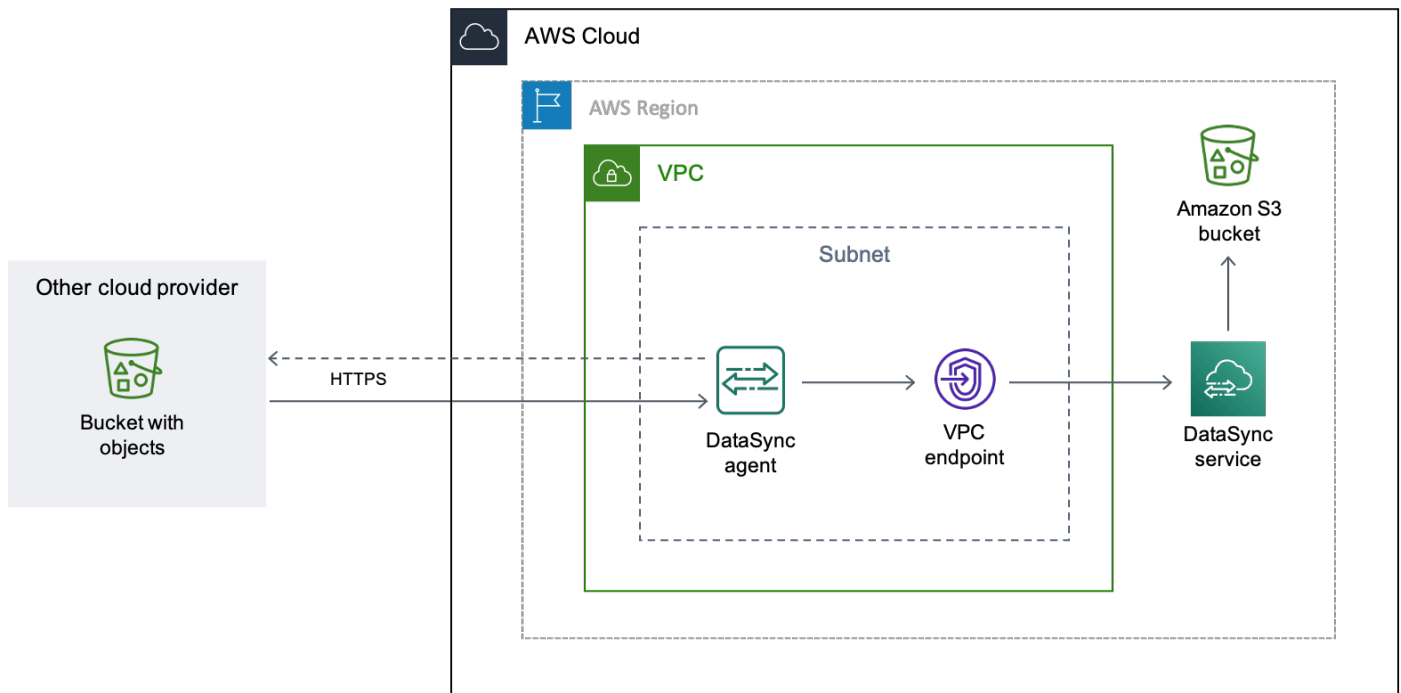
The following diagram illustrates the transfer.



### Agent in your VPC

1. You deploy a DataSync agent in a virtual private cloud (VPC) in your AWS environment.
2. The agent reads your Google Cloud Storage bucket by using a Hash-based Message Authentication Code (HMAC) key.
3. The objects from your Google Cloud Storage bucket move securely through TLS 1.3 into the AWS Cloud by using a private VPC endpoint.
4. The DataSync service writes the data to your S3 bucket.

The following diagram illustrates the transfer.



## Costs

The fees associated with this migration include:

- Running a Google [Compute Engine](#) virtual machine (VM) instance (if you deploy your DataSync agent in Google Cloud)
- Running an [Amazon EC2](#) instance (if you deploy your DataSync agent in a VPC within AWS)
- Transferring the data by using [DataSync](#), including request charges related to [Google Cloud Storage](#) and [Amazon S3](#) (if S3 is one of your transfer locations)
- Transferring data out of [Google Cloud Storage](#)
- Storing data in [Amazon S3](#)

## Prerequisites

Before you begin, do the following if you haven't already:

- [Create a Google Cloud Storage bucket](#) with the objects that you want to transfer to AWS.
- [Sign up for an AWS account.](#)
- [Create an Amazon S3 bucket](#) for storing your objects after they're in AWS.

## Creating an HMAC key for your Google Cloud Storage bucket

DataSync uses an HMAC key that's associated with your Google service account to authenticate with and read the bucket that you're transferring data from. (For detailed instructions on how to create HMAC keys, see the [Google Cloud Storage documentation](#).)

### To create an HMAC key

1. Create an HMAC key for your Google service account.
2. Make sure that your Google service account has at least Storage Object Viewer permissions.
3. Save your HMAC key's access ID and secret in a secure location.

You'll need these items later to configure your DataSync source location.

## Step 2: Configure your network

The network requirements for this migration depend on how you want to deploy your DataSync agent.

### For a DataSync agent in Google Cloud

If you want to host your DataSync agent in Google Cloud, configure your network to [allow DataSync transfers through a public endpoint](#).

### For a DataSync agent in your VPC

If you want to host your agent in AWS, you need a VPC with an interface endpoint. DataSync uses the VPC endpoint to facilitate the transfer.

### To configure your network for a VPC endpoint

1. If you don't have one, [create a VPC](#) in the same AWS Region as your S3 bucket.
2. [Create a private subnet for your VPC](#).
3. [Create a VPC service endpoint](#) for DataSync.
4. Configure your network to [allow DataSync transfers through a VPC endpoint](#).

To make the necessary configuration changes, you can modify the security group that's associated with your VPC endpoint. For more information, see [Control traffic to resources using security groups](#) in the *Amazon VPC User Guide*.

## Step 3: Create a DataSync agent

You need a DataSync agent that can access and read your Google Cloud Storage bucket.

### For Google Cloud

In this scenario, the DataSync agent runs in your Google Cloud environment.

**Before you begin:** [Install the Google Cloud CLI](#).

### To create the agent for Google Cloud

1. Open the AWS DataSync console at <https://console.aws.amazon.com/datasync/>.
2. In the left navigation pane, choose **Agents**, then choose **Create agent**.
3. For **Hypervisor**, choose **VMware ESXi**, then choose **Download the image** to download a .zip file that contains the agent.
4. Open a terminal. Unzip the image by running the following command:

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

5. Extract the contents of the agent's .ova file beginning with aws-datasync by running the following command:

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

6. Import the agent's .vmdk file into Google Cloud by running the following Google Cloud CLI command:

```
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

Importing the .vmdk file might take up to two hours.

7. Create and start a VM instance for the agent image that you just imported.



The instance needs the following configurations for your agent. (For detailed instructions on how to create an instance, see the [Google Cloud Compute Engine documentation](#).)

- For the machine type, choose one of the following:
  - **e2-standard-8** – For DataSync task executions working with up to 20 million objects.
  - **e2-standard-16** – For DataSync task executions working with more than 20 million objects.
- For the boot disk settings, go to the custom images section. Then choose the DataSync agent image that you just imported.
- For the service account setting, choose your Google service account (the same account that you used in [Step 1](#)).
- For the firewall setting, choose the option to allow HTTP (port 80) traffic.

To activate your DataSync agent, port 80 must be open on the agent. The port doesn't need to be publicly accessible. Once activated, DataSync closes the port.

8. After the VM instance is running, take note of its public IP address.

You'll need this IP address to activate the agent.

9. Go back to the DataSync console. On the **Create agent** screen where you downloaded the agent image, do the following to activate your agent:
  - For **Endpoint type**, choose the public service endpoints option (for example, **Public service endpoints in US East Ohio**).
  - For **Activation key**, choose **Automatically get the activation key from your agent**.
  - For **Agent address**, enter the public IP address of the agent VM instance that you just created.
  - Choose **Get key**.
10. Give your agent a name, and then choose **Create agent**.

Your agent is online and ready to move data.

### For your VPC

In this scenario, the agent runs as an Amazon EC2 instance in a VPC that's associated with your AWS account.

**Before you begin:** [Set up the AWS Command Line Interface \(AWS CLI\)](#).

### To create the agent for your VPC

1. Open a terminal. Make sure to configure your AWS CLI profile to use the account that's associated with your S3 bucket.
2. Copy the following command. Replace *vpc-region* with the AWS Region where your VPC resides (for example, us-east-1).

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

3. Run the command. In the output, take note of the "Value" property.

This value is the DataSync Amazon Machine Image (AMI) ID of the Region that you specified. For example, an AMI ID could look like `ami-1234567890abcdef0`.

4. Copy the following URL. Again, replace *vpc-region* with the AWS Region where your VPC resides. Then, replace *ami-id* with the AMI ID that you noted in the previous step.

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

5. Paste the URL into a browser.

The Amazon EC2 instance launch page in the AWS Management Console displays.

6. For **Instance type**, choose one of the [recommended Amazon EC2 instances for DataSync agents](#).
7. For **Key pair**, choose an existing key pair, or create a new one.
8. For **Network settings**, choose the VPC and subnet where you want to deploy the agent.
9. Choose **Launch instance**.
10. Once the Amazon EC2 instance is running, [choose your VPC endpoint](#).
11. [Activate your agent](#).

## Step 4: Create a DataSync source location for your Google Cloud Storage bucket

To set up a DataSync location for your Google Cloud Storage bucket, you need the access ID and secret for the HMAC key that you created in [Step 1](#).

## To create the DataSync source location

1. Open the AWS DataSync console at <https://console.aws.amazon.com/datasync/>.
2. In the left navigation pane, expand **Data transfer**, then choose **Locations** and **Create location**.
3. For **Location type**, choose **Object storage**.
4. For **Agents**, choose the agent that you created in [Step 3](#).
5. For **Server**, enter **storage.googleapis.com**.
6. For **Bucket name**, enter the name of your Google Cloud Storage bucket.
7. Expand **Additional settings**. For **Server protocol**, choose **HTTPS**. For **Server port**, choose **443**.
8. Scroll down to the **Authentication** section. Make sure that the **Requires credentials** check box is selected, and then do the following:
  - For **Access key**, enter your HMAC key's access ID.
  - For **Secret key**, enter your HMAC key's secret.
9. Choose **Create location**.

## Step 5: Create a DataSync destination location for your S3 bucket

You need a DataSync location for where you want your data to end up.

### To create the DataSync destination location

1. Open the AWS DataSync console at <https://console.aws.amazon.com/datasync/>.
2. In the left navigation pane, expand **Data transfer**, then choose **Locations** and **Create location**.
3. [Create a DataSync location for the S3 bucket](#).


If you deployed the DataSync agent in your VPC, this tutorial assumes that the S3 bucket is in the same AWS Region as your VPC and DataSync agent.

## Step 6: Create and start a DataSync task

With your source and destinations locations configured, you can start moving your data into AWS.

### To create and start the DataSync task

1. Open the AWS DataSync console at <https://console.aws.amazon.com/datasync/>.

2. In the left navigation pane, expand **Data transfer**, then choose **Tasks**, and then choose **Create task**.
  3. On the **Configure source location** page, do the following:
    - a. Choose **Choose an existing location**.
    - b. Choose the source location that you created in [Step 4](#), then choose **Next**.
  4. On the **Configure destination location** page, do the following:
    - a. Choose **Choose an existing location**.
    - b. Choose the destination location that you created in [Step 5](#), then choose **Next**.
  5. On the **Configure settings** page, do the following:
    - a. Under **Data transfer configuration**, expand **Additional settings** and clear the **Copy object tags** check box.
-  **Important**

Because DataSync communicates with Google Cloud Storage by using the Amazon S3 API, there's a limitation that might cause your DataSync task to fail if you try to copy object tags.
- b. Configure any other task settings that you want, and then choose **Next**.
  6. On the **Review** page, review your settings, and then choose **Create task**.
  7. On the task's details page, choose **Start**, and then choose one of the following:
    - To run the task without modification, choose **Start with defaults**.
    - To modify the task before running it, choose **Start with overriding options**.

When your task finishes, you'll see the objects from your Google Cloud Storage bucket in your S3 bucket.

## Configuring AWS DataSync transfers with Microsoft Azure Blob Storage

With AWS DataSync, you can transfer data between Microsoft Azure Blob Storage (including Azure Data Lake Storage Gen2 blob storage) and the following AWS storage services:

- [Amazon S3](#)

- [Amazon EFS](#)
- [Amazon FSx for Windows File Server](#)
- [Amazon FSx for Lustre](#)
- [Amazon FSx for OpenZFS](#)
- [Amazon FSx for NetApp ONTAP](#)

To set up this kind of transfer, you create a [location](#) for your Azure Blob Storage. You can use this location as a transfer source or destination.

## Accessing Azure Blob Storage

How DataSync accesses your Azure Blob Storage depends on several factors, including whether you're transferring to or from blob storage and what kind of [shared access signature \(SAS\) token](#) you're using. Your objects also must be in an [access tier](#) that DataSync can work with.

### Topics

- [SAS tokens](#)
- [Access tiers](#)

### SAS tokens

A SAS token specifies the access permissions for your blob storage. (For more information about SAS, see the [Azure Blob Storage documentation](#).)

You can generate SAS tokens to provide different levels of access. DataSync supports tokens with the following access levels:

- Account
- Container

The access permissions that DataSync needs depends on the scope of your token. Not having the correct permissions can cause your transfer to fail. For example, your transfer won't succeed if you're moving objects with tags to Azure Blob Storage but your SAS token doesn't have tag permissions.

### Topics

- [SAS token permissions for account-level access](#)

- [SAS token permissions for container-level access](#)
- [SAS expiration policies](#)

## SAS token permissions for account-level access

DataSync needs an account-level access token with the following permissions (depending on whether you're transferring to or from Azure Blob Storage).

### Transfers from blob storage

- **Allowed services** – Blob
- **Allowed resource types** – Container, Object

If you don't include these permissions, DataSync can't transfer your object metadata, including [object tags](#).

- **Allowed permissions** – Read, List
- **Allowed blob index permissions** – Read/Write (if you want DataSync to copy [object tags](#))

### Transfers to blob storage

- **Allowed services** – Blob
- **Allowed resource types** – Container, Object

If you don't include these permissions, DataSync can't transfer your object metadata, including [object tags](#).

- **Allowed permissions** – Read, Write, List, Delete (if you want DataSync to remove files that aren't in your transfer source)
- **Allowed blob index permissions** – Read/Write (if you want DataSync to copy [object tags](#))

## SAS token permissions for container-level access

DataSync needs a container-level access token with the following permissions (depending on whether you're transferring to or from Azure Blob Storage).

### Transfers from blob storage

- Read


- List
- Tag (if you want DataSync to copy [object tags](#))

 **Note**

You can't add the tag permission when generating a SAS token in the Azure portal. To add the tag permission, instead generate the token by using the [Azure Storage Explorer](#) app or generate a [SAS token that provides account-level access](#).

## Transfers to blob storage

- Read
- Write
- List
- Delete (if you want DataSync to remove files that aren't in your transfer source)
- Tag (if you want DataSync to copy [object tags](#))

 **Note**

You can't add the tag permission when generating a SAS token in the Azure portal. To add the tag permission, instead generate the token by using the [Azure Storage Explorer](#) app or generate a [SAS token that provides account-level access](#).

## SAS expiration policies

Make sure that your SAS doesn't expire before you expect to finish your transfer. For information about configuring a SAS expiration policy, see the [Azure Blob Storage documentation](#).

If the SAS expires during the transfer, DataSync can no longer access your Azure Blob Storage location. (You might see a Failed to open directory error.) If this happens, [update your location](#) with a new SAS token and restart your DataSync task.

## Access tiers

When transferring from Azure Blob Storage, DataSync can copy objects in the hot and cool tiers. For objects in the archive access tier, you must rehydrate those objects to the hot or cool tier before you can copy them.

When transferring to Azure Blob Storage, DataSync can copy objects into the hot, cool, and archive access tiers. If you're copying objects into the archive access tier, DataSync can't verify the transfer if you're trying to [verify all data in the destination](#).

DataSync doesn't support the cold access tier. For more information about access tiers, see the [Azure Blob Storage documentation](#).

## Considerations with Azure Blob Storage transfers

When planning to move data to or from Azure Blob Storage with DataSync, there are some things to keep in mind.

### Topics

- [Costs](#)
- [Blob types](#)
- [AWS Region availability](#)
- [Copying object tags](#)
- [Transferring to Amazon S3](#)
- [Deleting directories in a transfer destination](#)
- [Limitations](#)

### Costs

The fees associated with moving data in or out of Azure Blob Storage can include:

- Running an [Azure virtual machine \(VM\)](#) (if you deploy your DataSync agent in Azure)
- Running an [Amazon EC2](#) instance (if you deploy your DataSync agent in a VPC within AWS)
- Transferring the data by using [DataSync](#), including request charges related to [Azure Blob Storage](#) and [Amazon S3](#) (if S3 is one of your transfer locations)
- Transferring data in or out of [Azure Blob Storage](#)



- Storing data in an [AWS storage service](#) supported by DataSync

## Blob types

How DataSync works with blob types depends on whether you're transferring to or from Azure Blob Storage. When you're moving data into blob storage, the objects or files that DataSync transfers can only be block blobs. When you're moving data out of blob storage, DataSync can transfer block, page, and append blobs.

For more information about blob types, see the [Azure Blob Storage documentation](#).

## AWS Region availability

You can create an Azure Blob Storage transfer location in any [AWS Region that's supported by DataSync](#).

## Copying object tags

The ability for DataSync to preserve object tags when transferring to or from Azure Blob Storage depends on the following factors:

- **The size of an object's tags** – DataSync can't transfer an object with tags that exceed 2 KB.
- **Whether DataSync is configured to copy object tags** – DataSync [copies object tags](#) by default.
- **The namespace that your Azure storage account uses** – DataSync can copy object tags if your Azure storage account uses a flat namespace but not if your account uses a hierarchical namespace (a feature of Azure Data Lake Storage Gen2). Your DataSync task will fail if you try to copy object tags and your storage account uses a hierarchical namespace.
- **Whether your SAS token authorizes tagging** – The permissions that you need to copy object tags vary depending on the level of access that your token provides. Your task will fail if you try to copy object tags and your token doesn't have the right permissions for tagging. For more information, check the permission requirements for [account-level access tokens](#) or [container-level access tokens](#).

## Transferring to Amazon S3

When transferring to Amazon S3, DataSync won't transfer Azure Blob Storage objects larger than 5 TB or objects with metadata larger than 2 KB.

## Deleting directories in a transfer destination

When transferring to Azure Blob Storage, DataSync can [remove objects in your blob storage that aren't present in your transfer source](#). (You can configure this option by clearing the **Keep deleted files** setting in the DataSync console. Your [SAS token](#) must also have delete permissions.)

When you configure your transfer this way, DataSync won't delete directories in your blob storage if your Azure storage account is using a hierarchical namespace. In this case, you must manually delete the directories (for example, by using [Azure Storage Explorer](#)).

## Limitations

Remember the following limitations when transferring data to or from Azure Blob Storage:

- DataSync [creates some directories](#) in a location to help facilitate your transfer. If Azure Blob Storage is a destination location and your storage account uses a hierarchical namespace, you might notice task-specific subdirectories (such as `task-000011112222abcde`) in the `/.aws-datasync` folder. DataSync typically deletes these subdirectories following a transfer. If that doesn't happen, you can delete these task-specific directories yourself as long as a task isn't running.
- DataSync doesn't support using a SAS token to access only a specific folder in your Azure Blob Storage container.
- You can't provide DataSync a user delegation SAS token for accessing your blob storage.

## Creating your DataSync agent

To get started, you must create a DataSync agent that can connect to your Azure Blob Storage container. This process includes deploying and activating an agent.

### Tip

Although you can deploy your agent on an Amazon EC2 instance, using a Microsoft Hyper-V agent might result in decreased network latency and more data compression.

## Microsoft Hyper-V agents

You can deploy your DataSync agent directly in Azure with a Microsoft Hyper-V image.

**Tip**

Before you continue, consider using a shell script that might help you deploy your Hyper-V agent in Azure quicker. You can get more information and download the code on [GitHub](#). If you use the script, you can skip ahead to the section about [Getting your agent's activation key](#).

**Topics**

- [Prerequisites](#)
- [Downloading and preparing your agent](#)
- [Deploying your agent in Azure](#)
- [Getting your agent's activation key](#)
- [Activating your agent](#)

**Prerequisites**

To prepare your DataSync agent and deploy it in Azure, you must do the following:

- Enable Hyper-V on your local machine.
- Install [PowerShell](#) (including the Hyper-V Module).
- Install the [Azure CLI](#).
- Install [AzCopy](#).

**Downloading and preparing your agent**

Download an agent from the DataSync console. Before you can deploy the agent in Azure, you must convert it to a fixed-size virtual hard disk (VHD). For more information, see the [Azure documentation](#).

**To download and prepare your agent**

1. Open the AWS DataSync console at <https://console.aws.amazon.com/datasync/>.
2. In the left navigation pane, choose **Agents**, and then choose **Create agent**.
3. For **Hypervisor**, choose **Microsoft Hyper-V**, and then choose **Download the image**.

The agent downloads in a .zip file that contains a .vhdx file.

4. Extract the .vhdx file on your local machine.
5. Open PowerShell and do the following:
  - a. Copy the following Convert-VHD cmdlet:

```
Convert-VHD -Path .\local-path-to-vhdx-file\aws-datasync-2.0.1686143940.1-  
x86_64.xfs.gpt.vhdx `  
-DestinationPath .\local-path-to-vhdx-file\aws-datasync-2016861439401-  
x86_64.vhd -VHDType Fixed
```

- b. Replace each instance of *local-path-to-vhdx-file* with the location of the .vhdx file on your local machine.
  - c. Run the command.

Your agent is now a fixed-size VHD (with a .vhd file format) and ready to deploy in Azure.

## Deploying your agent in Azure

Deploying your DataSync agent in Azure involves:

- Creating a managed disk in Azure
- Uploading your agent to that managed disk
- Attaching the managed disk to a Linux virtual machine

## To deploy your agent in Azure

1. In PowerShell, go to the directory that contains your agent's .vhd file.
2. Run the `ls` command and save the Length value (for example, 85899346432).

This is the size of your agent image in bytes, which you need when creating a managed disk that can hold the image.

3. Do the following to create a managed disk:
  - a. Copy the following Azure CLI command:

```
az disk create -n your-managed-disk `
```

```
-g your-resource-group \  
-l your-azure-region \  
--upload-type Upload \  
--upload-size-bytes agent-size-bytes \  
--sku standard_lrs
```

- b. Replace *your-managed-disk* with a name for your managed disk.
- c. Replace *your-resource-group* with the name of the Azure resource group that your storage account belongs to.
- d. Replace *your-azure-region* with the Azure region where your resource group is located.
- e. Replace *agent-size-bytes* with the size of your agent image.
- f. Run the command.

This command creates an empty managed disk with a [standard SKU](#) where you can upload your DataSync agent.

4. To generate a shared access signature (SAS) that allows write access to the managed disk, do the following:
  - a. Copy the following Azure CLI command:

```
az disk grant-access -n your-managed-disk \  
-g your-resource-group \  
--access-level Write \  
--duration-in-seconds 86400
```

- b. Replace *your-managed-disk* with the name of the managed disk that you created.
- c. Replace *your-resource-group* with the name of the Azure resource group that your storage account belongs to.
- d. Run the command.

In the output, take note of the SAS URI. You need this URI when uploading the agent to Azure.

The SAS allows you to write to the disk for up to an hour. This means that you have an hour to upload your agent to the managed disk.

5. To upload your agent to your managed disk in Azure, do the following:

- a. Copy the following AzCopy command:

```
.\azcopy copy local-path-to-vhd-file sas-uri --blob-type PageBlob
```

- b. Replace *local-path-to-vhd-file* with the location of the agent's .vhd file on your local machine.
  - c. Replace *sas-uri* with the SAS URI that you got when you ran the `az disk grant-access` command.
  - d. Run the command.
6. When the agent upload finishes, revoke access to your managed disk. To do this, copy the following Azure CLI command:

```
az disk revoke-access -n your-managed-disk -g your-resource-group
```

- a. Replace *your-resource-group* with the name of the Azure resource group that your storage account belongs to.
  - b. Replace *your-managed-disk* with the name of the managed disk that you created.
  - c. Run the command.
7. Do the following to attach your managed disk to a new Linux VM:

- a. Copy the following Azure CLI command:

```
az vm create --resource-group your-resource-group `
--location eastus `
--name your-agent-vm `
--size Standard_E4as_v4 `
--os-type linux `
--attach-os-disk your-managed-disk
```

- b. Replace *your-resource-group* with the name of the Azure resource group that your storage account belongs to.
- c. Replace *your-agent-vm* with a name for the VM that you can remember.
- d. Replace *your-managed-disk* with the name of the managed disk that you're attaching to the VM.
- e. Run the command.

You've deployed your agent. Before you can start configuring your data transfer, you must activate the agent.

## Getting your agent's activation key

To manually get your DataSync agent's activation key, follow these steps.

Alternatively, [DataSync can automatically get the activation key for you](#), but this approach requires some network configuration.

### To get your agent's activation key

1. In the Azure portal, [enable boot diagnostics for the VM for your agent](#) by choosing the **Enable with custom storage account** setting and specifying your Azure storage account.

After you've enabled the boot diagnostics for your agent's VM, you can access your agent's local console to get the activation key.

2. While still in the Azure portal, go to your VM and choose **Serial console**.
3. In the agent's local console, log in by using the following default credentials:

- **Username – admin**
- **Password – password**

We recommend at some point changing at least the agent's password. In the agent's local console, enter **5** on the main menu, then use the `passwd` command to change the password.

4. Enter **0** to get the agent's activation key.
5. Enter the AWS Region where you're using DataSync (for example, **us-east-1**).
6. Choose the [service endpoint](#) that the agent will use to connect with AWS.
7. Save the value of the `Activation` key output.

## Activating your agent

After you have the activation key, you can finish creating your DataSync agent.

### To activate your agent

1. Open the AWS DataSync console at <https://console.aws.amazon.com/datasync/>.
2. In the left navigation pane, choose **Agents**, and then choose **Create agent**.

3. For **Hypervisor**, choose **Microsoft Hyper-V**.
4. For **Endpoint type**, choose the same type of service endpoint that you specified when you got your agent's activation key (for example, choose **Public service endpoints in *Region name***).
5. Configure your network to work with the service endpoint type that your agent is using. For service endpoint network requirements, see the following topics:
  - [VPC endpoints](#)
  - [Public endpoints](#)
  - [Federal Information Processing Standard \(FIPS\) endpoints](#)
6. For **Activation key**, do the following:
  - a. Choose **Manually enter your agent's activation key**.
  - b. Enter the activation key that you got from the agent's local console.
7. Choose **Create agent**.

Your agent is ready to connect with your Azure Blob Storage. For more information, see [Creating your Azure Blob Storage transfer location](#).

## Amazon EC2 agents

You can deploy your DataSync agent on an Amazon EC2 instance.

### To create an Amazon EC2 agent

1. [Deploy an Amazon EC2 agent](#).
2. [Choose a service endpoint](#) that the agent uses to communicate with AWS.

In this situation, we recommend using a virtual private cloud (VPC) service endpoint.

3. Configure your network to work with [VPC service endpoints](#).
4. [Activate the agent](#).

## Creating your Azure Blob Storage transfer location

You can configure DataSync to use your Azure Blob Storage as a transfer source or destination.

### Before you begin



Make sure that you know [how DataSync accesses Azure Blob Storage](#) and works with [access tiers](#) and [blob types](#). You also need a [DataSync agent](#) that can connect to your Azure Blob Storage container.

## Using the DataSync console

1. Open the AWS DataSync console at <https://console.aws.amazon.com/datasync/>.
2. In the left navigation pane, expand **Data transfer**, then choose **Locations** and **Create location**.
3. For **Location type**, choose **Microsoft Azure Blob Storage**.
4. For **Agents**, choose the DataSync agent that can connect with your Azure Blob Storage container.

You can choose more than one agent. For more information, see [Using multiple DataSync agents](#).

5. For **Container URL**, enter the URL of the container that's involved in your transfer.
6. (Optional) For **Access tier when used as a destination**, choose the [access tier](#) that you want your objects or files transferred into.
7. For **Folder**, enter path segments if you want to limit your transfer to a virtual directory in your container (for example, `/my/images`).
8. For **SAS token**, enter the SAS token that allows DataSync to access your blob storage.

The token is part of the SAS URI string that comes after the storage resource URI and a question mark (?). A token looks something like this:

```
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%2BZU178wccwqU%3D
```

9. (Optional) Enter values for the **Key** and **Value** fields to tag the location.

Tags help you manage, filter, and search for your AWS resources. We recommend creating at least a name tag for your location.

10. Choose **Create location**.

## Using the AWS CLI

1. Copy the following `create-location-azure-blob` command:

```
aws datasync 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" \
  --access-tier "access-tier-for-destination" \
  --tags [{"Key": "key1", "Value": "value1"}]
```

2. For the `--container-url` parameter, specify the URL of the Azure Blob Storage container that's involved in your transfer.
3. For the `--authentication-type` parameter, specify SAS.
4. For the `--sas-configuration` parameter's Token option, specify the SAS token that allows DataSync to access your blob storage.

The token is part of the SAS URI string that comes after the storage resource URI and a question mark (?). A token looks something like this:

```
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
```

5. For the `--agent-arns` parameter, specify the Amazon Resource Name (ARN) of the DataSync agent that can connect to your container.

Here's an example agent ARN: `arn:aws:datasync:us-east-1:123456789012:agent/agent-01234567890aaabfb`

You can specify more than one agent. For more information, see [Using multiple DataSync agents](#).

6. For the `--subdirectory` parameter, specify path segments if you want to limit your transfer to a virtual directory in your container (for example, `/my/images`).
7. (Optional) For the `--access-tier` parameter, specify the [access tier](#) (HOT, COOL, or ARCHIVE) that you want your objects or files transferred into.

This parameter applies only when you're using this location as a transfer destination.

8. (Optional) For the `--tags` parameter, specify key-value pairs that can help you manage, filter, and search for your location.

We recommend creating a name tag for your location.

9. Run the `create-location-azure-blob` command.

If the command is successful, you get a response that shows you the ARN of the location that you created. For example:

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

## Viewing your Azure Blob Storage transfer location

You can get details about the existing DataSync transfer location for your Azure Blob Storage.

### Using the DataSync console

1. Open the AWS DataSync console at <https://console.aws.amazon.com/datsync/>.
2. In the left navigation pane, expand **Data transfer**, then choose **Locations**.
3. Choose your Azure Blob Storage location.

You can see details about your location, including any DataSync transfer tasks that are using it.

### Using the AWS CLI

1. Copy the following `describe-location-azure-blob` command:

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

2. For the `--location-arn` parameter, specify the ARN for the Azure Blob Storage location that you created (for example, `arn:aws:datsync:us-east-1:123456789012:location/loc-12345678abcdefgh`).
3. Run the `describe-location-azure-blob` command.

You get a response that shows you details about your location. For example:

```
{
  "LocationArn": "arn:aws:datasync: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:datasync:us-east-1:123456789012:agent/
agent-01234567890deadfb"],
}
```

## Updating your Azure Blob Storage transfer location

If needed, you can modify your location's configuration in the console or by using the AWS CLI.

### Using the AWS CLI

1. Copy the following `update-location-azure-blob` command:

```
aws datasync 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" \
  --access-tier "access-tier-for-destination"
```

2. For the `--location-arn` parameter, specify the ARN for the Azure Blob Storage location that you're updating (for example, `arn:aws:datasync:us-east-1:123456789012:location/loc-12345678abcdefgh`).
3. For the `--authentication-type` parameter, specify SAS.
4. For the `--sas-configuration` parameter's `Token` option, specify the SAS token that allows DataSync to access your blob storage.

The token is part of the SAS URI string that comes after the storage resource URI and a question mark (?). A token looks something like this:

```
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%2BZU178wqwU%3D
```

5. For the `--agent-arns` parameter, specify the Amazon Resource Name (ARN) of the DataSync agent that you want to connect to your container.

Here's an example agent ARN: `arn:aws:datsync:us-east-1:123456789012:agent/agent-01234567890aaabfb`

You can specify more than one agent. For more information, see [Using multiple DataSync agents](#).

6. For the `--subdirectory` parameter, specify path segments if you want to limit your transfer to a virtual directory in your container (for example, `/my/images`).
7. (Optional) For the `--access-tier` parameter, specify the [access tier](#) (HOT, COOL, or ARCHIVE) that you want your objects to be transferred into.

This parameter applies only when you're using this location as a transfer destination.

## Next steps

After you finish creating a DataSync location for your Azure Blob Storage, you can continue setting up your transfer. Here are some next steps to consider:

1. If you haven't already, [create another location](#) where you plan to transfer your data to or from your Azure Blob Storage.
2. Learn how DataSync [handles metadata and special files](#), particularly if your transfer locations don't have a similar metadata structure.
3. Configure how your data gets transferred. For example, you can [move only a subset of your data](#) or delete files in your blob storage that aren't in your source location (as long as your [SAS token](#) has delete permissions).
4. [Start your transfer](#).

# Configuring AWS DataSync transfers with Microsoft Azure Files SMB shares

You can configure AWS DataSync to transfer data to or from a Microsoft Azure Files Server Message Block (SMB) share.

## Tip

For a full walkthrough on moving data from Azure Files SMB shares to AWS, see the [AWS Storage Blog](#).

## Accessing SMB shares

DataSync connects to your SMB share using the SMB protocol and authenticates with credentials that you provide it.

### Topics

- [Supported SMB protocol versions](#)
- [Required permissions](#)

## Supported SMB protocol versions

By default, DataSync automatically chooses a version of the SMB protocol based on negotiation with your SMB file server.

You also can configure DataSync to use a specific SMB version, but we recommend doing this only if DataSync has trouble negotiating with the SMB file server automatically. (DataSync supports SMB versions 1.0 and later.)

See the following table for a list of options in the DataSync console and API:

Console option	API option	Description
Automatic	AUTOMATIC	DataSync and the SMB file server negotiate the highest version of SMB that they mutually support between 2.1 and 3.1.1.

Console option	API option	Description
		This is the default and recommended option. If you instead choose a specific version that your file server doesn't support, you may get an <code>Operation Not Supported</code> error.
SMB 3.0.2	SMB3	Restricts the protocol negotiation to only SMB version 3.0.2.
SMB 2.1	SMB2	Restricts the protocol negotiation to only SMB version 2.1.
SMB 2.0	SMB2_0	Restricts the protocol negotiation to only SMB version 2.0.
SMB 1.0	SMB1	Restricts the protocol negotiation to only SMB version 1.0.

## Required permissions

DataSync needs a user who has permission to mount and access your SMB location. This can be a local user on your Windows file server or a domain user that's defined in your Microsoft Active Directory.

To set object ownership, DataSync requires the `SE_RESTORE_NAME` privilege, which is usually granted to members of the built-in Active Directory groups **Backup Operators** and **Domain Admins**. Providing a user to DataSync with this privilege also helps ensure sufficient permissions to files, folders, and file metadata, except for NTFS system access control lists (SACLs).

Additional privileges are required to copy SACLs. Specifically, this requires the Windows `SE_SECURITY_NAME` privilege, which is granted to members of the **Domain Admins** group. If you configure your task to copy SACLs, make sure that the user has the required privileges. To learn more about configuring a task to copy SACLs, see [Configuring how AWS DataSync handles files, objects, and metadata](#).

When you copy data between an SMB file server and Amazon FSx for Windows File Server file system, the source and destination locations must belong to the same Microsoft Active Directory domain or have an Active Directory trust relationship between their domains.

## Creating your Azure Files transfer location by using the console

1. Open the AWS DataSync console at <https://console.aws.amazon.com/datasync/>.
2. In the left navigation pane, expand **Data transfer**, then choose **Locations** and **Create location**.
3. For **Location type**, choose **Server Message Block (SMB)**.

You configure this location as a source or destination later.

4. For **Agents**, choose one or more DataSync agents that you want to connect to your SMB share.

If you choose more than one agent, make sure you understand using [multiple agents for a location](#).

5. For **SMB Server**, enter the Domain Name System (DNS) name or IP address of the SMB share that your DataSync agent will mount.

### Note

You can't specify an IP version 6 (IPv6) address.

6. For **Share name**, enter the name of the share exported by your SMB share where DataSync will read or write data.

You can include a subdirectory in the share path (for example, `/path/to/subdirectory`). Make sure that other SMB clients in your network can also mount this path.

To copy all the data in the subdirectory, DataSync must be able to mount the SMB share and access all of its data. For more information, see [Required permissions](#).

7. (Optional) Expand **Additional settings** and choose an **SMB Version** for DataSync to use when accessing your SMB share.

By default, DataSync automatically chooses a version based on negotiation with the SMB share. For information, see [Supported SMB versions](#).

8. For **User**, enter a user name that can mount your SMB share and has permission to access the files and folders involved in your transfer.

For more information, see [Required permissions](#).

9. For **Password**, enter the password of the user who can mount your SMB share and has permission to access the files and folders involved in your transfer.



10. (Optional) For **Domain**, enter the Windows domain name that your SMB share belongs to.

If you have multiple domains in your environment, configuring this setting makes sure that DataSync connects to the right share.

11. (Optional) Choose **Add tag** to tag your location.

*Tags* are key-value pairs that help you manage, filter, and search for your locations. We recommend creating at least a name tag for your location.

12. Choose **Create location**.

## Configuring AWS DataSync transfers with other cloud object storage

With AWS DataSync, you can transfer data between [AWS storage services](#) and the following cloud object storage providers:

- [Wasabi Cloud Storage](#)
- [DigitalOcean Spaces](#)
- [Oracle Cloud Infrastructure Object Storage](#)
- [Cloudflare R2 Storage](#)
- [Backblaze B2 Cloud Storage](#)
- [NAVER Cloud Object Storage](#)
- [Alibaba Cloud Object Storage Service](#)
- [IBM Cloud Object Storage](#)
- [Seagate Lyve Cloud](#)

To set up this kind of transfer, you need to create a DataSync [agent](#) that can connect to your cloud object storage. You must also create a transfer [location](#) for your cloud object storage (specifically an **Object storage** location). DataSync can use this location as a source or destination for your transfer.

### Accessing other cloud object storage

How DataSync accesses your cloud object storage depends on several factors, including whether your storage is compatible with the Amazon S3 API and the permissions and credentials that DataSync needs to access your storage.

## Topics

- [Amazon S3 API compatibility](#)
- [Storage permissions and endpoints](#)
- [Storage credentials](#)

### Amazon S3 API compatibility

Your cloud object storage must be compatible with the following [Amazon S3 API operations](#) for DataSync to connect to it:

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

### Storage permissions and endpoints

You must configure the permissions that allow DataSync to access your cloud object storage. If your object storage is a source location, DataSync needs read and list permissions for the bucket that you're transferring data from. If your object storage is a destination location, DataSync needs read, list, write, and delete permissions for the bucket.

DataSync also needs an endpoint (or server) to connect to your storage. The following table describes the endpoints that DataSync can use to access other cloud object storage:

Other cloud provider	Endpoint
Wasabi Cloud Storage	S3. <i>region</i> .wasabisys.com
DigitalOcean Spaces	<i>region</i> .digitaloceanspaces.com
Oracle Cloud Infrastructure Object Storage	<i>namespace</i> .compat.objectstorage. <i>region</i> .oraclecloud.com
Cloudflare R2 Storage	<i>account-id</i> .r2.cloudflarestorage.com
Backblaze B2 Cloud Storage	S3. <i>region</i> .backblazeb2.com
NAVER Cloud Object Storage	<i>region</i> .object.ncloudstorage.com (most regions)
Alibaba Cloud Object Storage Service	<i>region</i> .aliyuncs.com
IBM Cloud Object Storage	s3. <i>region</i> .cloud-object-storage.appdomain.cloud
Seagate Lyve Cloud	s3. <i>region</i> .lyvecloud.seagate.com

### Important

For details on how to configure bucket permissions and updated information on storage endpoints, see your cloud provider's documentation.

## Storage credentials

DataSync also needs the credentials to access the object storage bucket involved in your transfer. This might be an access key and secret key or something similar depending on how your cloud storage provider refers to these credentials.

For more information, see your cloud provider's documentation.

## Considerations when transferring from other cloud object storage

When planning to move objects to or from another cloud storage provider by using DataSync, there are some things to keep in mind.

### Topics

- [Costs](#)
- [Storage classes](#)
- [Object tags](#)
- [Transferring to Amazon S3](#)

### Costs

The fees associated with moving data in and out of another cloud storage provider can include:

- Running an [Amazon EC2](#) instance for your DataSync agent
- Transferring the data by using [DataSync](#), including request charges related to your cloud object storage and [Amazon S3](#) (if S3 is your transfer destination)
- Transferring data in or out of your cloud storage (check your cloud provider's pricing)
- Storing data in an [AWS storage service](#) supported by DataSync
- Storing data in another cloud provider (check your cloud provider's pricing)

### Storage classes

Some cloud storage providers have storage classes (similar to [Amazon S3](#)) which DataSync can't read without being restored first. For example, Oracle Cloud Infrastructure Object Storage has an archive storage class. You need to restore objects in that storage class before DataSync can transfer them. For more information, see your cloud provider's documentation.

### Object tags

Not all cloud providers support object tags. The ones that do might not allow querying tags through the Amazon S3 API. In either situation, your DataSync transfer task might fail if you try to copy object tags.

You can avoid this by clearing the **Copy object tags** checkbox in the DataSync console when creating, starting, or updating your task.

## Transferring to Amazon S3

When transferring to Amazon S3, DataSync can't transfer objects larger than 5 TB. DataSync also can only copy object metadata up to 2 KB.

## Creating your DataSync agent

To get started, you need a DataSync agent that can connect to your cloud object storage. This process includes deploying and activating an agent on an Amazon EC2 instance in your virtual private cloud (VPC) in AWS.

### To create an Amazon EC2 agent

1. [Deploy an Amazon EC2 agent](#).
2. [Choose a service endpoint](#) that the agent uses to communicate with AWS.

In this situation, we recommend using a VPC service endpoint.

3. Configure your network to work with [VPC service endpoints](#).
4. [Activate the agent](#).

## Creating a transfer location for your other cloud object storage

You can configure DataSync to use your cloud object storage as a source or destination location.

### Before you begin

Make sure that you know [how DataSync accesses your cloud object storage](#). You also need a [DataSync agent](#) that can connect to your cloud object storage.

1. Open the AWS DataSync console at <https://console.aws.amazon.com/datasync/>.
2. In the left navigation pane, expand **Data transfer**, then choose **Locations** and **Create location**.
3. For **Location type**, choose **Object storage**.
4. For **Agents**, choose the DataSync agent that can connect with your cloud object storage.

You can choose more than one agent. For more information, see [Using multiple DataSync agents](#).

5. For **Server**, enter the [endpoint](#) that DataSync can use to access your cloud object storage:
  - **Wasabi Cloud Storage** – `S3.region.wasabisys.com`
  - **DigitalOcean Spaces** – `region.digitaloceanspaces.com`
  - **Oracle Cloud Infrastructure Object Storage** – `namespace.compat.objectstorage.region.oraclecloud.com`
  - **Cloudflare R2 Storage** – `account-id.r2.cloudflarestorage.com`
  - **Backblaze B2 Cloud Storage** – `S3.region.backblazeb2.com`
  - **NAVER Cloud Object Storage** – `region.object.ncloudstorage.com` (most regions)
  - **Alibaba Cloud Object Storage Service** – `region.aliyuncs.com`
  - **IBM Cloud Object Storage** – `s3.region.cloud-object-storage.appdomain.cloud`
  - **Seagate Lyve Cloud** – `s3.region.lyvecloud.seagate.com`
6. For **Bucket name**, enter the name of the object storage bucket that you're transferring data to or from.
7. Expand **Additional settings**. For **Server protocol**, choose **HTTPS**. For **Server port**, choose **443**.
8. Scroll down to the **Authentication** section. Make sure that the **Requires credentials** check box is selected, and then provide DataSync your [storage credentials](#).
  - For **Access key**, enter the ID to access your cloud object storage.
  - For **Secret key**, enter the secret to access your cloud object storage.
9. (Optional) Enter values for the **Key** and **Value** fields to tag the location.

Tags help you manage, filter, and search for your AWS resources. We recommend creating at least a name tag for your location.

10. Choose **Create location**.

## Next steps

After you finish creating a DataSync location for your cloud object storage, you can continue setting up your transfer. Here are some next steps to consider:

1. If you haven't already, [create another location](#) where you plan to transfer your data to or from in AWS.
2. [Learn how DataSync handles metadata and special files](#) for object storage locations.

3. Configure how your data gets transferred. For example, maybe you only want to [move a subset of your data](#).

**⚠ Important**

Make sure that you configure how DataSync copies object tags correctly. For more information, see considerations with [object tags](#).

4. [Start your transfer](#).

## Transferring to or from edge storage with AWS DataSync

With AWS DataSync, you can transfer data to or from some AWS Snow Family devices. For more information, see [Where can I transfer my data with DataSync?](#).

### Topics

- [Configuring transfers with S3 compatible storage on Snowball Edge](#)
- [Configuring AWS DataSync transfers with AWS Snowcone](#)

## Configuring transfers with S3 compatible storage on Snowball Edge

With AWS DataSync, you can transfer objects between [Amazon S3 compatible storage on an AWS Snowball Edge](#) device or cluster and any of the following AWS storage services:

- [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 ONTAP](#)

### Prerequisites

Before you get started, make sure that you've done the following:

- Created an AWS storage resource in the AWS Region where you plan to transfer data to or from. For example, this could be an S3 bucket or Amazon EFS file system in US East (N. Virginia).
- Established a wide-area network (WAN) connection for traffic into and out of your on-premises storage environment. For example, you can establish this kind of connection with [AWS Direct Connect](#).

When you [create your DataSync agent](#), you'll configure this WAN connection so that DataSync can transfer data between your Amazon S3 compatible storage that's on-premises and your storage resource in AWS.

- Downloaded and installed the [Snowball Edge client](#).

## Accessing your Amazon S3 compatible storage

To access your Amazon S3 compatible storage bucket, DataSync needs the following:

- User credentials on your Snowball Edge device or cluster that can access the bucket that you're transferring data to or from.
- An HTTPS certificate that allows DataSync to verify the authenticity of the connection between the DataSync agent and the `s3api` endpoint on your device or cluster.

### Topics

- [Getting the user credentials to access your S3 bucket](#)
- [Getting a certificate for the s3api endpoint connection](#)

## Getting the user credentials to access your S3 bucket

DataSync needs the access key and secret key for a user who can access the bucket that you're working with on your Snowball Edge device or cluster.

### To get the user credentials to access your bucket

1. Open a terminal and run the Snowball Edge client.

For more information about running the Snowball Edge client, see [Using the Snowball Edge client](#) in the *AWS Snowball Edge Developer Guide*.

2. To get the access keys associated with your device or cluster, run the following `snowballEdge` command:



```
snowballEdge list-access-keys
```

3. In the output, locate the access key for the bucket that DataSync will work with (for example, AKIAIOSFODNN7EXAMPLE).
4. To get the secret access key, run the following snowballEdge command. Replace *access-key-for-datasync* with the access key that you located in the prior step.

```
snowballEdge get-secret-access-key --access-key-id access-key-for-datasync
```

The output includes the access key's corresponding secret key (for example, wJalrXUtnFEMI/K7MDENG/bPxrFiCYEXAMPLEKEY).

5. Save the access key and secret key somewhere that you can remember.

You will need these keys when you're [configuring the DataSync source location](#) for your transfer.

## Getting a certificate for the s3api endpoint connection

You need an HTTPS certificate that can verify the authenticity of the connection between your DataSync agent and an s3api endpoint on your Snowball Edge device or cluster.

### To get a certificate for the s3api endpoint connection

1. In the Snowball Edge client, run the following snowballEdge command:

```
snowballEdge get-certificate
```

2. Save the output to a base64-encoded .pem file.

You will specify this file when you're [configuring the DataSync source location](#) for your transfer.

## Creating a DataSync agent in your on-premises storage environment

During a transfer, DataSync uses an [agent](#) to read from or write to the Amazon S3 compatible storage on your Snowball Edge device or cluster.

This agent must be deployed in your on-premises storage environment where it can connect to your device or cluster through your network. For example, you can run the agent on a VMware ESXi hypervisor that has local network access to your cluster.

### To create a DataSync agent in your on-premises storage environment

1. Make sure that the [DataSync agent can run on your hypervisor](#) and that you [allocate the agent enough virtual machine \(VM\) resources](#).
2. Deploy the agent in your on-premises environment.

For instructions, see one of the following topics, depending on the type of hypervisor that you're deploying the agent on:

- [Deploy your agent on VMware](#)
- [Deploy your agent on Linux Kernel-based Machine \(KVM\)](#)
- [Deploy your agent on Microsoft Hyper-V](#)
- [Deploy your agent on Amazon EC2](#)

#### **Warning**

We don't recommend deploying an agent on Amazon EC2 agent to access on-premises storage because of increased network latency.

3. Configure your network to allow the following traffic between the agent and your Amazon S3 compatible storage:

From	To	Protocol and port
DataSync agent	A virtual network interface (VNI) for an s3api endpoint on your device or cluster. If you have a cluster, it can be any s3api endpoint VNI.	TCP 443 (HTTPS)

If you need to find a VNI on your device or cluster, see [describing your virtual network interfaces](#) on Snowball Edge.

4. [Choose a service endpoint](#) that the agent will use to communicate with AWS.
5. [Activate your agent](#).

## Configuring the source location for your transfer

After you create your agent, you can configure the source location for your DataSync transfer.

### Note

The following instructions assume that you're transferring from Amazon S3 compatible storage, but you can also use this location for a transfer destination.

### To configure the source location for your transfer by using the DataSync console

1. Open the AWS DataSync console at <https://console.aws.amazon.com/datasync/>.
2. In the left navigation pane, expand **Data transfer**. Choose **Tasks**, and then choose **Create task**.
3. On the **Configure source location** page, choose **Create a new location**.
4. For **Location type**, choose **Object storage**.
5. For **Agents**, choose the DataSync agent that you created in your on-premises storage environment.
6. For **Server**, enter the VNI for the `s3api` endpoint that's used by your Amazon S3 compatible storage.

If you have a Snowball Edge cluster instead of a single device, you can specify any of the cluster's `s3api` endpoint VNIs.

7. For **Bucket name**, enter the name of the Amazon S3 compatible storage bucket that you're transferring objects from.
8. For **Folder**, enter an object prefix.

DataSync only transfers objects with this prefix.

9. To configure the DataSync connection to the Snowball Edge device or cluster, expand **Additional settings** and do the following:
  - a. For **Server protocol**, choose **HTTPS**.
  - b. For **Server port**, enter **443**.

- c. For **Certificate**, choose the certificate file for the [s3api endpoint connection](#).
10. Select **Requires credentials**, and enter the **Access key** and **Secret key** to [access the Amazon S3 compatible storage bucket](#) on your Snowball Edge device or cluster.
11. Choose **Next**.

## Configuring the destination location for your transfer

Your transfer's destination location must be in the same AWS Region and AWS account where you created your agent.

**Before you begin:** Make sure you've [configured the source location](#) for your transfer.

### To configure the destination location for your transfer by using the DataSync console

1. On the **Configure destination location** page, choose **Create a new location** or **Choose an existing location** for the AWS storage resource where you're transferring objects to.

If you're creating a new location, see one of the following topics:

- [Amazon S3](#)
- [Amazon EFS](#)
- [FSx for Windows File Server](#)
- [FSx for Lustre](#)
- [FSx for OpenZFS](#)
- [FSx for ONTAP](#)

2. When you're done configuring the destination location, choose **Next**.

## Configuring your transfer settings

With DataSync, you can specify a transfer schedule, customize how your data integrity is verified, and specify whether you want to transfer only a subset of objects, among other options.

**Before you begin:** Make sure you've [configured the destination location](#) for your transfer.

### To configure your transfer settings by using the DataSync console

1. On the **Configure settings** page, change the transfer settings or use the defaults.

For more information about these settings, see [Working with AWS DataSync transfer tasks](#).

2. Choose **Next**.
3. Review your transfer details, and then choose **Create task**.

## Starting your transfer

After you create your transfer task, you're ready to start moving data. For instructions on starting a task by using the DataSync console or AWS CLI, see [Starting your task](#).

## Configuring AWS DataSync transfers with AWS Snowcone

To transfer data to or from an AWS Snowcone device, you must create an AWS DataSync transfer *location*. DataSync can use this location as a source or destination for transferring data.

### Creating your Snowcone transfer location

Before you begin, you need to enable Network File System (NFS) on your Snowcone device. For more information, see the [AWS Snowcone User Guide](#).

#### To create the location by using the console

1. Open the AWS DataSync console at <https://console.aws.amazon.com/datasync/>.
2. In the left navigation pane, expand **Data transfer**, then choose **Locations** and **Create location**.
3. For **Location type**, choose **Network File System (NFS)**.
4. For **Agents**, choose the Amazon EC2 agent that you launched on your Snowcone device.

For more information about the agent, see [AWS Snowcone User Guide](#).

5. For **NFS server**, enter the virtual IP address that you attached to the NFS server on your Snowcone device.
6. For **Mount path**, enter the NFS export path for the S3 bucket that you want to transfer data to or from.

The format of the export path for a bucket is `/buckets/bucket-name`. For more information, see [Using NFS file shares to manage file storage](#) in the *AWS Snowcone User Guide*.

7. (Optional) Expand **Additional settings** and choose a specific **NFS version** for DataSync to use when accessing your file server.

By default, DataSync uses NFS version 4.1. DataSync also supports NFS 3.x and 4.0.

8. (Optional) Choose **Add tag** to tag your location.

*Tags* are key-value pairs that help you manage, filter, and search for your locations. We recommend creating at least a name tag for your location.

9. Choose **Create location**.

## Configuring what AWS DataSync transfers

You can configure AWS DataSync to transfer what you want and how it handles the data it moves. Some examples include:

- Transferring an exact list of files or object by using a manifest.
- Including or excluding certain types of data in your transfer by using a filter.
- For recurring transfers, moving only the data that's changed since the last transfer
- Overwriting data in the destination location to match what's in the source location.
- Choosing which file or object metadata to preserve between your storage locations.
- Specifying what kind of integrity checks you want DataSync to perform on your transferred data.

### Topics

- [Specifying what AWS DataSync transfers by using a manifest](#)
- [Specifying what AWS DataSync transfers by using filters](#)
- [Metadata copied by AWS DataSync](#)
- [Links and directories copied by AWS DataSync](#)
- [Configuring how AWS DataSync handles files, objects, and metadata](#)
- [Configuring how AWS DataSync verifies data integrity](#)

## Specifying what AWS DataSync transfers by using a manifest

A *manifest* is a list of files or objects that you want AWS DataSync to transfer. For example, instead of having to transfer everything in an S3 bucket with potentially millions of objects, DataSync transfers only the objects that you list in your manifest.

Manifests are similar to [filters](#) but let you identify exactly which files or objects to transfer instead of data that matches a filter pattern.

## Creating your manifest

A manifest is a comma-separated values (CSV)-formatted file that lists the files or objects in your source location that you want DataSync to transfer. If your source is an S3 bucket, you can also include which version of an object to transfer.

### Topics

- [Guidelines](#)
- [Example manifests](#)

### Guidelines

Use these guidelines to help you create a manifest that works with DataSync.

#### Do

- Specify the full path of each file or object that you want to transfer.

You can't specify only a directory or folder with the intention of transferring all of its contents. For these situations, consider using an [include filter](#) instead of a manifest.

- Make sure that each file or object path is relative to the mount path, folder, directory, or prefix that you specified when configuring your DataSync source location.

For example, let's say you [configure an S3 location](#) with a prefix named photos. That prefix includes an object `my-picture.png` that you want to transfer. In the manifest, you then only need to specify the object (`my-picture.png`) instead of the prefix and object (`photos/my-picture.png`).

- To specify Amazon S3 object version IDs, separate the object's path and version ID by using a comma.

The following example shows a manifest entry with two fields. The first field includes an object named `picture1.png`. The second field is separated by a comma and includes a version ID of `111111`:

```
picture1.png,111111
```

- Use quotes in the following situations:
  - When a path contains special characters (commas, quotes, and line endings):

```
"filename,with,commas.txt"
```

- When a path spans multiple lines:

```
"this  
is  
a  
filename.txt"
```

- When a path includes quotes:

```
filename""with""quotes.txt
```

This represents a path named `filename"with"quotes.txt`.

These quote rules also apply to version ID fields. In general, if a manifest field has a quote, you must escape it with another quote.

- Separate each file or object entry with a new line.

You can separate lines by using Linux (line feed or carriage return) or Windows (carriage return followed by a line feed) style line breaks.

- Save your manifest (for example, `my-manifest.csv` or `my-manifest.txt`).
- Upload the manifest to an S3 bucket that [DataSync can access](#).

This bucket doesn't have to be in the same AWS Region or account where you're using DataSync.

## Don't

- Specify only a directory or folder with the intention of transferring all of its contents.

A manifest can only include full paths to the files or objects that you want to transfer. If you configure your source location to use a specific mount path, folder, directory, or prefix, you don't have to include that in your manifest.

- Specify a file or object path that exceeds 4,096 characters.
- Specify a file path, object path, or Amazon S3 object version ID that exceeds 1,024 bytes.



- Specify duplicate file or object paths.
- Include an object version ID if your source location isn't an S3 bucket.
- Include more than two fields in a manifest entry.

An entry can include only a file or object path and (if applicable) an Amazon S3 object version ID.

- Include characters that don't conform to UTF-8 encoding.
- Include unintentional spaces in your entry fields outside of quotes.

## Example manifests

Use these examples to help you create a manifest that works with DataSync.

### Manifest with full file or object paths

The following example shows a manifest with full file or object paths to transfer.

```
photos/picture1.png  
photos/picture2.png  
photos/picture3.png
```

### Manifest with only object keys

The following example shows a manifest with objects to transfer from an Amazon S3 source location. Since the [location is configured](#) with the prefix photos, only the object keys are specified.

```
picture1.png  
picture2.png  
picture3.png
```

### Manifest with object paths and version IDs

The first two entries in the following manifest example include specific Amazon S3 object versions to transfer.

```
photos/picture1.png,111111  
photos/picture2.png,121212
```

```
photos/picture3.png
```

## Manifest with UTF-8 characters

The following example shows a manifest with files that include UTF-8 characters.

```
documents/résumé1.pdf  
documents/résumé2.pdf  
documents/résumé3.pdf
```

## Providing DataSync access to your manifest

You need an AWS Identity and Access Management (IAM) role that gives DataSync access to your manifest in its S3 bucket. This role must include the following permissions:

- `s3:GetObject`
- `s3:GetObjectVersion`

You can generate this role automatically in the DataSync console or create the role yourself.

### Note

If your manifest is in a different AWS account, you must create this role manually.

## Creating the IAM role automatically

When creating or starting a transfer task in the console, DataSync can create an IAM role for you with the `s3:GetObject` and `s3:GetObjectVersion` permissions that you need to access your manifest.

## Required permissions to automatically create the role

To automatically create the role, make sure that the role that you're using to access the DataSync console has the following permissions:

- `iam:CreateRole`
- `iam:CreatePolicy`
- `iam:AttachRolePolicy`

## Creating the IAM role (same account)

You can manually create the IAM role that DataSync needs to access your manifest. The following instructions assume that you're in the same AWS account where you use DataSync and your manifest's S3 bucket is located.

1. Open the IAM console at <https://console.aws.amazon.com/iam/>.
2. In the left navigation pane, under **Access management**, choose **Roles**, and then choose **Create role**.
3. On the **Select trusted entity** page, for **Trusted entity type**, choose **AWS service**.
4. For **Use case**, choose **DataSync** in the dropdown list and select **DataSync**. Choose **Next**.
5. On the **Add permissions** page, choose **Next**. Give your role a name and choose **Create role**.
6. On the **Roles** page, search for the role that you just created and choose its name.
7. On the role's details page, choose the **Permissions** tab. Choose **Add permissions** then **Create inline policy**.
8. Choose the **JSON** tab and paste the following sample policy into the policy editor:

```
{
  "Version": "2012-10-17",
  "Statement": [{
    "Sid": "DataSyncAccessManifest",
    "Effect": "Allow",
    "Action": [
      "s3:GetObject",
      "s3:GetObjectVersion"
    ],
    "Resource": "arn:aws:s3:::manifest-bucket/my-manifest.csv"
  }]
}
```

9. In the sample policy that you just pasted, replace the following values with your own:
  - a. Replace *manifest-bucket* with the name of the S3 bucket that's hosting your manifest.
  - b. Replace *my-manifest.csv* with the file name of your manifest.
10. Choose **Next**. Give your policy a name and choose **Create policy**.
11. (Recommended) To prevent the [cross-service confused deputy problem](#), do the following:
  - a. On the role's details page, choose the **Trust relationships** tab. Choose **Edit trust policy**.

- b. Update the trust policy by using the following example, which includes the `aws:SourceArn` and `aws:SourceAccount` global condition context keys:

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

- Replace each instance *account-id* with the AWS account ID where you're using DataSync.
  - Replace *region* with the AWS Region where you're using DataSync.
- c. Choose **Update policy**.

You've created an IAM role that allows DataSync to access your manifest. Specify this role when [creating](#) or [starting](#) your task.

### Creating the IAM role (different account)

If your manifest is in an S3 bucket that belongs to a different AWS account, you must manually create the IAM role that DataSync uses to access the manifest. Then, in the AWS account where your manifest is located, you need to include the role in the S3 bucket policy.

### Creating the role

1. Open the IAM console at <https://console.aws.amazon.com/iam/>.

2. In the left navigation pane, under **Access management**, choose **Roles**, and then choose **Create role**.
3. On the **Select trusted entity** page, for **Trusted entity type**, choose **AWS service**.
4. For **Use case**, choose **DataSync** in the dropdown list and select **DataSync**. Choose **Next**.
5. On the **Add permissions** page, choose **Next**. Give your role a name and choose **Create role**.
6. On the **Roles** page, search for the role that you just created and choose its name.
7. On the role's details page, choose the **Permissions** tab. Choose **Add permissions** then **Create inline policy**.
8. Choose the **JSON** tab and paste the following sample policy into the policy editor:

```
{
  "Version": "2012-10-17",
  "Statement": [{
    "Sid": "DataSyncAccessManifest",
    "Effect": "Allow",
    "Action": [
      "s3:GetObject",
      "s3:GetObjectVersion"
    ],
    "Resource": "arn:aws:s3:::manifest-bucket/my-manifest.csv"
  }]
}
```

9. In the sample policy that you just pasted, replace the following values with your own:
  - a. Replace *manifest-bucket* with the name of the S3 bucket that's hosting your manifest.
  - b. Replace *my-manifest.csv* with the file name of your manifest.
10. Choose **Next**. Give your policy a name and choose **Create policy**.
11. (Recommended) To prevent the [cross-service confused deputy problem](#), do the following:
  - a. On the role's details page, choose the **Trust relationships** tab. Choose **Edit trust policy**.
  - b. Update the trust policy by using the following example, which includes the `aws:SourceArn` and `aws:SourceAccount` global condition context keys:

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

```

    "Principal": {
      "Service": "datasync.amazonaws.com"
    },
    "Action": "sts:AssumeRole",
    "Condition": {
      "StringEquals": {
        "aws:SourceAccount": "account-id"
      },
      "StringLike": {
        "aws:SourceArn": "arn:aws:datasync:region:account-id:*"
      }
    }
  }
}

```

- Replace each instance of *account-id* with the AWS account ID where you're using DataSync.
  - Replace *region* with the AWS Region where you're using DataSync.
- c. Choose **Update policy**.

You created the IAM role that you can include in your S3 bucket policy.

### Updating your S3 bucket policy with the role

Once you've created the IAM role, you must add it to the S3 bucket policy in the other AWS account where your manifest is located.

1. In the AWS Management Console, switch over to the account with your manifest's S3 bucket.
2. Open the Amazon S3 console at <https://console.aws.amazon.com/s3/>.
3. On the bucket's detail page, choose the **Permissions** tab.
4. Under **Bucket policy**, choose **Edit** and do the following to modify your S3 bucket policy:
  - a. Update what's in the editor to include the following policy statements:

```

{
  "Version": "2008-10-17",
  "Statement": [
    {
      "Sid": "DataSyncAccessManifestBucket",
      "Effect": "Allow",

```

```
"Principal": {
  "AWS": "arn:aws:iam::account-id:role/datasync-role"
},
"Action": [
  "s3:GetObject",
  "s3:GetObjectVersion"
],
"Resource": "arn:aws:s3::manifest-bucket"
}
]
```

- b. Replace *account-id* with the AWS account ID for the account that you're using DataSync with.
  - c. Replace *datasync-role* with the IAM role that you just created that allows DataSync to access your manifest.
  - d. Replace *manifest-bucket* with the name of the S3 bucket that's hosting your manifest in the other AWS account.
5. Choose **Save changes**.

You've created an IAM role that allows DataSync to access your manifest in the other account. Specify this role when [creating](#) or [starting](#) your task.

## Specifying your manifest when creating a task

You can specify the manifest that you want DataSync to use when creating a task.

### Using the DataSync console

1. Open the AWS DataSync console at <https://console.aws.amazon.com/datasync/>.
2. In the left navigation pane, choose **Tasks**, and then choose **Create task**.
3. Configure your task's source and destination locations.

For more information, see [Where can I transfer my data with AWS DataSync?](#)

4. For **Contents to scan**, choose **Specific files, objects, and folders**, then select **Using a manifest**.
5. For **S3 URI**, choose your manifest that's hosted on an S3 bucket.

Alternatively, you can enter the URI (for example, `s3://bucket/prefix/my-manifest.csv`).

6. For **Object version**, choose the version of the manifest that you want DataSync to use.

By default, DataSync uses the latest version of the object.

7. For **Manifest access role**, do one of the following:
  - Choose **Autogenerate** for DataSync to automatically create an IAM role with the permissions required to access your manifest in its S3 bucket.
  - Choose an existing IAM role that can access your manifest.

For more information, see [Providing DataSync access to your manifest](#).

8. Configure any other task settings you need, then choose **Next**.
9. Choose **Create task**.

## Using the AWS CLI

1. Copy the following create-task command:

```
aws datasync create-task \  
  --source-location-arn arn:aws:datasync:us-east-1:123456789012:location/  
loc-12345678abcdefgh \  
  --destination-location-arn arn:aws:datasync:us-east-1:123456789012:location/loc-  
abcdefgh12345678 \  
  --manifest-config {  
    "Source": {  
      "S3": {  
        "ManifestObjectPath": "s3-object-key-of-manifest",  
        "BucketAccessRoleArn": "bucket-iam-role",  
        "S3BucketArn": "DOC-EXAMPLE-BUCKET",  
        "ManifestObjectId": "manifest-version-to-use"  
      }  
    }  
  }  
}
```

2. For the `--source-location-arn` parameter, specify the Amazon Resource Name (ARN) of the location that you're transferring data from.
3. For the `--destination-location-arn` parameter, specify the ARN of the location that you're transferring data to.
4. For the `--manifest-config` parameter, do the following:



- `ManifestObjectPath` – Specify the S3 object key of your manifest.
- `BucketAccessRoleArn` – Specify the IAM role that allows DataSync to access your manifest in its S3 bucket.

For more information, see [Providing DataSync access to your manifest](#).

- `S3BucketArn` – Specify the ARN of the S3 bucket that's hosting your manifest.
- `ManifestObjectVersionId` – Specify the version of the manifest that you want DataSync to use.

By default, DataSync uses the latest version of the object.

5. Run the `create-task` command to create your task.

When you're ready, you can [start your transfer task](#).

## Specifying your manifest when starting a task

You can specify the manifest that you want DataSync to use when executing a task.

### Using the DataSync console

1. Open the AWS DataSync console at <https://console.aws.amazon.com/datasync/>.
2. In the left navigation pane, choose **Tasks**, and then choose the task that you want to start.
3. In the task overview page, choose **Start**, and then choose **Start with overriding options**.
4. For **Contents to scan**, choose **Specific files, objects, and folders**, then select **Using a manifest**.
5. For **S3 URI**, choose your manifest that's hosted on an S3 bucket.

Alternatively, you can enter the URI (for example, `s3://bucket/prefix/my-manifest.csv`).

6. For **Object version**, choose the version of the manifest that you want DataSync to use.

By default, DataSync uses the latest version of the object.

7. For **Manifest access role**, do one of the following:
  - Choose **Autogenerate** for DataSync to automatically create an IAM role to access your manifest in its S3 bucket.

- Choose an existing IAM role that can access your manifest.

For more information, see [Providing DataSync access to your manifest](#).

8. Choose **Start** to begin your transfer.

## Using the AWS CLI

1. Copy the following `start-task-execution` command:

```
aws datasync start-task-execution \  
  --task-arn arn:aws:datasync:us-east-1:123456789012:task/task-12345678abcdefgh \  
  --manifest-config {  
    "Source": {  
      "S3": {  
        "ManifestObjectPath": "s3-object-key-of-manifest",  
        "BucketAccessRoleArn": "bucket-iam-role",  
        "S3BucketArn": "DOC-EXAMPLE-BUCKET",  
        "ManifestObjectVersionId": "manifest-version-to-use"  
      }  
    }  
  }  
}
```

2. For the `--task-arn` parameter, specify the Amazon Resource Name (ARN) of the task that you're starting.
3. For the `--manifest-config` parameter, do the following:
  - `ManifestObjectPath` – Specify the S3 object key of your manifest.
  - `BucketAccessRoleArn` – Specify the IAM role that allows DataSync to access your manifest in its S3 bucket.

For more information, see [Providing DataSync access to your manifest](#).

- `S3BucketArn` – Specify the ARN of the S3 bucket that's hosting your manifest.
- `ManifestObjectVersionId` – Specify the version of the manifest that you want DataSync to use.

By default, DataSync uses the latest version of the object.

4. Run the `start-task-execution` command to begin your transfer.

## Limitations

- You can't use a manifest together with [filters](#).
- You can't specify only a directory or folder with the intention of transferring all of its contents. For these situations, consider using an [include filter](#) instead of a manifest.
- You can't use the **Keep deleted files** task option (PreserveDeletedFiles in the [API](#)) to [maintain files or objects in the destination that aren't in the source](#). DataSync only transfers what's listed in your manifest and doesn't delete anything in the destination.

## Troubleshooting

If you're transferring objects with specific version IDs from an S3 bucket, you might see an error related to HeadObject or GetObjectTagging. For example, here's an error related to GetObjectTagging:

```
[WARN] Failed to read metadata for file /picture1.png (versionId: 111111): S3 Get Object Tagging Failed
[ERROR] S3 Exception: op=GetObjectTagging photos/picture1.png, code=403, type=15, exception=AccessDenied,
msg=Access Denied req-hdrs: content-type=application/xml, x-amz-api-version=2006-03-01
rsp-hdrs: content-type=application/xml,
date=Wed, 07 Feb 2024 20:16:14 GMT, server=AmazonS3, transfer-encoding=chunked,
x-amz-id-2=I0WQ4fDEXAMPLEQM+ey7N9WgVhSnQ6JEXAMPLEZb7hSQDASK+Jd1vEXAMPLEa3Km, x-amz-request-id=79104EXAMPLEB723
```

If you see either of these errors, validate that the IAM role that DataSync uses to access your S3 source location has the following permissions:

- s3:GetObjectVersion
- s3:GetObjectVersionTagging

If you need to update your role with these permissions, see [Creating an IAM role for DataSync to access your Amazon S3 location](#).

## Next steps

If you haven't already, [start your task](#). Otherwise, [monitor your task's activity](#).

## Specifying what AWS DataSync transfers by using filters

AWS DataSync lets you apply filters to include or exclude data from your source location in a transfer. For example, if you don't want to transfer temporary files that end with `.tmp`, you can create an exclude filter so that these files don't make their way to your destination location.

You can use a combination of exclude and include filters in the same transfer task. If you modify a task's filters, those changes are applied the next time you run the task.

### Filtering terms, definitions, and syntax

Familiarize yourself with the concepts related to DataSync filtering:

#### Filter

The whole string that makes up a particular filter (for example, `*.tmp|*.temp` or `/folderA|/folderB`).

Filters are made up of patterns delimited by using a pipe (`|`). You don't need a delimiter when you add patterns in the DataSync console because you add each pattern separately.

#### Note

Filters are case sensitive. For example, filter `/folderA` won't match `/FolderA`.

#### Pattern

A pattern within a filter. For example, `*.tmp` is a pattern that's part of the `*.tmp|*.temp` filter. If your filter has multiple patterns, you delimit each pattern by using a pipe (`|`).

#### Folders

- All filters are relative to the source location path. For example, suppose that you specify `/my_source/` as the source path when you create your source location and task and specify the include filter `/transfer_this/`. In this case, DataSync transfers only the directory `/my_source/transfer_this/` and its contents.
- To specify a folder directly under the source location, include a forward slash (`/`) in front of the folder name. In the example preceding, the pattern uses `/transfer_this`, not `transfer_this`.

- DataSync interprets the following patterns the same way and matches both the folder and its content.

`/dir`

`/dir/`

- When you are transferring data from or to an Amazon S3 bucket, DataSync treats the `/` character in the object key as the equivalent of a folder on a file system.

## Special characters

Following are special characters for use with filtering.

Special character	Description
* (wildcard)	A character used to match zero or more characters. For example, <code>/movies_folder*</code> matches both <code>/movies_folder</code> and <code>/movies_folder1</code> .
(pipe delimiter)	A character used as a delimiter between patterns. It enables specifying multiple patterns, any of which can match the filter. For example, <code>*.tmp .temp</code> matches files ending with either <code>tmp</code> or <code>temp</code> . <div data-bbox="625 1197 1507 1465" style="border: 1px solid #add8e6; border-radius: 10px; padding: 10px; margin-top: 10px;"> <p><b>Note</b></p> <p>This delimiter isn't needed when you add patterns on the console because you add each pattern on a separate line.</p> </div>
\ (backslash)	A character used for escaping special characters ( <code>*</code> , <code> </code> , <code>\</code> ) in a file or object name. <p>A double backslash (<code>\\</code>) is required when a backslash is part of a file name. Similarly, <code>\\\\</code> represents two consecutive backslashes in a file name.</p> <p>A backslash followed by a pipe (<code>\\ </code>) is required when a pipe is part of a file name.</p>

Special character	Description
	A backslash (\) followed by any other character, or at the end of a pattern, is ignored.

## Example filters

The following examples show common filters you can use with DataSync.

### Note

There are limits to how many characters you can use in a filter. For more information, see [DataSync task quotas](#).

## Exclude some folders from your source location

In some cases, you might want to exclude folders in your source location to not copy them to your destination location. For example, if you have temporary work-in-progress folders, you can use something like the following filter:

```
*/.temp
```

To exclude folders with similar content (such as `/reports2021` and `/reports2022`), you can use an exclude filter like the following:

```
/reports*
```

To exclude folders at any level in the file hierarchy, you can use an exclude filter like the following.

```
*/folder-to-exclude-1|*/folder-to-exclude-2
```

To exclude folders at the top level of the source location, you can use an exclude filter like the following.

```
/top-level-folder-to-exclude-1|/top-level-folder-to-exclude-2
```

## Include a subset of the folders on your source location

In some cases, your source location might be a large share and you need to transfer a subset of the folders under the root. To include specific folders, start a task execution with an include filter like the following.

```
/folder-to-transfer/*
```

## Exclude specific file types

To exclude certain file types from the transfer, you can create a task execution with an exclude filter such as `*.temp`.

## Transfer individual files you specify

To transfer a list of individual files, start a task execution with an include filter like the following: `"/folder/subfolder/file1.txt|/folder/subfolder/file2.txt|/folder/subfolder/file2.txt"`

## Creating include filters

Include filters define the files, objects, and folders that you want DataSync to transfer. You can configure include filters when you create, edit, or start a task.

DataSync scans and transfers only files and folders that match the include filters. For example, to include a subset of your source folders, you might specify `/important_folder_1|important_folder_2`.

### Note

Include filters support the wildcard (\*) character only as the rightmost character in a pattern. For example, `/documents*/|code*` is supported, but `*.txt` isn't.

## Using the DataSync console

1. Open the AWS DataSync console at <https://console.aws.amazon.com/datasync/>.
2. In the left navigation pane, choose **Tasks**, and then choose **Create task**.
3. Configure your task's source and destination locations.

For more information, see [Where can I transfer my data with AWS DataSync?](#)

4. For **Contents to scan**, choose **Specific files, objects, and folders**, then select **Using filters**.
5. For **Includes**, enter your filter (for example, `/important_folders` to include an important directory), then choose **Add pattern**.

6. Add other include filters as needed.

## Using the AWS CLI

When using the AWS CLI, you must use single quotation marks (') around the filter and a | (pipe) as a delimiter if you have more than one filter.

The following example specifies two include filters `/important_folder1` and `important_folder2` when running the `create-task` command.

```
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' \
  --includes FilterType=SIMPLE_PATTERN,Value='/important_folder1|important_folder2'
```

## Creating exclude filters

Exclude filters define the files, objects, and folders in your source location that you don't want DataSync to transfer. You can configure these filters when you create, edit, or start a task.

### Topics

- [Data excluded by default](#)

### Data excluded by default

DataSync automatically excludes some data from being transferred:

- `.snapshot` – DataSync ignores any path ending with `.snapshot`, which typically is used for point-in-time snapshots of a storage system's files or directories.
- `/.aws-datasync` and `/.awssync` – DataSync creates these folders in your location to help facilitate your transfer.
- `/.zfs` – You might see this folder with Amazon FSx for OpenZFS locations.

### Using the DataSync console

1. Open the AWS DataSync console at <https://console.aws.amazon.com/datasync/>.



2. In the left navigation pane, choose **Tasks**, and then choose **Create task**.
3. Configure your task's source and destination locations.

For more information, see [Where can I transfer my data with AWS DataSync?](#)

4. For **Excludes**, enter your filter (for example, `*/temp` to exclude temporary folders), then choose **Add pattern**.
5. Add other exclude filters as needed.
6. If needed, add [include filters](#).

## Using the AWS CLI

When using the AWS CLI, you must use single quotation marks (') around the filter and a | (pipe) as a delimiter if you have more than one filter.

The following example specifies two exclude filters `*/temp` and `*/tmp` when running the `create-task` command.

```
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' \  
  \  
  --excludes FilterType=SIMPLE_PATTERN,Value='*/temp|*/tmp'
```

## Metadata copied by AWS DataSync

AWS DataSync can preserve your file or object metadata during a data transfer. How your metadata gets copied depends on your transfer locations and if those locations use similar types of metadata.

### System-level metadata

In general, DataSync doesn't copy system-level metadata. For example, when transferring from an SMB file server, the permissions you configured at the file system level aren't copied to the destination storage system.

There are exceptions. When transferring between Amazon S3 and other object storage, DataSync does copy some [system-defined object metadata](#).

## Metadata copied in Amazon S3 transfers

The following tables describe what metadata DataSync can copy when a transfer involves an Amazon S3 location.

### Topics

- [To Amazon S3](#)
- [Between Amazon S3 and other object storage](#)
- [Between Amazon S3 and HDFS](#)

### To Amazon S3

When copying from one of these locations	To this location	DataSync can copy
<ul style="list-style-type: none"> <li>• NFS</li> <li>• Amazon EFS</li> <li>• FSx for Lustre</li> <li>• FSx for OpenZFS</li> <li>• FSx for ONTAP (using NFS)</li> </ul>	<ul style="list-style-type: none"> <li>• Amazon S3</li> </ul>	<p>The following as Amazon S3 user metadata:</p> <ul style="list-style-type: none"> <li>• File and folder modification timestamps</li> <li>• File and folder access timestamps (DataSync can only do this on a best-effort basis)</li> <li>• User ID and group ID</li> <li>• POSIX permissions</li> </ul> <p>The file metadata stored in Amazon S3 user metadata is interoperable with NFS shares on file gateways using AWS Storage Gateway. A file gateway enables low-latency access from on-premises networks to data that was copied to Amazon S3 by</p>

When copying from one of these locations	To this location	DataSync can copy
		<p>DataSync. This metadata is also interoperable with FSx for Lustre.</p> <p>When DataSync copies objects that contain this metadata back to an NFS server, the file metadata is restored. Restoring metadata requires granting elevated permissions to the NFS server. For more information, see <a href="#">Configuring AWS DataSync transfers with an NFS file server</a>.</p>

### Between Amazon S3 and other object storage

When copying between these locations	DataSync can copy
<ul style="list-style-type: none"> <li>• Object storage</li> <li>• Amazon S3</li> </ul>	<ul style="list-style-type: none"> <li>• User-defined object metadata</li> <li>• Object tags</li> <li>• The following system-defined object metadata: <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><b>Note:</b> DataSync copies system-level metadata for all objects during an initial transfer. If you <a href="#">configure your task to transfer only data that has changed</a>,</p>

When copying between these locations	DataSync can copy
	<p>DataSync won't copy system metadata in subsequent transfers unless an object's content or user metadata has also been modified.</p> <p>DataSync doesn't copy other object metadata, such as object access control lists (ACLs), prior object versions, or the Last-Modified key.</p>

### Between Amazon S3 and HDFS

When copying between these locations	DataSync can copy
<ul style="list-style-type: none"> <li>• Hadoop Distributed File System (HDFS)</li> <li>• Amazon S3</li> </ul>	<p>The following as Amazon S3 user metadata:</p> <ul style="list-style-type: none"> <li>• File and folder modification timestamps</li> <li>• File and folder access timestamps (DataSync can only do this on a best-effort basis)</li> <li>• User ID and group ID</li> <li>• POSIX permissions</li> </ul> <p>HDFS uses strings to store file and folder user and group ownership, rather than numeric identifiers, such as UIDs and GIDs.</p>

### Metadata copied in NFS transfers

The following table describes what metadata DataSync can copy between locations that use Network File System (NFS).

When copying between these locations	DataSync can copy
<ul style="list-style-type: none"> <li>• NFS</li> </ul>	<ul style="list-style-type: none"> <li>• File and folder modification timestamps</li> </ul>

When copying between these locations	DataSync can copy
<ul style="list-style-type: none"> <li>• Amazon EFS</li> <li>• Amazon FSx for Lustre</li> <li>• Amazon FSx for OpenZFS</li> <li>• Amazon FSx for NetApp ONTAP (using NFS)</li> </ul>	<ul style="list-style-type: none"> <li>• File and folder access timestamps (DataSync can only do this on a best-effort basis)</li> <li>• User ID (UID) and group ID (GID)</li> <li>• POSIX permissions</li> </ul>

## Metadata copied in SMB transfers

The following table describes what metadata DataSync can copy between locations that use Server Message Block (SMB).

When copying between these locations	DataSync can copy
<ul style="list-style-type: none"> <li>• SMB</li> <li>• Amazon FSx for Windows File Server</li> <li>• FSx for ONTAP (using SMB)</li> </ul>	<ul style="list-style-type: none"> <li>• File timestamps: access time, modification time, and creation time</li> <li>• File owner security identifier (SID)</li> <li>• Standard file attributes: read-only (R), archive (A), system (S), hidden (H), compressed (C), not content indexed (I), encrypted (E), temporary (T), offline (O), and sparse (P)</li> </ul> <p>DataSync attempts to copy the archive (A), compressed (C), not context indexed (I), sparse (P), and temporary (T) attributes on a best-effort basis. If these attributes aren't applied on the destination, they're ignored during task verification.</p> <ul style="list-style-type: none"> <li>• NTFS discretionary access lists (DACLS), which determine whether to grant access to an object.</li> <li>• NTFS system access control lists (SACLs), which are used by administrators to log attempts to access a secured object.</li> </ul>

When copying between these locations	DataSync can copy
	<p><b>Note:</b> SACLs are not copied if you use SMB version 1.0.</p> <p>Copying DACLs and SACLs requires granting specific permissions to the Windows user that DataSync uses to access your location using SMB. For more information, see creating a location for <a href="#">SMB</a>, <a href="#">FSx for Windows File Server</a>, or <a href="#">FSx for ONTAP</a> (depending on the type of location in your transfer).</p>

## Metadata copied in other transfer scenarios

DataSync handles metadata the following ways when copying between these storage systems (most of which have different metadata structures).

When copying from one of these locations	To one of these locations	DataSync can copy
<ul style="list-style-type: none"> <li>• SMB</li> <li>• FSx for Windows File Server</li> <li>• FSx for ONTAP (using SMB)</li> </ul>	<ul style="list-style-type: none"> <li>• Amazon EFS</li> <li>• FSx for Lustre</li> <li>• FSx for OpenZFS</li> <li>• FSx for ONTAP (using NFS)</li> <li>• Amazon S3</li> <li>• Object storage</li> <li>• Azure Blob Storage</li> <li>• NFS</li> </ul>	<p><a href="#">Default POSIX metadata</a> for all files and folders on the destination file system or objects in the destination S3 bucket. This approach includes using the default POSIX user ID and group ID values.</p> <p>Windows-based metadata (such as ACLs) is not preserved.</p>
<ul style="list-style-type: none"> <li>• Object storage</li> </ul>	<ul style="list-style-type: none"> <li>• Amazon EFS</li> </ul>	<p><a href="#">Default POSIX metadata</a> on the destination files</p>

When copying from one of these locations	To one of these locations	DataSync can copy
<ul style="list-style-type: none"> <li>• Amazon S3</li> <li>• Azure Blob Storage</li> </ul>	<ul style="list-style-type: none"> <li>• FSx for Lustre</li> <li>• FSx for OpenZFS</li> <li>• FSx for ONTAP (using NFS)</li> </ul>	<p>and folders. This approach includes using the default POSIX user ID and group ID values.</p>
<ul style="list-style-type: none"> <li>• Amazon EFS</li> <li>• FSx for Lustre</li> <li>• FSx for OpenZFS</li> <li>• FSx for ONTAP (using NFS)</li> </ul>	<ul style="list-style-type: none"> <li>• Azure Blob Storage</li> </ul>	<p>The following as user-defined metadata:</p> <ul style="list-style-type: none"> <li>• File and folder modification timestamps</li> <li>• File and folder access timestamps (DataSync can only do this on a best-effort basis)</li> <li>• User ID and group ID</li> <li>• POSIX permissions</li> </ul>

When copying from one of these locations	To one of these locations	DataSync can copy
<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> <li>• FSx for OpenZFS</li> <li>• FSx for ONTAP (using NFS)</li> </ul>	<ul style="list-style-type: none"> <li>• File and folder modification timestamps</li> <li>• File and folder access timestamps (DataSync can only do this on a best-effort basis)</li> <li>• POSIX permissions</li> </ul> <p>HDFS stores file and folder user and group ownership as strings rather than numeric identifiers (such as UIDs and GIDs). Default values for UIDs and GIDs are applied on the destination file system. For more information, see <a href="#">Understanding when and how DataSync applies default POSIX metadata</a>.</p>
<ul style="list-style-type: none"> <li>• Amazon S3</li> <li>• Amazon EFS</li> <li>• FSx for Lustre</li> <li>• FSx for OpenZFS</li> <li>• FSx for Windows File Server</li> <li>• FSx for ONTAP</li> </ul>	<ul style="list-style-type: none"> <li>• HDFS</li> </ul>	<p>File and folder timestamps from the source location. The file or folder owner is set based on the HDFS user or Kerberos principal you specified when creating the <a href="#">HDFS transfer location</a>. The Groups Mapping configuration on the Hadoop cluster determines the group.</p>



When copying from one of these locations	To one of these locations	DataSync can copy
<ul style="list-style-type: none"> <li>Amazon S3</li> <li>Amazon EFS</li> <li>FSx for Lustre</li> <li>FSx for OpenZFS</li> <li>FSx for ONTAP (using NFS)</li> <li>Object storage</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>FSx for ONTAP (using SMB)</li> </ul>	File and folder timestamps from the source location. Ownership is set based on the Windows user that was specified in DataSync to access the Amazon FSx or SMB share. Permissions are inherited from the parent directory.
<ul style="list-style-type: none"> <li>Azure Blob Storage</li> </ul>	<ul style="list-style-type: none"> <li>FSx for Windows File Server</li> <li>FSx for ONTAP (using SMB)</li> </ul>	

## Understanding when and how DataSync applies default POSIX metadata

DataSync applies default POSIX metadata in the following situations:

- When your transfer's source and destination locations don't have similar metadata structures
- When metadata is missing from the source location

The following table describes how DataSync applies default POSIX metadata during these types of transfers:

Source	Destination	File permissions	Folder permissions	UID	GID
<ul style="list-style-type: none"> <li>Amazon S3<sup>1</sup></li> <li>Object storage<sup>1</sup></li> </ul>	<ul style="list-style-type: none"> <li>Amazon EFS</li> <li>FSx for Lustre</li> <li>FSx for OpenZFS</li> </ul>	0755	0755	65534	65534

Source	Destination	File permissions	Folder permissions	UID	GID
<ul style="list-style-type: none"> <li>Microsoft Azure Blob Storage<sup>1</sup></li> </ul>	<ul style="list-style-type: none"> <li>FSx for ONTAP (using NFS)</li> <li>NFS</li> </ul>				
<ul style="list-style-type: none"> <li>SMB</li> </ul>	<ul style="list-style-type: none"> <li>Amazon S3</li> <li>Object storage</li> <li>Amazon EFS</li> <li>FSx for Lustre</li> <li>FSx for OpenZFS</li> <li>FSx for ONTAP (using NFS)</li> <li>NFS</li> </ul>	0644	0755	65534	65534
<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> <li>FSx for OpenZFS</li> <li>FSx for ONTAP (using NFS)</li> <li>NFS</li> </ul>	0644	0755	65534	65534

<sup>1</sup> In cases where the objects don't have metadata that was previously applied by DataSync.

## Links and directories copied by AWS DataSync

AWS DataSync handles copied hard links, symbolic links, and directories differently depending on the storage locations involved in your transfer.

### Hard links

Here's how DataSync handles hard links in some common transfer scenarios:

- **When copying between an NFS file server, FSx for Lustre, FSx for OpenZFS, FSx for ONTAP (using NFS), and Amazon EFS**, hard links are preserved.
- **When copying to Amazon S3**, each underlying file referenced by a hard link is transferred only once. During incremental copies, separate objects are created in your S3 bucket. If a hard link is unchanged in Amazon S3, it's correctly restored when transferred to an NFS file server, FSx for Lustre, FSx for OpenZFS, FSx for ONTAP (using NFS), or Amazon EFS file system.
- **When copying to Microsoft Azure Blob Storage**, each underlying file referenced by a hard link is transferred only once. During incremental copies, separate objects are created in your blob storage if there are new references in the source. When copying from Azure Blob Storage, DataSync transfers hard links as if they are individual files.
- **When copying between an SMB file server, FSx for Windows File Server, and FSx for ONTAP (using SMB)**, hard links aren't supported. If DataSync encounters hard links in these situations, the transfer task completes with an error. To learn more, check your CloudWatch logs.
- **When copying to HDFS**, hard links aren't supported. CloudWatch logs show these links as skipped.

### Symbolic links

Here's how DataSync handles symbolic links in some common transfer scenarios:

- **When copying between an NFS file server, FSx for Lustre, FSx for OpenZFS, FSx for ONTAP (using NFS), and Amazon EFS**, symbolic links are preserved.
- **When copying to Amazon S3**, the link target path is stored in the Amazon S3 object. The link is correctly restored when transferred to an NFS file server, FSx for Lustre, FSx for OpenZFS, FSx for ONTAP, or Amazon EFS file system.
- **When copying to Azure Blob Storage**, symbolic links aren't supported. CloudWatch logs show these links as skipped.

- **When copying between an SMB file server, FSx for Windows File Server, and FSx for ONTAP (using SMB)**, symbolic links aren't supported. If DataSync encounters symbolic links in these situations, the task completes with an error. To learn more, check your CloudWatch logs.
- **When copying to HDFS**, symbolic links aren't supported. CloudWatch logs show these links as skipped.

## Directories

In general, DataSync preserves directories when transferring between storage systems. This isn't the case in the following situations:

- **When copying to Amazon S3**, directories are represented as empty objects that have prefixes and end with a forward slash (/).
- **When copying to Azure Blob Storage without a hierarchical namespace**, directories don't exist. What looks like a directory is just part of an object name.

## Configuring how AWS DataSync handles files, objects, and metadata

You can configure how AWS DataSync handles your files, objects, and their associated metadata when transferring between locations.

For example, with recurring transfers, you might want to overwrite files in your destination with changes in the source to keep the locations in sync. You can copy properties such as POSIX permissions for files and folders, tags associated with objects, and access control lists (ACLs).

**Before you begin:** The metadata that DataSync can preserve in a transfer depends on the storage systems involved and whether those systems use a similar metadata structure. Before configuring your task, make sure you understand how DataSync handles [metadata](#) and [special files](#).

### Using the DataSync console

The following instructions describe how to configure what DataSync transfers when creating a task. You also can configure this when editing or starting a task.

1. Open the AWS DataSync console at <https://console.aws.amazon.com/datasync/>.
2. In the left navigation pane, expand **Data transfer**, then choose **Tasks**, and then choose **Create task**.

3. Configure your task's source and destination locations.

For more information, see [Where can I transfer my data with AWS DataSync?](#)

4. For **Transfer mode**, choose one of the following options:
  - **Transfer only data that has changed** – DataSync copies only the data and metadata that differs between the source and destination location.
  - **Transfer all data** – DataSync copies everything in the source to the destination without comparing differences between the locations.
5. Select **Keep deleted files** if you want DataSync to maintain files or objects in the destination location that don't exist in the source.

If you don't choose this option and your task deletes objects from your Amazon S3 bucket, you might incur minimum storage duration charges for certain storage classes. For detailed information, see [Storage class considerations with Amazon S3 transfers](#).

 **Warning**

You can't deselect this option and enable **Transfer all data**. When you transfer all data, DataSync doesn't scan your destination location and doesn't know what to delete.

6. Select **Overwrite files** if you want DataSync to modify data in the destination location when the source data or metadata has changed.

If your task overwrites objects, you might incur additional charges for certain storage classes (for example, for retrieval or early deletion). For detailed information, see [Storage class considerations with Amazon S3 transfers](#).

If you don't choose this option, the destination data isn't overwritten even if the source data differs.

7. Under **Transfer options**, select how you want DataSync to copy metadata.

 **Important**

The settings you see in the console depend on your task's source and destination locations. In some cases, you may have to expand **Additional settings** to see these options.

Setting	Description
<b>Copy ownership</b>	DataSync copies POSIX file and folder ownership, such as the group ID of the file's owners and the user ID of the file's owner.
<b>Copy permissions</b>	DataSync copies POSIX permissions for files and folders from the source to the destination.
<b>Copy timestamps</b>	DataSync copies the timestamp metadata from the source to the destination.
<b>Copy object tags</b>	DataSync preserves the tags associated with your objects when transferring between object storage systems.

Setting	Description
<b>Copy ownership, DACLs, and SACLs</b>	<p>DataSync copies the following:</p> <ul style="list-style-type: none"><li>• The object owner.</li><li>• NTFS discretionary access lists (DACLs), which determine whether to grant access to an object.</li><li>• NTFS system access control lists (SACLs), which are used by administrators to log attempts to access a secured object.</li></ul> <p><b>Note:</b> SACLs are not copied if you use SMB version 1.0.</p> <p>Copying DACLs and SACLs requires granting specific permissions to the Windows user that DataSync uses to access your location using SMB. For more information, see <a href="#">creating a location for SMB</a>, <a href="#">FSx for Windows File Server</a>, or <a href="#">FSx for ONTAP</a> (depending on the type of location in your transfer).</p>
<b>Copy ownership and DACLs</b>	<p>DataSync copies the following:</p> <ul style="list-style-type: none"><li>• The object owner.</li><li>• DACLs, which determine whether to grant access to an object.</li></ul> <p>DataSync won't copy SACLs when you choose this option.</p>

Setting	Description
<b>Do not copy ownership or ACLs</b>	DataSync doesn't copy any ownership or permissions data. The objects that DataSync writes to your destination location are owned by the user whose credentials are provided for DataSync to access the destination. Destination object permissions are determined based on the permissions configured on the destination server.

## Using the DataSync API

You can configure these task settings by using the `Options` parameter with any of the following operations:

- [CreateTask](#)
- [StartTaskExecution](#)
- [UpdateTask](#)

## Configuring how AWS DataSync verifies data integrity

During a transfer, AWS DataSync always checks the integrity of your data, but you can specify how and when this verification happens with the following options:

- **Verify only the data transferred** (recommended) – DataSync calculates the checksum of transferred files and metadata at the source location. At the end of the transfer, DataSync then compares this checksum to the checksum calculated on those files at the destination.

We recommend this option when transferring to S3 Glacier Flexible Retrieval or S3 Glacier Deep Archive storage classes. For more information, see [Storage class considerations with Amazon S3 transfers](#).

- **Verify all data in the destination** – At the end of the transfer, DataSync scans the entire source and destination to verify that both locations are fully synchronized.

If you use a [manifest](#), DataSync only scans and verifies what's listed in the manifest.



You can't use this option when transferring to S3 Glacier Flexible Retrieval or S3 Glacier Deep Archive storage classes. For more information, see [Storage class considerations with Amazon S3 transfers](#).

- **Check integrity during transfer** – DataSync doesn't run additional verification at the end of the transfer. Data integrity is still checked during the transfer with checksum verification.

## Using the DataSync console

The following instructions describe how to configure data verification when creating a task. You also can configure this when editing a task or starting a task execution.

### To configure data verification by using the console

1. Open the AWS DataSync console at <https://console.aws.amazon.com/datasync/>.
2. In the left navigation pane, expand **Data transfer**, then choose **Tasks**, and then choose **Create task**.
3. Configure your task's source and destination locations.

For more information, see [Where can I transfer my data with AWS DataSync?](#)

4. For **Verification**, choose one of the following:
  - **Verify only the data transferred** (recommended)
  - **Verify all data in the destination**
  - **Check integrity during transfer**

## Using the DataSync API

You can configure how DataSync verifies data by using the `VerifyMode` parameter with any of the following operations:

- [CreateTask](#)
- [UpdateTask](#)
- [StartTaskExecution](#)

# Working with AWS DataSync transfer tasks

A *task* describes where and how AWS DataSync transfers data. Tasks consist of the following:

- [Source location](#) – The storage system or service where DataSync transfers data from.
- [Destination location](#) – The storage system or service where DataSync transfers data to.
- [Task settings](#) – Options for configuring how your task behaves, such as how it verifies data, when it runs, and more. Some task settings are optional. For instance, you don't have to give your task a name.
- [Task executions](#) – When you run a task, it's called a *task execution*.

## Topics

- [Setting bandwidth limits for your AWS DataSync task](#)
- [Scheduling your AWS DataSync task](#)
- [Tagging your AWS DataSync tasks](#)
- [Starting your AWS DataSync task](#)
- [Cancelling your AWS DataSync task execution](#)
- [AWS DataSync task statuses](#)

## Setting bandwidth limits for your AWS DataSync task

You can configure network bandwidth limits for your AWS DataSync task and each of its executions.

### Limiting bandwidth for a task

Set a bandwidth limit when creating, editing, or starting a task.

### Using the DataSync console

The following instructions describe how to configure a bandwidth limit for your task when you're creating it.

1. Open the AWS DataSync console at <https://console.aws.amazon.com/datasync/>.
2. In the left navigation pane, expand **Data transfer**, then choose **Tasks**, and then choose **Create task**.

3. Configure your task's source and destination locations.

For more information, see [Where can I transfer my data with AWS DataSync?](#)

4. For **Bandwidth limit**, choose one of the following:

- Select **Use available** to use all of the available network bandwidth for each task execution.
- Select **Set bandwidth limit (MiB/s)** and enter the maximum bandwidth that you want DataSync to use for each task execution.

## Using the DataSync API

You can configure a task's bandwidth limit by using the BytesPerSecond parameter with any of the following operations:

- [CreateTask](#)
- [UpdateTask](#)
- [StartTaskExecution](#)

## Throttling bandwidth for a task execution

You can modify the bandwidth limit for a running or queued task execution.

### Using the DataSync console

1. Open the AWS DataSync console at <https://console.aws.amazon.com/datasync/>.
2. In the navigation pane, expand **Data transfer**. then choose **Tasks**.
3. Choose the task and then select **History** to view the task's executions.
4. Choose the task execution that you want to modify and then choose **Edit**.
5. In the dialog box, choose one of the following:
  - Select **Use available** to use all of the available network bandwidth for the task execution.
  - Select **Set bandwidth limit (MiB/s)** and enter the maximum bandwidth that you want DataSync to use for the task execution.
6. Choose **Save changes**.

The new bandwidth limit takes effect within 60 seconds.

## Using the DataSync API

You can modify the bandwidth limit for a running or queued task execution by using the `BytesPerSecond` parameter with the [UpdateTaskExecution](#) operation.

## Scheduling your AWS DataSync task

You can set up an AWS DataSync task schedule to periodically transfer data between storage locations.

A scheduled task runs at a frequency that you specify, with a minimum interval of 1 hour. For example, your task could run every Sunday and Wednesday at 12:00 PM UTC. Here's how you'd configure this schedule with a cron expression:

```
0 12 ? * SUN,WED *
```

### Important

You can't use a cron expression for scheduling a task to run at an interval faster than 1 hour.

For information about cron expression syntax, see the [Amazon EventBridge User Guide](#).

## Creating a DataSync task schedule

You can schedule how frequently your task runs by using the DataSync console, AWS CLI, or DataSync API.

### Using the DataSync console

The following instructions describe how to set up a schedule when creating a task. You can modify the schedule later when editing the task.

1. Open the AWS DataSync console at <https://console.aws.amazon.com/datasync/>.
2. In the left navigation pane, expand **Data transfer**, then choose **Tasks**, and then choose **Create task**.
3. Configure your task's source and destination locations.

For more information, see [Where can I transfer my data with AWS DataSync?](#)

- For schedule **Frequency**, do one of the following:
  - Choose **Not scheduled** if you don't want to schedule your task to run periodically.
  - Choose **Hourly** and choose the minute in the hour that the task should run. The task runs every hour on the specified minute.
  - Choose **Daily** and enter the UTC time that you want the task to run in the format HH:MM. This task runs every day at the specified time.
  - Choose **Weekly** and the day of the week and enter the UTC time the task should run in the format HH:MM. This task runs every week on the specified day at the specified time.
  - Choose **Days of the week**, choose the specific day or days, and enter the UTC time that the task should run in the format HH:MM. This task runs on the days and the time that you specified.
  - Choose **Custom** if you want to use a custom cron expression to run your task, with a minimum interval of 1 hour. Then enter your expression in the **Cron expression** box.

## Using the AWS CLI

You can create a schedule for your DataSync task by using the `--schedule` parameter with the `create-task`, `update-task`, or `start-task-execution` command.

The following instructions describe how to do this with the `create-task` command.

- Copy the following `create-task` command:

```
aws datasync create-task \  
  --source-location-arn arn:aws:datasync:us-east-1:123456789012:location/  
loc-12345678abcdefgh \  
  --destination-location-arn arn:aws:datasync:us-east-1:123456789012:location/  
loc-abcdefgh12345678 \  
  --schedule '{  
    "ScheduleExpression": "cron(0 12 ? * SUN,WED *)"
```

- For the `--source-location-arn` parameter, specify the Amazon Resource Name (ARN) of the location that you're transferring data from.
- For the `--destination-location-arn` parameter, specify the ARN of the location that you're transferring data to.
- For the `--schedule` parameter, specify a cron expression for your schedule.

In the example, the expression `cron(0 12 ? * SUN,WED *)` sets a task schedule that runs every Sunday and Wednesday at 12:00 PM UTC.

5. Run the `create-task` command to create your task with the schedule.

## Pausing a DataSync task schedule

There can be situations where you need to pause your DataSync task schedule. For example, you might need to temporarily disable a recurring transfer to fix an issue with your task or perform maintenance on your storage system.

DataSync might disable your schedule automatically if your task fails repeatedly with the same error.

### Using the DataSync console

1. Open the AWS DataSync console at <https://console.aws.amazon.com/datasync/>.
2. In the left navigation pane, expand **Data transfer**, and then choose **Tasks**.
3. Choose the task that you want to pause the schedule for, and then choose **Edit**.
4. For **Schedule**, turn off **Enable schedule**. Choose **Save changes**.

### Using the AWS CLI

1. Copy the following `update-task` command:

```
aws datasync update-task \  
  --task-arn arn:aws:datasync:us-east-1:123456789012:task/task-12345678abcdefgh \  
  --schedule '{  
    "ScheduleExpression": "cron(0 12 ? * SUN,WED *)",  
    "Status": "DISABLED"  
  }'
```

2. For the `--task-arn` parameter, specify the ARN of the task that you want to pause the schedule for.
3. For the `--schedule` parameter, do the following:
  - For `ScheduleExpression`, specify a cron expression for your schedule.

In the example, the expression `cron(0 12 ? * SUN,WED *)` sets a task schedule that runs every Sunday and Wednesday at 12:00 PM UTC.

- For Status, specify DISABLED to pause the task schedule.
4. Run the `update-task` command.
  5. To resume the schedule, run the same `update-task` command with Status set to ENABLED.

## Checking the status of a DataSync task schedule

You can see whether your DataSync task schedule is enabled.

### Using the DataSync console

1. Open the AWS DataSync console at <https://console.aws.amazon.com/datasync/>.
2. In the left navigation pane, expand **Data transfer**, and then choose **Tasks**.
3. In the **Schedule** column, check whether the task's schedule is enabled or disabled.

### Using the AWS CLI

1. Copy the following `describe-task` command:

```
aws datasync describe-task \  
  --task-arn arn:aws:datasync:us-east-1:123456789012:task/task-12345678abcdefgh
```

2. For the `--task-arn` parameter, specify the ARN of the task that you want information about.
3. Run the `describe-task` command.

You get a response that provides details about your task, including its schedule. (The following example focuses primarily on the task schedule configuration and doesn't show a full `describe-task` response.)

The example shows that the task's schedule is manually disabled. If the schedule is disabled by the DataSync SERVICE, you see an error message for `DisabledReason` to help you understand why the task keeps failing. For more information, see [???](#).

```
{
```

```
"TaskArn": "arn:aws:datsync:us-east-1:123456789012:task/task-12345678abcdefgh",
>Status": "AVAILABLE",
>Schedule": {
  >ScheduleExpression": "cron(0 12 ? * SUN,WED *)",
  >Status": "DISABLED",
  >StatusUpdateTime": 1697736000,
  >DisabledBy": "USER",
  >DisabledReason": "Manually disabled by user."
},
...
}
```

## Tagging your AWS DataSync tasks

*Tags* are key-value pairs that help you manage, filter, and search for your AWS DataSync resources. You can add up to 50 tags to each DataSync task and task execution.

For example, you might create a task for a large data migration and tag the task with the key **Project** and value **Large Migration**. To further organize the migration, you could tag one run of the task with the key **Transfer Date** and value **May 2021** (subsequent task executions might be tagged **June 2021**, **July 2021**, and so on).

### Tagging your task

You can tag your DataSync task only when creating the task.

#### Using the DataSync console

1. Open the AWS DataSync console at <https://console.aws.amazon.com/datsync/>.
2. In the left navigation pane, expand **Data transfer**, then choose **Tasks**, and then choose **Create task**.
3. Configure your task's source and destination locations.

For more information, see [Where can I transfer my data with AWS DataSync?](#)

4. On the **Configure settings** page, choose **Add new tag** to tag your task.

#### Using the AWS CLI

1. Copy the following create-task command:



```
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. Specify the following parameters in the command:

- `--source-location-arn` – Specify the Amazon Resource Name (ARN) of the source location in your transfer.
- `--destination-location-arn` – Specify the ARN of the destination location in your transfer.
- `--tags` – Specify the tags that you want to apply to the task.

For more than one tag, separate each key-value pair with a space.

3. (Optional) Specify other parameters that make sense for your transfer scenario.

For a list of `--options`, see the [create-task](#) command.

4. Run the `create-task` command.

You get a response that shows the task that you just created.

```
{  
  "TaskArn": "arn:aws:datasync:us-east-2:123456789012:task/task-  
abcdef01234567890"  
}
```

To view the tags you added to this task, you can use the [list-tags-for-resource](#) command.

## Tagging your task execution

You can tag each run of your DataSync task.

If your task already has tags, remember the following about using tags with task executions:

- If you start your task with the console, its user-created tags are applied automatically to the task execution. However, system-created tags that begin with `aws:` are not applied.

- If you start your task with the DataSync API or AWS CLI, its tags are not applied automatically to the task execution.

## Using the DataSync console

To add, edit, or remove tags from a task execution, you must start the task with overriding options.

1. Open the AWS DataSync console at <https://console.aws.amazon.com/datasync/>.
2. In the left navigation pane, expand **Data transfer**, then choose **Tasks**.
3. Choose the task.
4. Choose **Start**, then choose one of the following options:
  - **Start with defaults** – Applies any tags associated with your task.
  - **Start with overriding options** – Allows you to add, edit, or remove tags for this particular task execution.

## Using the AWS CLI

1. Copy the following `start-task-execution` command:

```
aws datasync start-task-execution \  
  --task-arn 'arn:aws:datasync:region:account-id:task/task-id' \  
  --tags Key=tag-key,Value=tag-value
```

2. Specify the following parameters in the command:
  - `--task-arn` – Specify the ARN of the task that you want to start.
  - `--tags` – Specify the tags that you want to apply to this specific run of the task.

For more than one tag, separate each key-value pair with a space.

3. (Optional) Specify other parameters that make sense for your situation.

For more information, see the [start-task-execution](#) command.

4. Run the `start-task-execution` command.

You get a response that shows the task execution that you just started.

```
{
```

```
"TaskExecutionArn": "arn:aws:datsync:us-east-2:123456789012:task/task-
abcdef01234567890"
}
```

To view the tags you added to this task, you can use the [list-tags-for-resource](#) command.

## Starting your AWS DataSync task

Once you create your AWS DataSync transfer task, you can start moving data. Each run of a task is called a *task execution*.

When you start your task, DataSync performs a recursive directory listing to discover all the files, objects, and associated metadata in the transfer's source and destination locations. DataSync identifies differences between the locations and determines what to copy, a process that usually takes between a few minutes and a few hours.

### Important

If you're planning to transfer data to or from an Amazon S3 location, review [how DataSync can affect your S3 request charges](#) and the [DataSync pricing page](#) before you begin.

## Running multiple tasks

There are scenarios where you may want to run multiple DataSync tasks. For example, you may be copying data on a recurring basis or migrating a lot of files incrementally (something you can do with [task filtering](#)).

Remember the following about running multiple tasks:

- Tasks run in a series (first in, first out). If you run the same task more than once, by default each execution goes into a queue. If you start a task twice, for instance, the second task execution won't start until the first one finishes.
- You can run separate tasks at the same time, but these tasks will be queued if they're using the same DataSync agent.

## Starting your task

Once you've created your task, you can begin moving data right away.

## Using the DataSync console

1. Open the AWS DataSync console at <https://console.aws.amazon.com/datasync/>.
2. In the left navigation pane, expand **Data transfer**, then choose **Tasks**.
3. Choose the task that you want to run.

Make sure that the task has an **Available** status. You also can select multiple tasks.

4. Choose **Actions** and then choose one of the following options:
  - **Start** – Runs the task (or tasks if you selected more than one).
  - **Start with overriding options** – Allows you to modify some of your task settings before you begin moving data. When you're ready, choose **Start**.
5. Choose **See execution details** to view details about the running task execution.

## Using the DataSync API

You can start your task by using the [StartTaskExecution](#) operation. Use the [DescribeTaskExecution](#) operation to get details about the running task execution.

Once started, you can [check the task execution's status](#) as DataSync moves your data. You also can [throttle the task execution's bandwidth](#) if needed.

## Cancelling your AWS DataSync task execution

You can stop any running or queued AWS DataSync task execution.

### To cancel a task execution by using the console

1. Open the AWS DataSync console at <https://console.aws.amazon.com/datasync/>.
2. In the left navigation pane, expand **Data transfer**, then choose **Tasks**.
3. Select the **Task ID** for the running task that you want to monitor.

The task status should be **Running**.

4. Choose **History** to view the task's executions.
5. Select the task execution that you want to stop, and then choose **Stop**.
6. In the dialog box, choose **Stop**.

To cancel a running or queued task by using the DataSync API, see [CancelTaskExecution](#).

## AWS DataSync task statuses

There are statuses to help you understand if your AWS DataSync task is ready to run, in progress, or having an issue.

### Task statuses

When you create an AWS DataSync task, you might see these statuses.

Console status	API status	Description
Available	AVAILABLE	The task is ready to start transferring data.
Running	RUNNING	A task execution is in progress.
Unavailable	UNAVAILABLE	A DataSync agent used by the task is offline.
Queued	QUEUED	Another task execution using the same DataSync agent is in progress. For more information, see <a href="#">Running multiple tasks</a> .

### Task execution statuses

When you start a DataSync task, you might see these statuses.

Console status	API status	Description
Queueing	QUEUED	Another task execution is running and using the same DataSync agent. For more information, see <a href="#">Running multiple tasks</a> .
Launching	LAUNCHING	DataSync is initializing the task execution. This status usually goes quickly but can take up to a few minutes.
Preparing	PREPARING	DataSync is determining what data needs to transfer. This can take just minutes or a few hours depending

Console status	API status	Description
		on the number of files, objects, or directories in both locations.
Transferring	TRANSFERRING	DataSync is performing the actual data transfer.
Verifying	VERIFYING	DataSync is performing a data-integrity check at the end of the transfer.
Success	SUCCESS	The task execution succeeded.
Cancelling	CANCELLING	The task execution is in the process of being cancelled.
Error	ERROR	The task execution failed.

# Discovering your storage with AWS DataSync Discovery

AWS DataSync Discovery helps you accelerate your migration to AWS. With DataSync Discovery, you can do the following:

- **Understand how your on-premises storage is used** – DataSync Discovery provides detailed reporting about your storage system resources, including utilization, capacity, and configuration information.
- **Get recommendations about migrating your data to AWS** – DataSync Discovery can suggest AWS storage services (such as Amazon FSx for NetApp ONTAP, Amazon EFS, and Amazon FSx for Windows File Server) for your data. Recommendations include a cost estimate and help you understand how to configure a suggested storage service. When you're ready, you can then use DataSync to migrate your data to AWS.

## Topics

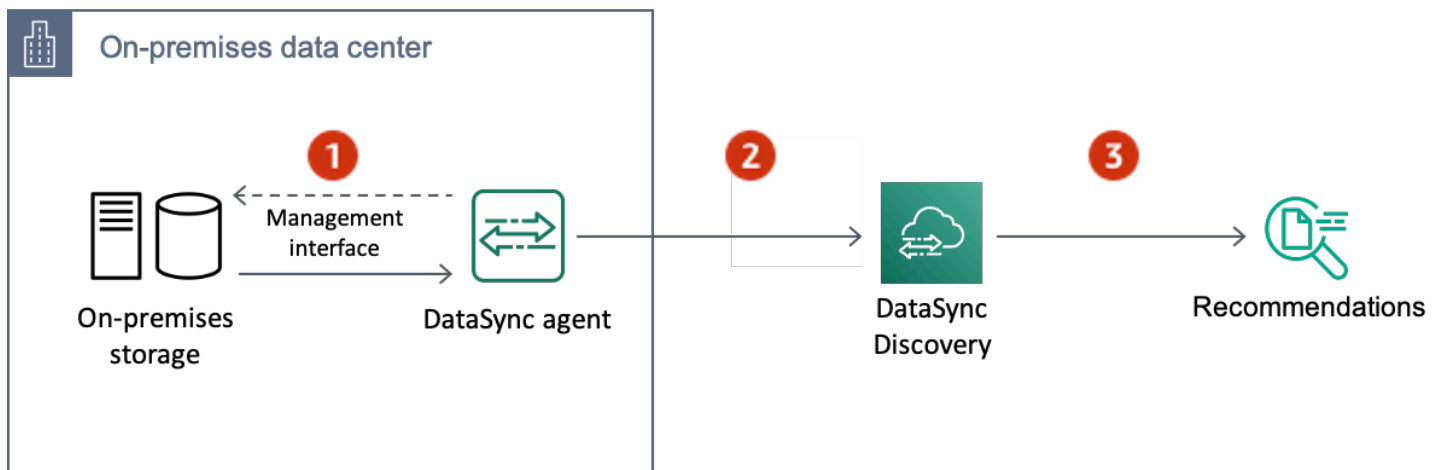
- [How AWS DataSync Discovery works](#)
- [Adding your on-premises storage system to DataSync Discovery](#)
- [Working with DataSync discovery jobs](#)
- [Viewing storage resource information collected by AWS DataSync Discovery](#)
- [Getting recommendations from AWS DataSync Discovery](#)
- [AWS DataSync Discovery statuses](#)

## How AWS DataSync Discovery works

Learn the key concepts and terminology related to AWS DataSync Discovery.

### DataSync Discovery architecture

The following diagram illustrates how DataSync Discovery collects information and provides recommendations for migrating data from an on-premises storage system to AWS.



Reference	Description
1	A DataSync agent connects to your on-premises storage system's management interface (using port 443, for example). You then run a discovery job to collect information about your system.
2	The agent sends the information that it collects to DataSync Discovery through a <a href="#">public service endpoint</a> .
3	Using the information that it collects, DataSync Discovery recommends AWS storage services that you can migrate your data to.

## Concepts and terminology

Familiarize yourself with DataSync Discovery features.

### Topics

- [Agent](#)
- [Discovery job](#)
- [Storage system resource information](#)
- [AWS storage recommendations](#)



## Agent

An *agent* is a virtual machine (VM) appliance that DataSync Discovery uses to access the management interface of your on-premises storage system. The agent collects (reads) information about how your storage resources are performing and being used.

You can deploy an agent in your storage environment on VMware ESXi, Linux Kernel-based Virtual Machine (KVM), or Microsoft Hyper-V hypervisors. For storage in a virtual private cloud (VPC) in AWS, you can deploy an agent as an Amazon EC2 instance.

A DataSync Discovery agent is no different than an agent that you can use for DataSync transfers, but we don't recommend using the same agent for these scenarios.

To get started, see [Deploying your AWS DataSync agent](#).

## Discovery job

You run a *discovery job* to collect information about your on-premises storage system through the storage system's management interface.

You can run a discovery job between 1 hour and 31 days. You'll get more accurate AWS storage recommendations the longer your discovery job runs.

For more information, see [Working with DataSync discovery jobs](#).

## Storage system resource information

DataSync Discovery can give you performance and utilization information about your on-premises storage system's resources. For example, get an idea about how much storage capacity is being used in a specific storage volume compared to how much capacity you originally provisioned.

You can view this information as your discovery job collects it by using the following:

- The [DescribeStorageSystemResources](#) operation
- The [DescribeStorageSystemResourceMetrics](#) operation

For more information, see [Viewing storage resource information collected by AWS DataSync Discovery](#).

## AWS storage recommendations

Using the information that it collects about your on-premises storage system's resources, DataSync Discovery recommends AWS storage services to help plan your migration to AWS.

You can view recommendations by using the [DescribeStorageSystemResources](#) operation.

For more information, see [Getting recommendations from AWS DataSync Discovery](#).

## Limitations

- Currently, you can only activate DataSync Discovery agents with [public service endpoints](#).

## Adding your on-premises storage system to DataSync Discovery

Specify an on-premises storage system that you want AWS DataSync Discovery to collect information about and provide AWS storage migration recommendations for.

### Note

DataSync Discovery currently supports NetApp Fabric-Attached Storage (FAS) and All Flash FAS (AFF) systems that are running ONTAP 9.7 or later.

## Accessing your on-premises storage system

To collect information about your on-premises storage system, DataSync Discovery needs credentials that provide read access to your storage system's management interface. For security, DataSync Discovery stores these credentials in AWS Secrets Manager.

### Important

If you update these credentials on your storage system, make sure to also update them in DataSync Discovery. You can do this by using the [UpdateStorageSystem](#) operation.

## How DataSync Discovery uses AWS Secrets Manager

AWS Secrets Manager is a secret storage service that protects database credentials, API keys, and other secret information. DataSync Discovery uses Secrets Manager to protect the credentials that you provide for accessing your on-premises storage system.

Secrets Manager encrypts secrets using AWS Key Management Service keys. For more information, see [Secret encryption and decryption](#).

You can configure Secrets Manager to automatically rotate secrets for you according to a schedule that you specify. This enables you to replace long-term secrets with short-term ones, which helps to significantly reduce the risk of compromise. For more information, see [Rotate AWS Secrets Manager secrets](#).

You pay for credentials stored in Secrets Manager. For more information, see [AWS Secrets Manager Pricing](#).

## Adding your on-premises storage system

You must provide some information about your storage system before DataSync Discovery can collect information about it.

### Using the AWS CLI

Using the AWS Command Line Interface (AWS CLI), configure DataSync Discovery to work with your on-premises storage system.

**Before you begin:** We recommend that you [enable logging with CloudWatch](#).

### To add an on-premises storage system by using the AWS CLI

1. Copy the following add-storage-system command:

```
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. Specify the following required parameters in the command:

- `--server-configuration ServerHostname` – Specify the domain name or IP address of your storage system's management interface.
- `--server-configuration ServerPort` – Specify the network port that's needed to connect with the system's management interface.
- `--system-type` – Specify the type of storage system that you're adding.
- `--credentials` – Include the following options:
  - `Username` – Specify the user name needed to access your storage system's management interface.
  - `Password` – Specify the password needed to access your storage system's management interface.

For more information, see [Accessing your on-premises storage system](#).

- `--agent-arns` – Specify the DataSync agent that you want to connect to your storage system's management interface.

If you don't have an agent, see [Deploying your AWS DataSync agent](#).

3. (Optional) Add any of the following parameters to the command:

- `--cloud-watch-log-group-arn` – Specify the Amazon Resource Name (ARN) of the CloudWatch log group that you want to use to log DataSync Discovery activity.
- `--tags` – Specify a Key and Value to tag the DataSync resource that's representing your storage system.

A *tag* is a key-value pair that helps you manage, filter, and search for your DataSync resources.

- `--name` – Specify a name for your storage system.

4. Run the `add-storage-system` command.

You get a response that shows you the storage system ARN that you just added.

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

After you add the storage system, you can run a discovery job to collect information about the storage system.

## Removing your on-premises storage system

When you remove an on-premises storage system from DataSync Discovery, you permanently delete any associated discovery jobs, collected data, and recommendations.

### Using the AWS CLI

1. Copy the following `remove-storage-system` command:

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

2. For `--storage-system-arn`, specify the ARN of your storage system.
3. Run the `remove-storage-system` command.

If successful, you get an HTTP 200 response with an empty HTTP body.

## Logging DataSync Discovery activity to Amazon CloudWatch

When you enable logging with Amazon CloudWatch, you can more easily troubleshoot issues with DataSync Discovery. For example, if your discovery job is interrupted, you can check the logs to locate the issue. If you resolve the problem within 12 hours of when it occurred, your discovery job picks up where it left off.

If you configure your system by using the AWS CLI, you must [create a log group](#) with a resource policy that allows DataSync to log events to the log group. You can use a [log group resource policy](#) similar to one for DataSync tasks, with some differences:

- For the service principal, use `discovery-datasync.amazonaws.com`.
- If you're using the `ArnLike` condition, specify a storage system ARN like this:

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

## Working with DataSync discovery jobs

After you deploy your AWS DataSync agent and add your on-premises storage system to DataSync Discovery, you can run discovery jobs to collect information about the system and get AWS migration recommendations.

### Starting a discovery job

You can run a discovery job for up to 31 days. A storage system can have only one active discovery job at a time. The information that a discovery job collects is available for up to 60 days following the end of the job (unless you remove the related storage system from DataSync Discovery before that).

#### Tip

DataSync Discovery can provide more accurate recommendations the longer your discovery job runs. We recommend running a discovery job for at least 14 days.

### Using the AWS CLI

With the AWS Command Line Interface (AWS CLI), you can run a discovery job for as short as 1 hour.

1. Copy the following `start-discovery-job` command:

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

2. Specify the following parameters in the command:
  - `--storage-system-arn` – Specify the Amazon Resource Name (ARN) of the [on-premises storage system that you added](#) to DataSync Discovery.
  - `--collection-duration-minutes` – Specify how long that you want the discovery job to run in minutes. Enter a value between 60 (1 hour) and 44640 (31 days).
3. Run the `start-discovery-job` command.

You get a response that shows the discovery job that you just started.

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

Shortly after starting the discovery job, you can begin [looking at the information that the job collects](#) (including storage system capacity and usage).

## Stopping a discovery job

Stop a discovery job at any time. You can still [get recommendations](#) for a stopped job.

### Using the AWS CLI

1. Copy the following `stop-discovery-job` command:

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

2. For `--discovery-job-arn`, specify the ARN of the discovery job that's currently running.
3. Run the `stop-discovery-job` command.

If successful, you get an HTTP 200 response with an empty HTTP body.

## Viewing storage resource information collected by AWS DataSync Discovery

AWS DataSync Discovery collects information about your on-premises storage system that can help you understand how its storage resources are configured, performing, and utilized. DataSync Discovery uses this information to generate recommendations for migrating your data to AWS.

A discovery job can give you the following information about your storage system's resources (such as its volumes):

- Total, available, and in use storage capacity
- Number of Common Internet File System (CIFS) shares in a resource and whether a resource is available via Network File System (NFS)
- Data transfer protocols

- Performance (such as IOPS, throughput, and latency)

## Viewing information collected about your storage system

You can begin to see what kind of information DataSync Discovery is collecting about your on-premises storage system shortly after you start a discovery job.

You can view this information by using the following options:

- The [DescribeStorageSystemResources](#) operation – Get data about all of the storage system resources that DataSync Discovery can collect information about, including utilization, capacity, and configuration data.
- The [DescribeStorageSystemResourceMetrics](#) operation – Get performance and capacity information that DataSync Discovery can collect about a specific resource in your storage system.

### Using the AWS CLI

The following steps show how to use the [DescribeStorageSystemResources](#) operation with the AWS CLI.

1. Copy the following `describe-storage-system-resources` command:

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

2. Specify the following parameters in the command:
  - `--discovery-job-arn` – Specify the Amazon Resource Name (ARN) of the [discovery job](#) that you ran.
  - `--resource-type` – Specify one of the following values, depending on what kind of storage system resources you want information about:
    - CLUSTER
    - SVM
    - VOLUME
3. (Optional) Specify the `--resource-ids` parameter with the IDs of the storage system resources that you want information about.



#### 4. Run the `describe-storage-system-resources` command.

The following example response returns information that a discovery job collected about two volumes in a storage system.

Note that the `RecommendationStatus` is `NONE` for each volume. To get AWS storage recommendations, you must run the `generate-recommendations` command before the `describe-storage-system-resources` command. For more information, see [Getting recommendations](#).

```
{
  "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-EXAMPLEaaaaa",
    "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"
  }
]
}
```

## Getting recommendations from AWS DataSync Discovery

After AWS DataSync Discovery collects information about your on-premises storage system, it can recommend moving your data on a per-resource basis to one or more of the following AWS storage services:

- [Amazon FSx for NetApp ONTAP](#)
- [Amazon Elastic File System \(Amazon EFS\)](#)
- [Amazon FSx for Windows File Server](#)

## What's included in the recommendations?

DataSync Discovery recommendations include storage configurations and cost estimates to help you choose the AWS storage service that works for your data.

### AWS storage configuration

DataSync Discovery provides information about how you might want to configure a recommended AWS storage service. The storage configuration is designed to optimize costs while helping meet storage performance and capacity needs based on information that's collected during a discovery job.

The storage configuration is only an approximation and might not account for all capabilities provided by an AWS storage service. For more information, see [What's not included in the recommendations?](#)

### Estimated cost

DataSync Discovery provides an estimated monthly cost for each AWS storage service that it recommends. The cost is based on standard AWS pricing and provides only an estimate of your AWS fees. It does not include any taxes that might apply. Your actual fees depend on a variety of factors, including your usage of AWS services.

The estimated cost also doesn't include the one-time or periodic fees for migrating your data to AWS.

## What's not included in the recommendations?

DataSync Discovery won't recommend an AWS storage service that doesn't meet your storage configuration needs.

Additionally, the following AWS storage capabilities currently aren't accounted for when recommendations are determined:

- **Amazon FSx for NetApp ONTAP** – Single-AZ deployments and backup storage
- **Amazon EFS** – EFS One Zone storage classes and backup storage
- **Amazon FSx for Windows File Server** – Single-AZ deployments and backup storage

## Getting recommendations

You can generate AWS storage recommendations after your discovery job completes, when you stop the job, and even sometimes if the job completes but had some issues collecting information from your storage system.

There might be situations when you can't get recommendations (for example, if your discovery job fails). For more information, see [Recommendation statuses](#).

### Tip

Before starting your migration to AWS, review the DataSync Discovery recommendations with your AWS account team.

## Using the AWS CLI

1. Copy the following `describe-discovery-job` command:

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

2. For the `--discovery-job-arn` parameter, specify the Amazon Resource Name (ARN) of the [discovery job](#) that you ran on the storage system.
3. Run the `describe-discovery-job` command.

If your response includes a `Status` that isn't `FAILED`, you can continue. If you see `FAILED`, you must run another discovery job on your storage system to try to generate recommendations.

4. If your discovery job completed successfully, skip this step. Otherwise, do the following to manually generate recommendations:
  - a. Copy the following `generate-recommendations` command:


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

- b. For the `--discovery-job-arn` parameter, specify the ARN of the same discovery job that you specified in Step 2.

- c. For the `--resource-type` parameter, specify `CLUSTER`, `SVM`, or `RESOURCE` depending on the kind of resource you want recommendations on.
  - d. For the `--resource-ids` parameter, specify universally unique identifiers (UUIDs) of the resources that you want recommendations on.
  - e. Run the `generate-recommendations` command.
  - f. Wait until the `RecommendationStatus` element in the response has a `COMPLETED` status, then move to the next step.
5. Copy the following `describe-storage-system-resources` command:

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

6. Specify the following parameters in the command:
- `--discovery-job-arn` – Specify the ARN of the same discovery job that you specified in Step 2.
  - `--resource-type` – Specify the resource type you generated recommendations on (for example, `VOLUME`).
7. Run the `describe-storage-system-resources` command.

 **Note**

In the response, if you don't see `COMPLETED` for `RecommendationStatus`, check the [recommendation statuses](#) for more information. You may need to retry generating recommendations.

In this example response, the `Recommendations` element suggests a couple AWS storage services where you can migrate a specific volume, how you might configure the service, and estimated monthly AWS storage costs.

```
{  
  "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 Discovery statuses

You can check the status of your discovery jobs and whether AWS DataSync Discovery can provide storage recommendations for your AWS migrations.

### Discovery job statuses

Use the following table to understand what's going on with your discovery job.

API status	Description
RUNNING	Your discovery job is running. The job collects data about your on-premises storage system for the duration that you specified.
WARNING	Your discovery job has encountered errors and currently can't collect data. Review the

API status	Description
	Amazon CloudWatch logs and address these issues within 12 hours, or the job will be terminated.
STOPPED	You stopped your discovery job before the job was expected to finish.
COMPLETED	Your discovery job successfully collected all data from your on-premises storage system.
COMPLETED_WITH_ISSUES	There were times during the discovery job when DataSync Discovery couldn't collect data. For details, see your CloudWatch logs.
TERMINATED	Your discovery job was canceled because of unresolved issues and some data wasn't collected. For details, see your CloudWatch logs.
FAILED	Your discovery job encountered issues and couldn't collect data from your on-premises storage system. For details, see your CloudWatch logs.

## Recommendation statuses

Use the following table to understand whether DataSync Discovery recommendations for a specific on-premises storage resource are ready to view.

API status	Description
NONE	You can't generate recommendations yet. Try generating recommendations when your discovery job completes.

API status	Description
NONE	Your discovery job collected enough data for DataSync Discovery to provide recommendations. You may be able to generate recommendations if you stopped the discovery job early or the job completed but had issues with data collection.
IN_PROGRESS	DataSync Discovery is working on your recommendations. How long this takes depends on how many resources you're generating recommendations for. If you're using the console, it may take a few minutes to generate recommendations for a storage resource.
COMPLETED	You can view your recommendations.
FAILED	DataSync Discovery couldn't generate recommendations. You can review your CloudWatch logs to identify the issue and try generating the recommendations again.
NONE	Recommendations aren't available. You may see this status for a failed discovery job or issue with the storage resource.
COMPLETED	DataSync Discovery currently doesn't support an AWS storage service that meets the needs of the storage resource.



# Monitoring AWS DataSync activity

Monitoring is important for maintaining the reliability and performance of your AWS DataSync transfer and storage discovery activities. We recommend that you collect monitoring data so that you can more easily debug errors if one occurs. Before you start monitoring DataSync, however, create a monitoring plan that includes answers to the following questions:

- What are your monitoring goals?
- What resources will you monitor?
- How often will you monitor these resources?
- What monitoring tools will you use?
- Who will perform the monitoring tasks?
- Who should be notified when something goes wrong?

AWS provides various services and tools for monitoring DataSync. You can configure some of these to do the monitoring for you, but some require manual intervention. We recommend that you automate monitoring tasks as much as possible.

## Topics

- [Monitoring your AWS DataSync transfers with task reports](#)
- [Monitoring AWS DataSync with Amazon CloudWatch](#)
- [Logging AWS DataSync API calls with AWS CloudTrail](#)
- [Monitoring AWS DataSync events by using Amazon EventBridge](#)
- [Manual monitoring tools for AWS DataSync](#)

## Monitoring your AWS DataSync transfers with task reports

*Task reports* provide detailed information about what AWS DataSync attempts to transfer, skip, verify, and delete during a [task execution](#).

Task reports are generated in JSON format. You can customize the level of detail in your reports:

- [Summary only task reports](#) give you the necessary details about your task execution, such as how many files transferred and whether DataSync could verify the data integrity of those files.

- [Standard task reports](#) include a summary plus detailed reports that list each file, object, or folder that DataSync attempts to transfer, skip, verify, and delete. With a standard task report, you can also specify the [report level](#) to show only the task execution's errors or its successes and errors.

## Use cases

Here are some situations where task reports can help you monitor and audit your data transfers:

- When migrating millions of files, quickly identify files that DataSync has issues transferring.
- Verify chain-of-custody processes for your files.

## Summary only task reports

A report that's only a summary of a task execution includes the following details:

- The AWS account that ran the task execution
- The source and destination locations
- The total number of files, objects, and folders that were skipped, transferred, verified, and deleted
- The total bytes (logical and physical) that were transferred
- If the task execution was completed, canceled, or encountered an error
- The start and end times (including the total time of the transfer)
- The task's settings (such as bandwidth limits, data integrity verification, and other options for your DataSync transfer)

## Standard task reports

A standard task report includes a [summary](#) of your task execution plus detailed reports of what DataSync attempts to transfer, skip, verify, and delete.

### Topics

- [Report level](#)
- [Transferred reports](#)
- [Skipped reports](#)

- [Verified reports](#)
- [Deleted reports](#)

## Report level

With standard task reports, you can choose one of the following report levels:

- Errors only
- Successes and errors (essentially a list of everything that happened during your task execution)

For example, you might want to see which files DataSync skipped successfully during your transfer and which ones it didn't. Files that DataSync skipped successfully might be ones that you purposely want DataSync to exclude because they already exist in your destination location. However, a skipped error for instance might indicate that DataSync doesn't have the right permissions to read a file.

## Transferred reports

A list of files, objects, and directories that DataSync attempted to transfer during your task execution. A transferred report includes the following details:

- The paths for the transferred data
- What was transferred (content, metadata, or both)
- The metadata, which includes the data type, content size (objects and files only), and more
- The time when an item was transferred
- The object version (if the destination is an Amazon S3 bucket that has versioning enabled)
- If something was overwritten in the destination
- Whether an item transferred successfully

### Note

When moving data between S3 buckets, the prefix that you specify in your [source location](#) can show up in your report (or in Amazon CloudWatch logs), even if that prefix doesn't exist as an object in your destination location. (In the DataSync console, you might also notice this prefix showing up as skipped or verified data.)

## Skipped reports

A list of files, objects, and directories that DataSync discovered in your source location but didn't attempt to transfer. The reasons DataSync skips data can depend on several factors, such as how you configure your task and file permissions. Here are some examples:

- There's a file that exists in your source and destination locations. The file in the source hasn't been modified since the previous task execution. Since you're [only transferring data that has changed](#), DataSync skips that file and doesn't transfer it during your next task execution.
- An object that exists in your source and destination locations changes in your source. When you run your task, DataSync skips this object in your destination because your task doesn't [overwrite data in the destination](#).
- DataSync skips a directory in your source location because it can't read it.

If this happens and isn't expected, check your access permissions and make sure that DataSync can read what was skipped.

A skipped report includes the following details:

- The paths for skipped data
- The time when an item was skipped
- The reason it was skipped
- Whether an item was skipped successfully

### Note

Skipped reports can be large when they include successes and errors, you configure your task to [transfer only the data that has changed](#), and source data already exists in the destination.

## Verified reports

A list of files, objects, and directories that DataSync attempted to verify the integrity of during your task execution. A verified data report includes the following details:

- The paths for verified data

- The time when an item was verified
- The reason for the verification error (if any)
- The source and destination SHA256 checksums (files only)
- Whether an item was successfully verified

### Note

When you configure your task to [verify only the data that's transferred](#), DataSync doesn't verify directories in some situations or files that fail to transfer. In either case, DataSync doesn't include unverified data in this report.

## Deleted reports

A list of files, directories, and objects that were deleted during your task execution. DataSync generates this report only if you [configure your task](#) to delete data in the destination location that isn't in the source. A deleted data report includes the following details:

- The paths for deleted data
- Whether an item was successfully deleted
- The time when an item was deleted

## Example task reports

The level of detail in your task report is up to you. Here are some example transferred data reports with the following configuration:

- **Report type** – Standard
- **Report level** – Successes and errors

### Note

Reports use the ISO-8601 standard for the timestamp format. Times are in UTC and measured in nanoseconds. This behavior differs from how some other task report metrics

are measured. For example, [task execution details](#), such as `TransferDuration` and `VerifyDuration`, are measured in milliseconds.

### Example transferred data report with success status

This report shows that an object named `object1.txt` successfully transferred.

```
{
  "TaskExecutionId": "exec-abcdefgh12345678",
  "Transferred": [{
    "RelativePath": "/object1.txt",
    "SrcMetadata": {
      "Type": "Regular",
      "ContentSize": 6,
      "Mtime": "2022-01-07T16:59:26.136114671Z",
      "Atime": "2022-01-07T16:59:26.136114671Z",
      "Uid": 0,
      "Gid": 0,
      "Mode": "0644"
    },
    "Overwrite": "False",
    "DstS3VersionId": "jTqRtX3jN4J2G8k0sFSGYK1f35KqpAVP",
    "TransferTimestamp": "2022-01-07T16:59:45.747270957Z",
    "TransferType": "CONTENT_AND_METADATA",
    "TransferStatus": "SUCCESS"
  }]
}
```

### Example transferred data report with error status

This report shows that an object named `object1.txt` didn't transfer because of an S3 bucket permissions issue. (If you get an error like this, see [Providing DataSync access to S3 buckets](#).)

```
{
  "TaskExecutionId": "exec-abcdefgh12345678",
  "Transferred": [{
    "RelativePath": "/object1.txt",
    "SrcMetadata": {
      "Type": "Regular",
      "ContentSize": 6,
      "Mtime": "2022-01-07T16:59:26.136114671Z",
```

```
        "Atime": "2022-01-07T16:59:26.136114671Z",
        "Uid": 0,
        "Gid": 0,
        "Mode": "0644"
    },
    "Overwrite": "False",
    "DstS3VersionId": "jTqRtX3jN4J2G8k0sFSGYK1f35KqpAVP",
    "TransferTimestamp": "2022-01-07T16:59:45.747270957Z",
    "TransferType": "CONTENT_AND_METADATA",
    "TransferStatus": "FAILED",
    "FailureReason": "S3 Get Object Failed",
    "FailureCode": 40974
  ]
}
```

## Prerequisites

Before you can create a task report, you must do the following.

### Topics

- [Create an S3 bucket for your task reports](#)
- [Allow DataSync to upload task reports to your S3 bucket](#)

## Create an S3 bucket for your task reports

If you don't already have one, [create an S3 bucket](#) where DataSync can upload your task report. Reports are stored in the S3 Standard storage class.

We recommend the following for this bucket:

- If you're planning to transfer data to an S3 bucket, don't use the same bucket for your task report if you [disable the Keep deleted files option](#). Otherwise, DataSync will delete any previous task reports each time you execute a task since those reports don't exist in your source location.
- To avoid a complex access permissions setup, make sure that your task report bucket is in the same AWS account and Region as your DataSync transfer task.

## Allow DataSync to upload task reports to your S3 bucket

You must configure an AWS Identity and Access Management (IAM) role that allows DataSync to upload a task report to your S3 bucket.

In the DataSync console, you can create an IAM role that in most cases automatically includes the permissions to upload a task report to your bucket. Keep in mind that this automatically generated role might not meet your needs from a least-privilege standpoint. This role also won't work if your bucket is encrypted with a customer managed AWS Key Management Service (AWS KMS) key (SSE-KMS). In these cases, you can create the role manually as long as the role does at least the following:

- [Prevents the cross-service confused deputy problem](#) in the role's trusted entity.

The following full example shows how you can use the `aws:SourceArn` and `aws:SourceAccount` global condition context keys to prevent the confused deputy problem with DataSync.

```
{
  "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:*"
        }
      }
    }
  ]
}
```

- Allows DataSync to upload a task report to your S3 bucket.



The following example does this by including the `s3:PutObject` action only for a specific prefix (`reports/`) in your bucket.

```
{
  "Version": "2012-10-17",
  "Statement": [{
    "Action": [
      "s3:PutObject"
    ],
    "Effect": "Allow",
    "Resource": "arn:aws:s3:::your-task-reports-bucket/reports/*"
  }]
}
```

- If your S3 bucket is encrypted with a customer managed SSE-KMS key, the [key's policy](#) must include the IAM role that DataSync uses to access the bucket.

For more information, see [Accessing S3 buckets using server-side encryption](#).

## Creating a summary only task report

You can configure a task report that includes a [summary only](#) when creating your DataSync task, starting your task, or updating your task.

The following steps show how to configure a summary only task report when creating a task.

### Using the DataSync console

1. Open the AWS DataSync console at <https://console.aws.amazon.com/datasync/>.
2. In the left navigation pane, expand **Data transfer**, then choose **Tasks**, and then choose **Create task**.
3. Configure your task's source and destination locations.

For more information, see [Where can I transfer my data with AWS DataSync?](#)

4. Scroll down to the **Task report** section. For **Report type**, choose **Summary only**.
5. For **S3 bucket for reports**, choose an S3 bucket where you want DataSync to upload your task report.

**Tip**

If you're planning to transfer data to an S3 bucket, don't use the same bucket for your task report if you [disable the Keep deleted files option](#). Otherwise, DataSync will delete any previous task reports each time you execute a task since those reports don't exist in your source location.

6. For **Folder**, enter a prefix to use for your task report when DataSync uploads the report to your S3 bucket (for example, **reports/**).

Make sure to include the appropriate delimiter character at the end of your prefix. This character is usually a forward slash (/). For more information, see [Organizing objects by using prefixes](#) in the *Amazon S3 User Guide*.

7. For **IAM role**, do one of the following:

- Choose **Autogenerate** to have DataSync automatically create an IAM role with the permissions that are required to access the S3 bucket.

If DataSync previously created an IAM role for this S3 bucket, that role is chosen by default.

- Choose a custom IAM role that you created.

In some cases, you might need to create the role yourself. For more information, see [Allow DataSync to upload task reports to your S3 bucket](#).

**Important**

If your S3 bucket is encrypted with a customer managed SSE-KMS key, the key's policy must include the IAM role that DataSync uses to access the bucket. For more information, see [Accessing S3 buckets using server-side encryption](#).

8. Finish creating your task, and then [start the task](#) to begin transferring your data.

When your transfer is complete, you can [view your task report](#).

**Using the AWS CLI**

1. Copy the following create-task AWS Command Line Interface (AWS CLI) command:

```
aws datasync create-task \
  --source-location-arn arn:aws:datasync:us-east-1:123456789012:location/
loc-12345678abcdefgh \
  --destination-location-arn arn:aws:datasync:us-east-1:123456789012:location/
loc-abcdefgh12345678 \
  --task-report-config '{
    "Destination":{
      "S3":{
        "Subdirectory":"reports/",
        "S3BucketArn":"arn:aws:s3:::your-task-reports-bucket",
        "BucketAccessRoleArn":"arn:aws:iam::123456789012:role/bucket-iam-role"
      }
    },
    "OutputType":"SUMMARY_ONLY"
  }'
```

2. For the `--source-location-arn` parameter, specify the Amazon Resource Name (ARN) of the source location in your transfer. Replace `us-east-1` with the appropriate AWS Region, replace `123456789012` with the appropriate AWS account number, and replace `12345678abcdefgh` with the appropriate source location ID.
3. For the `--destination-location-arn` parameter, specify the ARN of the destination location in your transfer. Replace `us-east-1` with the appropriate AWS Region, replace `123456789012` with the appropriate AWS account number, and replace `abcdefgh12345678` with the appropriate destination location ID.
4. For the `--task-report-config` parameter, do the following:
  - `Subdirectory` – Replace `reports/` with the prefix in your S3 bucket where you want DataSync to upload your task reports.

Make sure to include the appropriate delimiter character at the end of your prefix. This character is usually a forward slash (/). For more information, see [Organizing objects by using prefixes](#) in the *Amazon S3 User Guide*.

- `S3BucketArn` – Specify the ARN of the S3 bucket where you want to upload your task report.

#### Tip

If you're planning to transfer data to an S3 bucket, don't use the same bucket for your task report if you [disable the Keep deleted files option](#). Otherwise, DataSync

will delete any previous task reports each time you execute a task since those reports don't exist in your source location.

- `BucketAccessRoleArn` – Specify the IAM role that allows DataSync to upload a task report to your S3 bucket.

For more information, see [Allow DataSync to upload task reports to your S3 bucket](#).

**⚠ Important**

If your S3 bucket is encrypted with a customer managed SSE-KMS key, the key's policy must include the IAM role that DataSync uses to access the bucket. For more information, see [Accessing S3 buckets using server-side encryption](#).

- `OutputType` – Specify `SUMMARY_ONLY`.

For more information, see [Summary only task reports](#).

5. Run the `create-task` command to create your task.

You get a response like the following that shows you the ARN of the task that you created. You will need this ARN to run the `start-task-execution` command.

```
{
  "TaskArn": "arn:aws:datsync:us-east-1:123456789012:task/task-12345678abcdefgh"
}
```

6. Copy the following `start-task-execution` command.

```
aws datasync-task-report start-task-execution \
  --task-arn arn:aws:datsync:us-east-1:123456789012:task/task-12345678abcdefgh
```

7. For the `--task-arn` parameter, specify the ARN of the task that you're starting. Use the ARN that you received from running the `create-task` command.
8. Run the `start-task-execution` command.

When your transfer is complete, you can [view your task report](#).

## Creating a standard task report

You can configure a [standard task report](#) when creating your DataSync task, starting your task, or updating your task.

The following steps show how to configure a standard task report when creating a task.

### Using the DataSync console

1. Open the AWS DataSync console at <https://console.aws.amazon.com/datasync/>.
2. In the left navigation pane, expand **Data transfer**, then choose **Tasks**, and then choose **Create task**.
3. Configure your task's source and destination locations.

For more information, see [Where can I transfer my data with AWS DataSync?](#)

4. Scroll down to the **Task report** section. For **Report type**, choose **Standard report**.
5. For **Report level**, choose one of the following:
  - **Errors only** – Your task report includes only issues with what DataSync tried to transfer, skip, verify, and delete.
  - **Successes and errors** – Your task report includes what DataSync successfully transferred, skipped, verified, and deleted and what it didn't.
  - **Custom** – Allows you to choose whether you want to see errors only or successes and errors for specific aspects of your task report.

For example, you can choose **Successes and errors** for the transferred files list but **Errors only** for the rest of the report.

6. If you're transferring to an S3 bucket that uses object versioning, keep **Include Amazon S3 object versions** selected if you want your report to include the new version for each transferred object.
7. For **S3 bucket for reports**, choose an S3 bucket where you want DataSync to upload your task report.

#### Tip

If you're planning to transfer data to an S3 bucket, don't use the same bucket for your task report if you [disable the Keep deleted files option](#). Otherwise, DataSync will

delete any previous task reports each time you execute a task since those reports don't exist in your source location.

8. For **Folder**, enter a prefix to use for your task report when DataSync uploads the report to your S3 bucket (for example, **reports/**). Make sure to include the appropriate delimiter character at the end of your prefix. This character is usually a forward slash (/). For more information, see [Organizing objects by using prefixes](#) in the *Amazon S3 User Guide*.
9. For **IAM role**, do one of the following:
  - Choose **Autogenerate** to have DataSync automatically create an IAM role with the permissions that are required to access the S3 bucket.

If DataSync previously created an IAM role for this S3 bucket, that role is chosen by default.

- Choose a custom IAM role that you created.

In some cases, you might need to create the role yourself. For more information, see [Allow DataSync to upload task reports to your S3 bucket](#).

#### **Important**

If your S3 bucket is encrypted with a customer managed SSE-KMS key, the key's policy must include the IAM role that DataSync uses to access the bucket. For more information, see [Accessing S3 buckets using server-side encryption](#).

10. Finish creating your task and [start the task](#) to begin transferring your data.

When your transfer is complete, you can [view your task report](#).

## Using the AWS CLI

1. Copy the following create-task command:

```
aws datasync create-task \  
  --source-location-arn arn:aws:datasync:us-east-1:123456789012:location/  
loc-12345678abcdefgh \  
  --destination-location-arn arn:aws:datasync:us-east-1:123456789012:location/  
loc-abcdefgh12345678 \  
  --task-report-config '{  
    "Destination":{
```

```
"S3":{
  "Subdirectory":"reports/",
  "S3BucketArn":"arn:aws:s3:::your-task-reports-bucket",
  "BucketAccessRoleArn":"arn:aws:iam::123456789012:role/bucket-iam-role"
},
"OutputType":"STANDARD",
"ReportLevel":"level-of-detail",
"ObjectVersionIds":"include-or-not"
}'
```

- For the `--source-location-arn` parameter, specify the ARN of the source location in your transfer. Replace `us-east-1` with the appropriate AWS Region, replace `123456789012` with the appropriate AWS account number, and replace `12345678abcdefgh` with the appropriate source location ID.
- For the `--destination-location-arn` parameter, specify the ARN of the destination location in your transfer. Replace `us-east-1` with the appropriate AWS Region, replace `123456789012` with the appropriate AWS account number, and replace `abcdefgh12345678` with the appropriate destination location ID.
- For the `--task-report-config` parameter, do the following:
  - `Subdirectory` – Replace `reports/` with the prefix in your S3 bucket where you want DataSync to upload your task reports. Make sure to include the appropriate delimiter character at the end of your prefix. This character is usually a forward slash (/). For more information, see [Organizing objects by using prefixes](#) in the *Amazon S3 User Guide*.
  - `S3BucketArn` – Specify the ARN of the S3 bucket where you want to upload your task report.

#### Tip

If you're planning to transfer data to an S3 bucket, don't use the same bucket for your task report if you [disable the Keep deleted files option](#). Otherwise, DataSync will delete any previous task reports each time you execute a task since those reports don't exist in your source location.

- `BucketAccessRoleArn` – Specify the IAM role that allows DataSync to upload a task report to your S3 bucket.

For more information, see [Allow DataSync to upload task reports to your S3 bucket](#).

**⚠ Important**

If your S3 bucket is encrypted with a customer managed SSE-KMS key, the key's policy must include the IAM role that DataSync uses to access the bucket.

For more information, see [Accessing S3 buckets using server-side encryption](#).

- `OutputType` – Specify STANDARD report.

For more information, see [Standard task reports](#) Types of task reports.

- (Optional) `ReportLevel` – Specify whether you want `ERRORS_ONLY` (the default) or `SUCCESSES_AND_ERRORS` in your report.
- (Optional) `ObjectVersionIds` – If you're transferring to an S3 bucket that uses object versioning, specify `NONE` if you don't want to include the new version for each transferred object in the report.

By default, this option is set to `INCLUDE`.

- (Optional) `Overrides` – Customize the `ReportLevel` of a particular aspect of your report.

For example, you might want to see `SUCCESSES_AND_ERRORS` for the list of what DataSync deletes in your destination location, but you want `ERRORS_ONLY` for everything else. In this example, you would add the following `Overrides` option to the `--task-report-config` parameter:

```
"Overrides":{
  "Deleted":{
    "ReportLevel":"SUCCESSES_AND_ERRORS"
  }
}
```

If you don't use `Overrides`, your entire report uses the `ReportLevel` that you specify.

5. Run the `create-task` command to create your task.

You get a response like the following that shows you the ARN of the task that you created. You will need this ARN to run the `start-task-execution` command.

```
{
  "TaskArn": "arn:aws:datsync:us-east-1:123456789012:task/task-12345678abcdefgh"
```



```
}
```

6. Copy the following `start-task-execution` command.

```
aws datasync-task-report start-task-execution \  
  --task-arn arn:aws:datasync:us-east-1:123456789012:task/task-12345678abcdefgh
```

7. For the `--task-arn` parameter, specify the ARN of the task you're running. Use the ARN that you received from running the `create-task` command.
8. Run the `start-task-execution` command.

When your transfer is complete, you can [view your task report](#).

## Viewing your task reports

DataSync creates task reports for every task execution. When your execution completes, you can find the related task reports in your S3 bucket. Task reports are organized under prefixes that include the IDs of your tasks and their executions.

To help locate task reports in your S3 bucket, use these examples:

- **Summary only task report** – *reports-prefix/Summary-Reports/task-id-folder/task-execution-id-folder*
- **Standard task report** – *reports-prefix/Detailed-Reports/task-id-folder/task-execution-id-folder*

Because task reports are in JSON format, you have several options for viewing your reports:

- View a report by using [Amazon S3 Select](#).
- Visualize reports by using AWS services such as AWS Glue, Amazon Athena, and Amazon QuickSight. For more information about visualizing your task reports, see the [AWS Storage Blog](#).

## Limitations

- Individual task reports can't exceed 5 MB. If you're copying a large number of files, your task report might be split into multiple reports.

- There are situations when creating task reports can affect the performance of your data transfer. For example, you might notice this when your network connection has high latency and the files you're transferring are small or you're copying only metadata changes.

## Monitoring AWS DataSync with Amazon CloudWatch

You can monitor AWS DataSync by using Amazon CloudWatch, which collects and processes raw data from DataSync into readable, near real-time metrics. These statistics are retained for a period of 15 months.

By default, DataSync metrics data is automatically sent to CloudWatch in 5-minute intervals. For more information, see [What is Amazon CloudWatch?](#) in the *Amazon CloudWatch User Guide*.

### CloudWatch metrics for DataSync

Amazon CloudWatch provides metrics that you can use to get information about DataSync performance and to troubleshoot issues. To see CloudWatch metrics for DataSync, you can use the following tools:

- The CloudWatch console
- The CloudWatch CLI
- The CloudWatch API
- The DataSync console (on the task execution's details page)

For more information, see [Using Amazon CloudWatch metrics](#) in the *Amazon CloudWatch User Guide*.

DataSync metrics use the `aws/datasync` namespace and provide metrics for the following dimensions:

- **AgentId** – The unique ID of the agent.
- **TaskId** – The unique ID of the task. It takes the form of `task-01234567890abcdef`.

The `aws/datasync` namespace includes the following metrics.

Metric	Description
BytesCompressed	<p>The physical number of bytes transferred over the network after compression was applied. In most cases, this number is less than BytesTransferred unless the data isn't compressible.</p> <p>Unit: Bytes</p>
BytesPreparedDestination	<p>The total number of bytes of data that are prepared at the destination location.</p> <p>Unit: Bytes</p>
BytesPreparedSource	<p>The total number of bytes of data that are prepared at the source location.</p> <p>Unit: Bytes</p>
BytesTransferred	<p>The total number of bytes that are involved in the transfer. For the number of bytes sent over the network, see BytesCompressed .</p> <p>Unit: Bytes</p>
BytesVerifiedDestination	<p>The total number of bytes of data that are verified at the destination location.</p> <p>Unit: Bytes</p>
BytesVerifiedSource	<p>The total number of bytes of data that are verified at the source location.</p> <p>Units: Bytes</p>
BytesWritten	<p>The total logical size of all files, objects, and directories that transferred to the destination location.</p> <p>Unit: Bytes</p>
FilesPreparedDestination	<p>The total number of files, objects, and directories that are prepared at the destination location.</p>

Metric	Description
	Unit: Count
FilesPreparedSource	<p>The total number of files, objects, and directories that are prepared at the source location.</p> <p>Unit: Count</p>
FilesTransferred	<p>The actual number of files, objects, directories, and metadata that transferred over the network. This value is calculated and updated on an ongoing basis during the <a href="#">transferring</a> phase of your task execution . It's updated periodically when each piece of data is read from the source location and sent over the network.</p> <p>If failures occur during a transfer, this value can be less than <code>EstimatedFilesToTransfer</code> . This value can also be greater than <code>EstimatedFilesTransferred</code> in some cases. This element is implementation-specific for some location types, so don't use it as an indicator for a correct transfer total or to monitor your task execution.</p> <p>Unit: Count</p>
FilesVerifiedDestination	<p>The total number of files, objects, and directories that are verified at the destination location.</p> <p>Unit: Count</p>
FilesVerifiedSource	<p>The total number of files, objects, and directories that are verified at the source location.</p> <p>Unit: Count</p>

## Allowing DataSync to upload logs to CloudWatch log groups

DataSync requires sufficient permissions to send logs to a CloudWatch log group. When you create a task by using the console, DataSync can many times create an AWS Identity and Access Management (IAM) resource policy with the correct permissions for you.

If you want to use an existing CloudWatch log group or if you want to create your tasks programmatically, you must create this IAM resource policy yourself.

The following example is a resource policy that grants these permissions.

```
{
  "Version": "2012-10-17",
  "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:*:*"
    }
  ]
}
```

The policy uses `Condition` statements to help ensure that only DataSync tasks from the specified account have access to the specified CloudWatch log group. We recommend using the [aws:SourceArn](#) and [aws:SourceAccount](#) global condition context keys in these `Condition` statements to protect against the confused deputy problem. For more information, see [Cross-service confused deputy prevention](#).

To specify the DataSync task or tasks, replace *region* with the Region code for the AWS Region where the tasks are located (for example, `us-west-2`), and replace *account-id* with the AWS account ID of the account that contains the tasks. To specify the CloudWatch log group, replace the

same values. You can also modify the Resource statement to target specific log groups. For more information about using SourceArn and SourceAccount, see [Global condition keys](#) in the *IAM User Guide*.

To apply the policy, save this policy statement to a file on your local computer. Then run the following AWS CLI command to apply the resource policy. To use this example command, replace *full-path-to-policy-file* with the path to the file that contains your policy statement.

```
aws logs put-resource-policy --policy-name trust-datasync --policy-document
file://full-path-to-policy-file
```

### Note

Run this command by using the same AWS account and AWS Region where you activated your DataSync agent.

For more information, see [Working with log groups and log streams](#) in the *Amazon CloudWatch Logs User Guide*.

## Configuring logging for your DataSync transfer task

You can publish details about your DataSync transfer task to a CloudWatch log group.

### Before you begin

DataSync needs permission to upload logs to a CloudWatch log group. You can set up this permission through an IAM resource policy in a couple different ways:

- When you create your task by using the console, DataSync can create a log group and the associated resource policy for you. DataSync can also apply this resource policy for you.
- If you want to use an existing log group, see an [example of how to create a resource policy yourself](#).

### Using the DataSync console

The following instructions describe how to configure CloudWatch logging when creating a task. You also can configure CloudWatch logging when editing a task.

1. Open the AWS DataSync console at <https://console.aws.amazon.com/datasync/>.
2. In the left navigation pane, expand **Data transfer**, then choose **Tasks**, and then choose **Create task**.
3. Configure your task's source and destination locations.

For more information, see [Where can I transfer my data with AWS DataSync?](#)

4. On the **Configure settings** page, give your task a name, configure your task execution, configure your data transfer, set a schedule, and optionally add tags, and configure a task report.
5. Scroll down to the **Logging** section. For **Log level**, choose one of the following options:
  - **Log basic information such as transfer errors** – Publish logs with only basic information (such as transfer errors).
  - **Log all transferred objects and files** – Publish logs for all files or objects that your DataSync task transfers and performs data-integrity checks on.
  - **Do not send logs to CloudWatch**
6. For **CloudWatch log group**, specify a log group that DataSync has permission to upload logs to by doing one of the following:
  - Choose **Autogenerate** to automatically create a log group that allows DataSync to upload logs to it.
  - Choose an existing log group in your current AWS Region.

If you choose an existing log group, make sure that you have a [resource policy](#) that allows DataSync to upload logs to the log group.

## Using the DataSync API

You can configure CloudWatch logging for your task by using the `CloudWatchLogGroupArn` parameter with any of the following operations:

- [CreateTask](#)
- [UpdateTask](#)

## Logging AWS DataSync API calls with AWS CloudTrail

AWS DataSync is integrated with AWS CloudTrail, a service that provides a record of actions taken by a user, role, or an AWS service in DataSync. CloudTrail captures all API calls for DataSync as events. The calls that are captured include calls from the DataSync console and code calls to DataSync API operations.

If you create a trail, you can enable continuous delivery of CloudTrail events to an Amazon S3 bucket, including events for AWS DataSync. If you don't configure a trail, you can still view the most recent events in the CloudTrail console in **Event history**. Using the information collected by CloudTrail, you can determine the request that was made to AWS DataSync, the IP address from which the request was made, who made the request, when it was made, and additional details.

To learn more about CloudTrail, see the [AWS CloudTrail User Guide](#).

### Working with DataSync information in CloudTrail

CloudTrail is enabled on your AWS account when you create the account. When activity occurs in AWS DataSync, that activity is recorded in a CloudTrail event along with other AWS service events in **Event history**. You can view, search, and download recent events in your AWS account. For more information, see [Viewing events with CloudTrail event history](#).

For an ongoing record of events in your AWS account, including events for AWS DataSync, create a trail. A *trail* enables CloudTrail to deliver log files to an Amazon S3 bucket. By default, when you create a trail in the console, the trail applies to all AWS Regions. The trail logs events from all AWS Regions in the same AWS partition and delivers the log files to the Amazon S3 bucket that you specify. Additionally, you can configure other AWS services to further analyze and act upon the event data collected in CloudTrail logs. For more information, see the following:

- [Overview for creating a trail](#)
- [CloudTrail supported services and integrations](#)
- [Configuring Amazon SNS notifications for CloudTrail](#)
- [Receiving CloudTrail log files from multiple Regions](#) and [Receiving CloudTrail log files from multiple accounts](#)

All DataSync actions are logged by CloudTrail. (For more information, see the DataSync [API reference](#).)



For example, calls to the `CreateAgent`, `CreateTask`, and `ListLocations` operations generate entries in the CloudTrail log files.

Every event or log entry contains information about who generated the request. The identity information helps you determine the following:

- Whether the request was made with root or AWS Identity and Access Management (IAM) credentials.
- Whether the request was made with temporary security credentials for a role or federated user.
- Whether the request was made by another AWS service.

For more information, see [CloudTrail userIdentity element](#) in the *AWS CloudTrail User Guide*.

## Understanding DataSync log file entries

A trail is a configuration that enables delivery of events as log files to an Amazon S3 bucket that you specify. CloudTrail log files contain one or more log entries. An event represents a single request from any source and includes information about the requested action, the date and time of the action, the request parameters, and so on. CloudTrail log files aren't an ordered stack trace of the public API calls, so they don't appear in any specific order.

The following example shows a CloudTrail log entry that demonstrates the `CreateTask` operation.

```
{
  "eventVersion": "1.05",
  "userIdentity": {
    "type": "IAMUser",
    "principalId": "1234567890abcdef0",
    "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": "192.0.2.1",
"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-abcdef01234567890",
  "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-021345abcdef6789"
},
"responseElements": {
  "taskArn": "arn:aws:datasync:ap-southeast-1:123456789012:task/
task-1234567890abcdef0"
},
"requestID": "a1b2c3d4-5678-90ab-cdef-EXAMPLE11111",
"eventID": "a1b2c3d4-5678-90ab-cdef-EXAMPLE22222",
"eventType": "AwsApiCall",
"recipientAccountId": "123456789012"
}
```

## Monitoring AWS DataSync events by using Amazon EventBridge

Amazon EventBridge events describe changes in DataSync resources. You can set up rules to match these events and route them to one or more target functions or streams. Events are emitted on a best-effort basis.

## DataSync transfer events

The following EventBridge events are available for DataSync transfers.

### Agent state changes

#### Event

#### Description

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 virtual machine (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

#### Event

#### Description

Adding

DataSync is adding a location.

Available

The location is created and is available to use.

### Task state changes

#### Event

#### Description

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 might see this event when an agent associated with the task goes offline.

## Agent state changes

Queued

Another task is running and using the same agent. DataSync runs tasks in series (first in, first out).

## Task execution state changes

### Event

### Description

Queueing

Another task execution is running and using the same DataSync agent. For more information, see [Running multiple tasks](#).

Launching

DataSync is initializing the task execution. This status usually goes quickly but can take up to a few minutes.

Preparing

DataSync is determining what data needs to transfer. This can take just minutes or a few hours depending on the number of files, objects, or directories in both locations.

Transferring

DataSync is performing the actual data transfer.

Verifying

DataSync is performing a data-integrity check at the end of the transfer.

Success

The task execution succeeded.

Cancelling

The task execution is in the process of being cancelled.

Error

The task execution failed.

## DataSync Discovery events

The following EventBridge events are available for DataSync Discovery.

## Storage system state changes

Event	Description
Storage System Connectivity Status Change	The connection between your DataSync agent and your on-premises storage system changed. For details, see your CloudWatch logs.

## Discovery job state changes

Event	Description
Discovery Job State Change	The status of your discovery job changed. For more information, see <a href="#">discovery job statuses</a> .
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 <a href="#">DescribeStorageSystemResources</a> and <a href="#">DescribeStorageSystemResourceMetrics</a> operations.

## Manual monitoring tools for AWS DataSync

You can track your AWS DataSync transfers from the console or the command line.

### Monitoring your transfer by using the DataSync console

You can monitor your DataSync transfer by using the console, which provides real-time metrics such as data transferred, data and file throughput, and data compression.

#### To monitor your transfer by using the DataSync console

1. After you [start your DataSync task](#), choose **See execution details**.
2. View metrics about your transfer.

## Monitoring your transfer by using the AWS CLI

You can monitor your DataSync transfer by using the AWS Command Line Interface (AWS CLI).

Copy the following `describe-task-execution` command. To use this example command, replace the *user input placeholders* with your own information.

```
aws datasync describe-task-execution \  
  --task-execution-arn 'arn:aws:datasync:region:account-id:task/task-id/execution/task-  
execution-id'
```

This command returns information about a task execution similar to that shown following.

```
{  
  "BytesCompressed": 3500,  
  "BytesTransferred": 5000,  
  "BytesWritten": 5000,  
  "EstimatedBytesToTransfer": 5000,  
  "EstimatedFilesToDelete": 10,  
  "EstimatedFilesToTransfer": 100,  
  "FilesDeleted": 10,  
  "FilesSkipped": 0,  
  "FilesTransferred": 100,  
  "FilesVerified": 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:datasync:us-east-1:111222333444:task/task-
aaaabbbbccccdddf/execution/exec-1234abcd1234abcd1",
  "TaskReportConfig": {
    "Destination": {
      "S3": {
        "BucketAccessRoleArn": "arn:aws:iam::111222333444:role/my-datasync-
role",
        "S3BucketArn": "arn:aws:s3:::DOC-EXAMPLE-BUCKET/*",
        "Subdirectory": "reports"
      }
    },
    "ObjectVersionIds": "INCLUDE",
    "OutputType": "STANDARD",
    "Overrides": {
      "Deleted": {
        "ReportLevel": "ERRORS_ONLY"
      },
      "Skipped": {
        "ReportLevel": "SUCCESSSES_AND_ERRORS"
      },
      "Transferred": {
        "ReportLevel": "ERRORS_ONLY"
      },
      "Verified": {
        "ReportLevel": "ERRORS_ONLY"
      }
    },
    "ReportLevel": "ERRORS_ONLY"
  }
}

```

- If the task execution succeeds, the value of **Status** changes to **SUCCESS**. For information about what the response elements mean, see [DescribeTaskExecution](#).
- If the task execution fails, the result sends error codes that can help you troubleshoot issues. For information about the error codes, see [TaskExecutionResultDetail](#).

## Monitoring your transfer by using the watch utility

To monitor the progress of your task in real time from the command line, you can use the standard Unix watch utility. Task execution duration values are measured in milliseconds.

The watch utility doesn't recognize the DataSync alias. The following example shows how to invoke the CLI directly. To use this example command, replace the *user input placeholders* with your own information.

```
# 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'"
```



# Managing AWS DataSync resources

Learn how to manage your AWS DataSync resources, such as agents, locations, and tasks.

## Managing your DataSync agent

Once you create a DataSync agent, AWS manages the virtual machine (VM) appliance for you (including software updates). [Learn more](#)

## Working with your DataSync agent's local console

While AWS manages your DataSync agent once it's deployed and activated, there might be cases where you need to change your agent's settings or troubleshoot an issue. [Learn more](#)

## Replacing your DataSync agent

To replace a DataSync agent, you must create a new agent and update any locations that are using the old agent. [Learn more](#)

## Cleaning up DataSync resources

If you used DataSync for a test or just no longer need its resources, delete those resources so that you aren't charged for them. [Learn more](#)

## Reusing a DataSync agent's infrastructure

After you delete an agent resource from DataSync, you can still use the agent's virtual machine or Amazon EC2 instance to activate a new agent. [Learn more](#)

## Managing your AWS DataSync agent

Once you create an AWS DataSync agent in your storage environment, AWS manages the virtual machine (VM) appliance for you.

## Agent software updates

AWS automatically updates your agent's software, including the underlying operating system and related DataSync software packages.

DataSync updates your agent only when it's idle. For example, your agent won't be updated until your transfer is complete.

The agent might go offline briefly following updates. This can happen, for instance, shortly after [agent activation](#) when AWS updates the agent.

### Warning

DataSync doesn't support updating an Amazon EC2 agent manually with cloud-init directives. If you update an agent this way, you may encounter interoperability problems with DataSync where you can't activate or use the agent.

## Agent statuses

The following table describes the status of DataSync agents.

Agent status	Meaning
Online	The agent is configured properly and ready to use. This is the normal running status for an agent.
Offline	The agent has been out of contact with DataSync for five minutes or longer. This can happen for a few reasons. For more information, see <a href="#">What do I do if my agent is offline?</a>

## Troubleshooting your agent

While AWS manages the DataSync agent for you, there are situations when you might need to again work directly with it. For example, if your agent goes offline or loses its connection to your on-premises storage system, you can try to resolve these issues in the [agent's local console](#).

For more information, see [troubleshooting DataSync agents](#).

## Working with your AWS DataSync agent's local console

While AWS manages your AWS DataSync agent once it's deployed and activated, there might be cases where you need to change your agent's settings or troubleshoot an issue. Here are some examples of why you'd work with your agent through its local console:

- Manually assign an IP address to the agent.
- Test your agent's connection to AWS or a storage system.
- Provide AWS Support access to your agent to help with an issue (such as a firewall misconfiguration).

### Important

You don't need to use the agent's local console for standard DataSync functionality.

## Accessing the agent's local console

How you access the local console depends on the type of agent you're using.

### Accessing the local console (VMware ESXi, Linux KVM, or Microsoft Hyper-V)

For security reasons, you can't remotely connect to the local console of the DataSync agent virtual machine (VM).

- If this is your first time using the local console, log in with the default credentials. The default user name is **admin** and the password is **password**.

**Note**

We recommend changing the default password. To do this, on the console main menu enter **5** (or **6** for VMware VMs), then run the `passwd` command to change the password.

## Accessing the local console (Amazon EC2)

To connect to an Amazon EC2 agent's local console, you must use SSH.

**Before you begin:** Make sure that your EC2 instance's security group allows access with SSH (TCP port 22).

1. Open a terminal and copy the following `ssh` command:

```
ssh -i /path/key-pair-name.pem instance-user-name@instance-public-ip-address
```

- For `/path/key-pair-name`, specify the path and file name (`.pem`) of the private key required to connect to your instance.
  - For `instance-user-name`, specify `admin`.
  - For `instance-public-ip-address`, specify the public IP address of your instance.
2. Run the `ssh` command to connect to the instance.

Once connected, the main menu of the agent's local console displays.

## Getting an agent activation key

If your agent isn't activated yet, you can obtain its activation key from the local console. This option is displayed only until the agent has been activated.

### To get an activation key for your agent from the local console

1. Log in to your agent's local console.
2. On the **AWS DataSync Activation - Configuration** main menu, enter **0** to get an activation key.
3. Enter the AWS Region that your agent will be activated in.

4. Enter the service endpoint type that your agent will be using. Options include public, Federal Information Processing Standard (FIPS), and virtual private cloud (VPC) with AWS PrivateLink.
5. The activation key is automatically generated and displayed on screen. Select and copy this value.
6. Using the activation key copied from the last step, use the following `create-agent` CLI command to create and activate the agent:

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

On successful activation, this command returns something similar to the following.

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

You can also insert the activation key in the DataSync console by using the agent creation wizard.

After the agent is activated, the console menu displays the **Agent ID** and **AWS Region**. The option for getting an activation key is no longer visible in the console menu.

## Configuring your agent's network settings


The default network configuration for the agent is Dynamic Host Configuration Protocol (DHCP). With DHCP, your agent is automatically assigned an IP address. In some cases, you might need to manually assign your agent's IP as a static IP address, as described following.

### To configure your agent to use static IP addresses

1. Log in to your agent's local console.
2. On the **AWS DataSync Activation - Configuration** main menu, enter **1** to begin configuring your network.
3. On the **Network Configuration** menu, choose one of the following options.

To	Do this
Get information about your network adapter	<p>Enter <b>1</b>.</p> <p>A list of adapter names appears, and you are prompted to enter an adapter name—for example, <b>eth0</b>. If the adapter you specify is in use, the following information about the adapter is displayed:</p> <ul style="list-style-type: none"><li>• Media access control (MAC) address</li><li>• IP address</li><li>• Netmask</li><li>• Agent IP address</li><li>• DHCP enabled status</li></ul> <p>You use the same adapter name when you configure a static IP address (option <b>3</b>) as when you set your agent's default route adapter (option <b>5</b>).</p>
Configure DHCP	<p>Enter <b>2</b>.</p> <p>You are prompted to configure the network interface to use DHCP.</p>

To	Do this
Configure a static IP address for your agent	<p>Enter <b>3</b>.</p> <p>You are prompted to enter the Network adapter name.</p> <div data-bbox="829 510 1507 825" style="border: 1px solid #f08080; border-radius: 10px; padding: 10px;"><p><b>⚠ Important</b></p><p>If your agent has already been activated, you must shut it down and restart it from the DataSync console for the settings to take effect.</p></div>
Reset all your agent's network configuration to DHCP	<p>Enter <b>4</b>.</p> <p>All network interfaces are set to use DHCP.</p> <div data-bbox="829 1192 1507 1507" style="border: 1px solid #f08080; border-radius: 10px; padding: 10px;"><p><b>⚠ Important</b></p><p>If your agent has already been activated, you must shut down and restart your agent from the DataSync console for the settings to take effect.</p></div>
Set your agent's default route adapter	<p>Enter <b>5</b>.</p> <p>The available adapters for your agent are shown, and you are prompted to choose one of the adapters—for example, <b>eth0</b>.</p>

To	Do this
Edit your agent's Domain Name System (DNS) configuration	Enter <b>6</b> . The available adapters of the primary and secondary DNS servers are displayed. You are prompted to provide the new IP address.
View your agent's DNS configuration	Enter <b>7</b> . The available adapters of the primary and secondary DNS servers are displayed. <div data-bbox="829 684 1507 953" style="border: 1px solid #add8e6; border-radius: 10px; padding: 10px; margin-top: 10px;"> <p> <b>Note</b></p> <p>For some versions of the VMware hypervisor, you can edit the adapter configuration in this menu.</p> </div>
View routing tables	Enter <b>8</b> . The default route of your agent is displayed.

## Testing your agent's connection to AWS

You can use your agent's local console to test your internet connection. This test can be useful when you are troubleshooting network issues with your agent.

### To test your agent's connection to AWS DataSync endpoints

1. Log in to your agent's local console.
2. On the **AWS DataSync Activation - Configuration** main menu, enter **2** to begin testing network connectivity.
3. Enter the service endpoint type that your agent is connecting to. Valid endpoint types include public, FIPS, and VPC endpoints that are using AWS PrivateLink.



When the agent is activated, the **Test Network Connectivity** option can be initiated without any additional user input, because the Region and endpoint type are taken from the activated agent information.

- a. To test public endpoint connectivity, enter **1**, followed by the AWS Region in which your agent is activated. Connectivity test results against the correct endpoints for your agent's Region are displayed. For information about AWS Regions and endpoints, see [Where can I use DataSync?](#)

Each endpoint in the selected AWS Region displays either a **PASSED** or **FAILED** message.

- b. To test FIPS endpoint connectivity, enter **2**, followed by the AWS Region in which your agent is activated. Connectivity test results against the correct endpoints for your agent's Region are displayed. For information about AWS Regions and endpoints, see [Where can I use DataSync?](#)

Each endpoint in the selected AWS Region displays either a **PASSED** or **FAILED** message.

- c. To test VPC endpoint connectivity, enter **3**. Network connectivity test results for your agent's VPC endpoints are displayed.

Each VPC endpoint displays either a **PASSED** or **FAILED** message.

For information about network and firewall requirements, see [AWS DataSync network requirements](#).

## Testing your agent's connection to your storage

You can test whether your DataSync agent can connect to the storage involved in your transfer. This test can help verify that you configured your transfer location correctly.

### To test your agent's connection to your storage

1. Log in to your agent's local console.
2. On the **AWS DataSync Activation - Configuration** main menu, enter **3**.
3. Enter one of the following options:
  - a. Enter **1** to test an NFS server connection.
  - b. Enter **2** to test an SMB server connection.

- c. Enter **3** to test an object storage server connection.
  - d. Enter **4** to test an HDFS connection.
  - e. Enter **5** to test a Microsoft Azure Blob Storage connection.
4. Enter the storage server's IP address or domain name.

Remember the following when entering the IP address or domain name:

- Don't include a protocol. For example, enter **mystorage.com** instead of **https://mystorage.com**.
  - For HDFS, enter the IP address or domain name of the NameNode or DataNode in the Hadoop cluster.
5. If requested, enter the TCP port for connecting to the storage server (for example, **80** or **443**).

You'll see if the connectivity test **PASSED** or **FAILED**.

## Checking your agent's system resources

When you log in to your agent console, virtual CPU cores, root volume size, and RAM are automatically checked. If there are any errors or warnings, they're flagged on the console menu display with a banner that provides details about those errors or warnings.

If there are no errors or warnings when the console starts, the menu displays white text. The **View System Resource Check** option will display (0 Errors).

If there are errors or warnings, the console menu displays the number of errors and warnings, in red and yellow respectively, in a banner across the top of the menu. For example, (1 ERROR, 1 WARNING).

### To check your agent's system resources

1. Log in to your agent's local console.
2. On the **AWS DataSync Activation - Configuration** main menu, enter **4** to view the results of the system resource check.

The console displays an **[OK]**, **[WARNING]**, or **[FAIL]** message for each resource as described in the table following.

For Amazon EC2 instances, the system resource check verifies that the instance type is one of the instances recommended for use with DataSync. If the instance type matches that list, a single result is displayed in green text, as follows.

```
[ OK ] Instance Type Check
```

If the Amazon EC2 instance is not on the recommended list, the system resource check verifies the following resources.

- CPU cores check: At least four cores are required.
- Disk size check: A minimum of 80 GB of available disk space is required.
- RAM check:
  - 32 GB of RAM assigned to the instance for task executions working with up to 20 million files, objects, or directories.
  - 64 GB of RAM assigned to the instance for task executions working with more than 20 million files, objects, or directories.
- CPU flags check: The agent VM CPU must have either SSSE3 or SSE4 instruction set flags.

If the Amazon EC2 instance is not on the list of recommended instances for DataSync, but it has sufficient resources, the result of the system resource check displays four results, all in green text.

The same resources are verified for agents deployed in Hyper-V, Linux Kernel-based Virtual Machine (KVM), and VMware VMs.

VMware agents are also checked for supported version; unsupported versions cause a red banner error. Supported versions include VMware versions 6.5 and 6.7.

## Synchronizing the time on your VMware agent

If you are using a VMware VM, you can view Network Time Protocol (NTP) server configurations and synchronize the VM time on your agent with your VMware hypervisor host.

### To manage system time

1. Log in to your agent's local console.

2. On the **AWS DataSync Activation - Configuration** main menu, enter **5** to manage your system's time.
3. On the **System Time Management** menu, enter **1** to view and synchronize the VM system time.

To	Do this
View and synchronize your VM time with NTP server time	<p>Enter <b>1</b>.</p> <p>The current time of your agent is displayed . Your agent determines the time difference between your agent VM and your NTP server time, and prompts you to synchronize the agent time with NTP time.</p> <p>After your agent is deployed and running, in some scenarios the agent's time can drift. For example, suppose that there is a prolonged network outage and your hypervisor host and agent don't get time updates. In this case, the agent's time is different from the true time. When there is a time drift, a discrepancy occurs between the stated times when operations such as snapshots occur and the actual times that the operations occur.</p>
Edit your NTP server configuration	<p>Enter <b>2</b>.</p> <p>You are prompted to provide a preferred and a secondary NTP server.</p>
View your NTP server configuration	<p>Enter <b>3</b>.</p> <p>Your NTP server configuration is displayed.</p>

## Configuring other agent settings

In a DataSync agent's local console, you can perform some maintenance tasks and diagnose issues with your agent.

### To run a configuration or diagnostic command in your agent's local console

1. Log in to your agent's local console.
2. On the **AWS DataSync Activation - Configuration** main menu, enter **5** (or for **6** a VMware VM) for the **Command Prompt**.
3. Use the following commands to perform the following tasks with your agent.

Command	Description
<b>dig</b>	Look up DNS information about the host.
<b>diskclean</b>	Perform disk cleanup.
<b>exit</b>	Return to the console configuration menu.
<b>h</b>	Display a list of available commands.
<b>ifconfig</b>	Display or configure network interfaces.
<b>ip</b>	Display or configure routing, devices, and tunnels.
<b>iptables</b>	Set up and maintain IPv4 packet filtering and network address translation (NAT).
<b>ncport</b>	Test connectivity to a specific network TCP port.
<b>nping</b>	Get information to troubleshoot network issues.
<b>open-support-channel</b>	Connect the agent to AWS Support.
<b>save-iptables</b>	Save IP table firewall rules permanently.

Command	Description
<code>save-routing-table</code>	Save a newly added routing table entry.
<code>sslcheck</code>	Verify whether an SSL certificate is valid.
<code>tcptracert</code>	Collect <code>tracert</code> output on TCP traffic to a destination.

4. Follow the onscreen instructions.

## Getting help with your agent from AWS Support

You can allow AWS Support to access your AWS DataSync agent and assist you with troubleshooting agent issues. By default, AWS Support access to your agent is disabled. You enable this access through the host's local console. To give AWS Support access to DataSync, you first log in to the local console for the host and then connect to the support server.

To log in to an agent running on Amazon EC2, create a rule for the instance's security group that opens TCP port 22 for Secure Shell (SSH) access.

### Note

If you add a new rule to an existing security group, the new rule applies to all instances that use that security group. For more information about security groups and how to add a security group rule, see [Amazon EC2 security groups for Linux instances](#) in the *Amazon EC2 User Guide*.

### To enable AWS Support access to AWS DataSync

1. Log in to your host's local console.

If this is your first time logging in to the local console, see [Accessing the agent's local console](#).


2. At the prompt, enter **5** to open the command prompt (for VMware VMs, use **6**).
3. Enter **h** to open the **AVAILABLE COMMANDS** window.
4. In the **AVAILABLE COMMANDS** window, enter the following to connect to AWS Support:

```
open-support-channel
```

If you are using the agent with VPC endpoints, you must provide a VPC endpoint IP address for your support channel, as follows:

**open-support-channel** *vpc-ip-address*

Your firewall must allow the outbound TCP port 22 to initiate a support channel to AWS. When you connect to AWS Support, DataSync assigns you a support number. Make a note of your support number.

 **Note**

The channel number isn't a Transmission Control Protocol/User Datagram Protocol (TCP/UDP) port number. Instead, it makes an SSH (TCP 22) connection to servers and provides the support channel for the connection.

5. When the support channel is established, provide your support service number to AWS Support so that they can provide troubleshooting assistance.
6. When the support session is finished, press **Enter** to end it.
7. Enter **exit** to log out of the DataSync local console.
8. Follow the prompts to exit the local console.

## Replacing your AWS DataSync agent

To replace an AWS DataSync agent, you must create a new agent and update any transfer locations that are using the old agent.

### Creating a new agent

To create your new DataSync agent, follow the same process when you created your old agent:

1. [Deploy an agent](#) in your storage environment.
2. [Choose a service endpoint](#) that the agent uses to communicate with AWS.
3. [Configure your network](#) so that the agent can communicate with your storage and AWS.
4. [Activate your agent](#).
5. Once activated, make note of the agent's Amazon Resource Name (ARN).

You need this ARN when updating your DataSync location to use the new agent.

## Updating your location with the new agent

Once you create a new agent, you can update an existing DataSync location to use this agent. In most cases, you also have to re-enter access credentials to update the location. This is because DataSync stores location credentials in a way that only your agent can use them.

### Using the DataSync console

The following instructions describe how to update locations with a new agent by using the DataSync console.

#### NFS

1. Open the AWS DataSync console at <https://console.aws.amazon.com/datasync/>.
2. In the left navigation pane, expand **Data transfer**, then choose **Locations**.
3. Choose the location that you want to update, then choose **Edit**.
4. For **Agents**, choose your new agent.

You can choose more than one agent if you're replacing [multiple agents](#) for a location.

5. Choose **Save changes**.

#### SMB

1. Open the AWS DataSync console at <https://console.aws.amazon.com/datasync/>.
2. In the left navigation pane, expand **Data transfer**, then choose **Locations**.
3. Choose the location that you want to update, then choose **Edit**.
4. For **Agents**, choose your new agent.

You can choose more than one agent if you're replacing [multiple agents](#) for a location.

5. For **Password**, enter the password of the user that can mount your SMB file server and has permission to access the files and folders involved in your transfer.
6. Choose **Save changes**.



## HDFS

1. Open the AWS DataSync console at <https://console.aws.amazon.com/datasync/>.
2. In the left navigation pane, expand **Data transfer**, then choose **Locations**.
3. Choose the location that you want to update, then choose **Edit**.
4. For **Agents**, choose your new agent.

You can choose more than one agent if you're replacing [multiple agents](#) for a location.

5. If you're using Kerberos authentication, upload your **Keytab file** and **Kerberos configuration file**.
6. Choose **Save changes**.

## Object storage

1. Open the AWS DataSync console at <https://console.aws.amazon.com/datasync/>.
2. In the left navigation pane, expand **Data transfer**, then choose **Locations**.
3. Choose the location that you want to update, then choose **Edit**.
4. For **Agents**, choose your new agent.

You can choose more than one agent if you're replacing [multiple agents](#) for a location.

5. If your location requires credentials, enter the **Secret key** that allows DataSync to access your object storage bucket.
6. Choose **Save changes**.

## Azure Blob Storage

Do the following to update your Microsoft Azure Blob Storage location:

1. Open the AWS DataSync console at <https://console.aws.amazon.com/datasync/>.
2. In the left navigation pane, expand **Data transfer**, then choose **Locations**.
3. Choose the location that you want to update, then choose **Edit**.
4. For **Agents**, choose your new agent.

You can choose more than one agent if you're replacing [multiple agents](#) for a location.

5. For **SAS token**, enter the [shared access signature \(SAS\) token](#) that allows DataSync to access your blob storage.
6. Choose **Save changes**.

## Using the AWS CLI

The following instructions describe how to update locations with a new agent by using the AWS CLI. (You can also do this by using the [DataSync API](#).)

### NFS

1. Copy the following [update-location-nfs](#) command:

```
aws datasync update-location-nfs \  
  --location-arn datasync-nfs-location-arn \  
  --on-prem-config AgentArns=new-datasync-agent-arn
```

2. For the `--location-arn` parameter, specify the ARN of the NFS location that you're updating.
3. For the `--on-prem-config` parameter's `AgentArns` option, specify the ARN of your new agent.

You can specify more than one ARN if you're replacing [multiple agents](#) for a location.

4. Run the `update-location-nfs` command to update the location.

### SMB

1. Copy the following [update-location-smb](#) command:

```
aws datasync update-location-smb \  
  --location-arn datasync-smb-location-arn \  
  --agent-arns new-datasync-agent-arn \  
  --password smb-file-server-password
```

2. For the `--location-arn` parameter, specify the ARN of the SMB location that you're updating.
3. For the `--agent-arns` parameter, specify the ARN of your new agent.

You can specify more than one ARN if you're replacing [multiple agents](#) for a location.

4. For the `--password` parameter, specify the password of the user that can mount your SMB file server and has permission to access the files and folders involved in your transfer.
5. Run the `update-location-smb` command to update the location.

## HDFS

1. Copy the following [update-location-hdfs](#) command:

```
aws datasync update-location-hdfs \  
  --location-arn datasync-hdfs-location-arn \  
  --agent-arns new-datasync-agent-arn \  
  --kerberos-keytab keytab-file \  
  --kerberos-krb5-conf krb5-conf-file
```

2. For the `--location-arn` parameter, specify the ARN of the HDFS location that you're updating.
3. For the `--agent-arns` parameter, specify the ARN of your new agent.

You can specify more than one ARN if you're replacing [multiple agents](#) for a location.

4. If you're using Kerberos authentication, include the `--kerberos-keytab` and `--kerberos-krb5-conf` parameters:

- For the `--kerberos-keytab` parameter, specify the Kerberos key table (keytab) that contains mappings between the defined Kerberos principal and encrypted keys.

You can specify the keytab file by providing the file's address.

- For the `--kerberos-krb5-conf` parameter, specify the file that contains the configuration for your Kerberos realm.

You can specify the `krb5.conf` file by providing the file's address.

If you're using simple authentication, you don't need to include these Kerberos-related parameters in your command.

5. Run the `update-location-hdfs` command to update the location.

## Object storage

1. Copy the following [update-location-object-storage](#) command:

```
aws datasync update-location-object-storage \
  --location-arn datasync-object-storage-location-arn \
  --agent-arns new-datasync-agent-arn \
  --secret-key bucket-secret-key
```

- For the `--location-arn` parameter, specify the ARN of the object storage location that you're updating.
- For the `--agent-arns` parameter, specify the ARN of your new agent.

You can specify more than one ARN if you're replacing [multiple agents](#) for a location.

- Do the following depending on if your object storage location requires access credentials:
  - If your location requires credentials** – For the `--secret-key` parameter, specify the secret key that allows DataSync to access your object storage bucket.
  - If your location requires credentials** – Specify empty strings for the `--access-key` and `--secret-key` parameters. Here's an example command:

```
aws datasync update-location-object-storage \
  --location-arn arn:aws:datasync:us-east-2:111122223333:location/  
loc-abcdef01234567890 \
  --agent-arns arn:aws:datasync:us-east-2:111122223333:agent/  
agent-1234567890abcdef0 \
  --access-key "" \
  --secret-key ""
```

- Run the `update-location-object-storage` command to update the location.

## Azure Blob Storage

- Copy the following [update-location-azure-blob](#) command:

```
aws datasync update-location-azure-blob \
  --location-arn datasync-azure-blob-storage-location-arn \
  --agent-arns new-datasync-agent-arn \
  --sas-configuration '{
    "Token": "sas-token-for-azure-blob-storage"
  }'
```

2. For the `--location-arn` parameter, specify the ARN of the Azure Blob Storage location that you're updating.
3. For the `--agent-arns` parameter, specify the ARN of your new agent.

You can specify more than one ARN if you're replacing [multiple agents](#) for a location.

4. For the `--sas-configuration` parameter's Token option, specify the [SAS token](#) that allows DataSync to access your blob storage.
5. Run the `update-location-azure-blob` command to update the location.

## Next steps

1. [Delete your old agent](#). If you have any running DataSync tasks using this agent, wait until those tasks finish before deleting it.
2. If you need to replace agents for multiple locations, repeat the previous steps.
3. When you're done, you can resume [running your tasks](#).

### Note

**Replacing agents for scheduled tasks** – If you replace an agent for a [scheduled task](#), you must start that task manually if the new agent is using a different type of [service endpoint](#) than your old agent. If you don't run the task manually before its next scheduled run, the task fails.

For example, if your old agent used a public service endpoint, but the new agent uses a VPC endpoint, start that task manually by using the console or `StartTaskExecution` operation. After that, your task will resume running on its schedule.

## Cleaning up your AWS DataSync resources

If you used AWS DataSync for a test or don't need the AWS resources that you created, delete them so that you aren't charged for resources you don't plan to use.

### Deleting a DataSync agent

When you delete an agent from AWS DataSync, the agent resource is no longer associated with your AWS account and can't be undone.

Keep in mind that deleting an agent from DataSync doesn't remove its virtual machine (VM) or Amazon EC2 instance from your storage environment. You can delete the VM or instance or reuse it to activate a new agent.

## Prerequisites

Don't delete an agent until you update or remove the DataSync resources that depend on it. If you're replacing an agent, [update your transfer locations](#) with the new agent. If you aren't replacing an agent, delete transfer [tasks](#) and [locations](#) using that agent first.

## Deleting the agent

1. Open the AWS DataSync console at <https://console.aws.amazon.com/datasync/>.
2. In the left navigation pane, choose **Agents**.
3. Choose the agent that you want to delete.
4. Choose **Delete**, enter **delete** in the text box that appears, and then choose **Delete**.
5. If you aren't planning to [reuse the agent's infrastructure](#) for other DataSync activities, delete the agent's VM or Amazon EC2 instance to remove it from your storage environment.

## Deleting a DataSync location

As a best practice, remove the AWS DataSync locations that you no longer need.

### To remove a location by using the DataSync console

1. Open the AWS DataSync console at <https://console.aws.amazon.com/datasync/>.
2. In the left navigation pane, expand **Data transfer**, then choose **Locations**.
3. Choose the location that you want to remove.
4. Choose **Delete**. Confirm the deletion by entering **delete**, and then choose **Delete**.

## Deleting a DataSync task

If you no longer need an AWS DataSync task, you can delete it and its related AWS resources.

## Prerequisites

When you run a task, DataSync automatically creates and manages [network interfaces](#) for data transfer traffic. When you delete a task, you also delete its related network interfaces as long as you have the following permissions:

- `ec2:DeleteNetworkInterface`
- `ec2:DescribeNetworkInterfaces`
- `ec2:ModifyNetworkInterfaceAttribute`

These permissions are available in the AWS managed policy `AWSDatasyncFullAccess`. For more information, see [AWS managed policies for AWS DataSync](#).

## Deleting the task

Once you delete a task, you can't restore it.

### Using the DataSync console

1. Open the AWS DataSync console at <https://console.aws.amazon.com/datasync/>.
2. In the left navigation pane, expand **Data transfer**, then choose **Tasks**.
3. Select the task that you want to delete.
4. Choose **Actions**, then choose **Delete**.
5. In the dialog box, choose **Delete**.

### Using the AWS CLI

1. Copy the following `delete-task` command:

```
aws datasync delete-task \  
    --task-arn "task-to-delete"
```

2. For the `--task-arn` parameter, specify the Amazon Resource Name (ARN) of the task you're deleting (for example, `arn:aws:datasync:us-east-2:123456789012:task/task-012345678abcd0123`).
3. Run the `delete-task` command.

## Reusing an AWS DataSync agent's infrastructure

You can delete an agent resource from DataSync and still use the agent's underlying VM or Amazon EC2 instance to activate a new agent.

### To reuse an agent's infrastructure

1. [Test the agent's connection to AWS](#). If the network tests pass, go to the next step.

The network tests must pass before you can move to the next step.

2. [Delete the agent](#) resource from DataSync but don't delete the agent's VM or Amazon EC2 instance.
3. Repeat step 1 to test the agent's connection to AWS again. If the network tests pass, go to the next step.
4. About three minutes after deleting the agent resource from DataSync, check if port 80 is open on the agent VM or Amazon EC2 instance. If it is, go to the next step.
5. [Activate a new agent](#) with the existing VM or Amazon EC2 instance.

You can activate the new agent in a different AWS Region, AWS account, and with another type of [service endpoint](#). If you use a different type of service endpoint, you have to adjust your [network configuration](#).



# Security in AWS DataSync

Cloud security at AWS is the highest priority. As an AWS customer, you benefit from a data center and network architecture that is built to meet the requirements of the most security-sensitive organizations.

Security is a shared responsibility between AWS and you. The [shared responsibility model](#) describes this as security *of* the cloud and security *in* the cloud:

- **Security of the cloud** – AWS is responsible for protecting the infrastructure that runs AWS services in the AWS Cloud. AWS also provides you with services that you can use securely. Third-party auditors regularly test and verify the effectiveness of our security as part of the [AWS compliance programs](#). To learn about the compliance programs that apply to AWS DataSync, see [AWS services in scope by compliance program](#).
- **Security in the cloud** – Your responsibility is determined by the AWS service that you use. You are also responsible for other factors including the sensitivity of your data, your company's requirements, and applicable laws and regulations.

This documentation helps you understand how to apply the shared responsibility model when using DataSync. The following topics show you how to configure DataSync to meet your security and compliance objectives. You also learn how to use other AWS services that help you to monitor and secure your DataSync resources.

## Topics

- [Data protection in AWS DataSync](#)
- [Identity and access management in AWS DataSync](#)
- [Compliance validation for AWS DataSync](#)
- [Resilience in AWS DataSync](#)
- [Infrastructure security in AWS DataSync](#)

## Data protection in AWS DataSync

AWS DataSync securely transfers data between self-managed storage systems and AWS storage services and also between AWS storage services. How your storage data is encrypted in transit depends in part on the locations involved in the transfer.

After the transfer completes, data is encrypted at rest by the system or service that's storing the data (not DataSync).

## Topics

- [AWS DataSync encryption in transit](#)
- [AWS DataSync encryption at rest](#)
- [Internet network traffic privacy](#)

## AWS DataSync encryption in transit

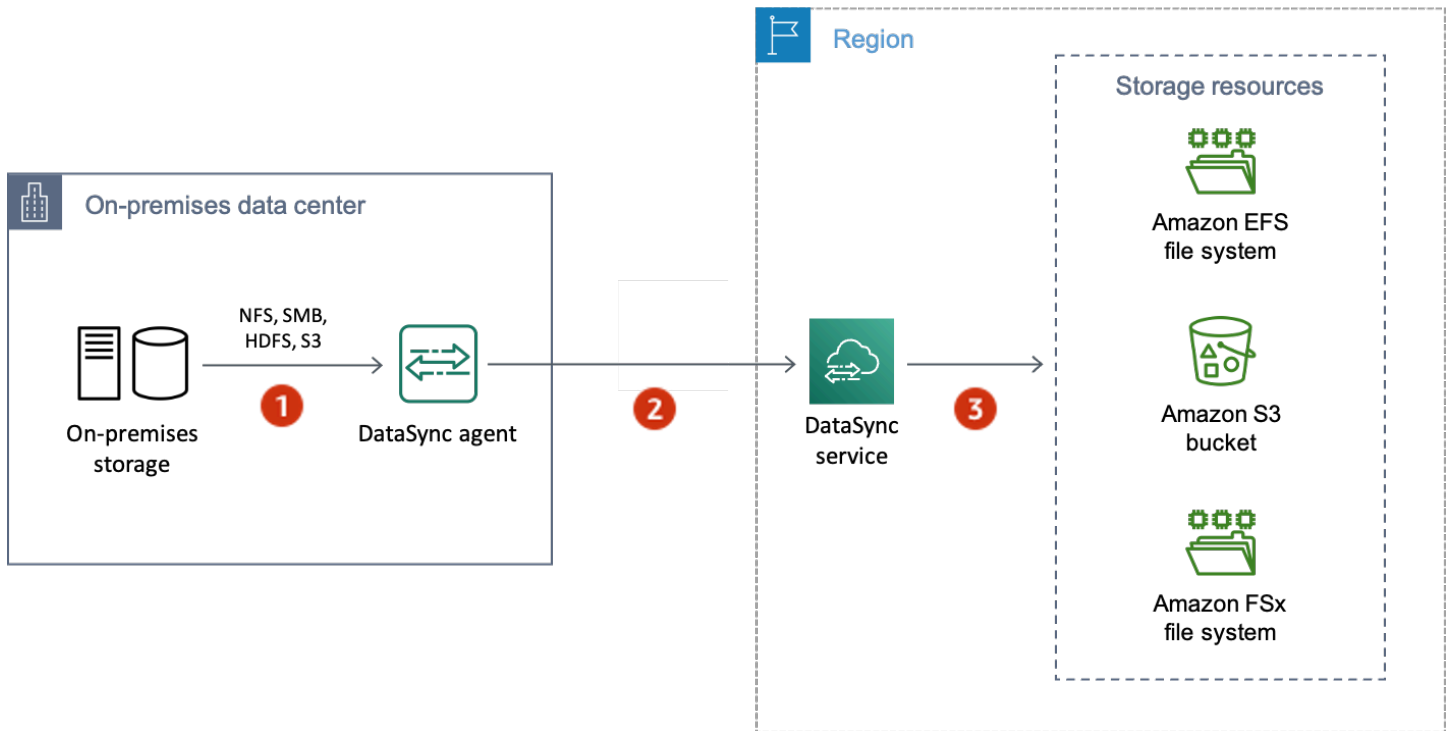
Your storage data (including metadata) is encrypted in transit, but how it's encrypted throughout the transfer depends on your source and destination locations.

When connecting with a location, DataSync uses the most secure options provided by that location's data access protocol. For example, when connecting with a file system using Server Message Block (SMB), DataSync uses the security features provided by SMB.

## Network connections in a transfer

DataSync requires three network connections to copy data: a connection to read data from a source location, another to transfer data between locations, and one more to write data to a destination location.

The following diagram is an example of the network connections that DataSync uses to transfer data from an on-premises storage system to an AWS storage service. To understand where the connections happen and how data is protected as it moves through each connection, use the accompanying table.



Reference	Network connection	Description
1	Reading data from the source location	DataSync connects by using the storage system's protocol for accessing data (for example, SMB or the Amazon S3 API). For this connection, data is protected by using the security features of the storage system.
2	Transferring data between locations	For this connection, DataSync encrypts all network traffic with Transport Layer Security (TLS) 1.3.
3	Writing data to the destination location	Like it did with the source location, DataSync connects by using the storage system's protocol for accessing data. Data is again protected by using the security features of the storage system.

Learn how your data is encrypted in transit when DataSync connects to the following AWS storage services:

- [Amazon EFS](#)
- [Amazon FSx for Windows File Server](#)
- [Amazon FSx for Lustre](#)
- [Amazon FSx for OpenZFS](#)
- [Amazon FSx for NetApp ONTAP](#)
- [Amazon S3](#)

## TLS ciphers

When transferring data between locations, DataSync uses different TLS ciphers. The TLS cipher depends on the type of service endpoint that your agent uses to communicate with DataSync. (For more information, see [Choosing a service endpoint for your AWS DataSync agent.](#))

### Contents

- [Public or VPC endpoints](#)
- [FIPS endpoints](#)

### Public or VPC endpoints

For public and virtual private cloud (VPC) service endpoints, DataSync uses one of the following TLS ciphers:

- TLS\_ECDHE\_RSA\_WITH\_AES\_256\_GCM\_SHA384 (ecdh\_x25519)
- TLS\_ECDHE\_RSA\_WITH\_CHACHA20\_POLY1305\_SHA256 (ecdh\_x25519)
- TLS\_ECDHE\_RSA\_WITH\_AES\_128\_GCM\_SHA256 (ecdh\_x25519)

### FIPS endpoints

For Federal Information Processing Standard (FIPS) service endpoints, DataSync uses the following TLS cipher:

- TLS\_AES\_128\_GCM\_SHA256 (secp256r1)

## AWS DataSync encryption at rest

Because AWS DataSync is a transfer service, it generally doesn't manage your storage data at rest. The storage services and systems that DataSync supports are responsible for protecting data in that state. However, there is some service-related data that DataSync manages at rest.

### What's encrypted?

The only data that DataSync handles at rest relates to the information that it discovers about your on-premises storage system and the details needed to complete your transfer. DataSync stores the following data with full at-rest encryption in Amazon DynamoDB:

- Information collected about your on-premises storage system (if you use DataSync Discovery). This information is also stored with full at-rest encryption in Amazon S3.
- Task configurations (for example, details about the locations in your transfer).
- User credentials that allow your DataSync agent to authenticate with a location. These credentials are encrypted by using your agent's public keys. The agent can decrypt these keys as needed with its private keys.

For more information, see [DynamoDB encryption at rest](#) in the *Amazon DynamoDB Developer Guide*.

### Contents

- [Information collected by DataSync Discovery](#)
- [Key management](#)

### Information collected by DataSync Discovery

DataSync Discovery stores and manages the data that it collects about your on-premises storage system for up to 60 days. You can use Amazon EventBridge to notify you when that expiration date is approaching. For more information, see [DataSync Discovery events](#).

When you remove an on-premises storage system resource from DataSync Discovery, you permanently delete any associated discovery jobs, collected data, and recommendations.

## Key management

You can't manage the encryption keys that DataSync uses to store information in DynamoDB related to running your task. This information includes your task configurations and the credentials that agents use to authenticate with a storage location.

## What's not encrypted?

Though DataSync doesn't control how your storage data is encrypted at rest, we still recommend configuring your locations with the highest level of security that they support. For example, you can encrypt objects with Amazon S3 managed encryption keys (SSE-S3) or AWS Key Management Service (AWS KMS) keys (SSE-KMS).

Learn more about how AWS storage services encrypt data at rest:

- [Amazon S3](#)
- [Amazon EFS](#)
- [Amazon FSx for Windows File Server](#)
- [Amazon FSx for Lustre](#)
- [Amazon FSx for OpenZFS](#)
- [Amazon FSx for NetApp ONTAP](#)

## Internetwork traffic privacy

We recommend configuring your source and destination locations with the highest level of security that each one supports. When connecting to a location, AWS DataSync works with the most secure version of the data access protocol that the storage system uses. Additionally, consider limiting subnet traffic to known protocols and services.

DataSync secures the connection between locations—including between AWS accounts, AWS Regions, and Availability Zones—by using Transport Layer Security (TLS) 1.3.

## Identity and access management in AWS DataSync

AWS uses security credentials to identify you and to grant you access to your AWS resources. You can use features of AWS Identity and Access Management (IAM) to allow other users, services,

and applications to use your AWS resources fully or in a limited way, without sharing your security credentials.

By default, IAM identities (users, groups, and roles) don't have permission to create, view, or modify AWS resources. To allow users, groups, and roles to access AWS DataSync resources and interact with the DataSync console and API, we recommend that you use an IAM policy that grants them permission to use the specific resources and API actions that they will need. You then attach the policy to the IAM identity that requires access. For an overview of the basic elements for a policy, see [Access management for AWS DataSync](#).

## Topics

- [Access management for AWS DataSync](#)
- [AWS managed policies for AWS DataSync](#)
- [IAM customer managed policies for AWS DataSync](#)
- [Using service-linked roles for AWS DataSync](#)
- [Granting permission to tag AWS DataSync resources during creation](#)
- [Cross-service confused deputy prevention](#)
- [DataSync API permissions: Actions and resources](#)

## Access management for AWS DataSync

Every AWS resource is owned by an AWS account. Permissions to create or access a resource are governed by permissions policies. An account administrator can attach permissions policies to AWS Identity and Access Management (IAM) identities. Some services (such as AWS Lambda) also support attaching permissions policies to resources.

### Note

An *account administrator* is a user with administrator privileges in an AWS account. For more information, see [IAM best practices](#) in the *IAM User Guide*.

## Topics

- [DataSync resources and operations](#)
- [Understanding resource ownership](#)
- [Managing access to resources](#)

- [Specifying policy elements: Actions, effects, resources, and principals](#)
- [Specifying conditions in a policy](#)

## DataSync resources and operations

In DataSync, the primary resources are agent, location, task, and task execution.

These resources have unique Amazon Resource Names (ARNs) associated with them, as shown in the following table.

Resource type	ARN format
Agent ARN	arn:aws:datsync: <i>region:account-id</i> :agent/ <i>agent-id</i>
Location ARN	arn:aws:datsync: <i>region:account-id</i> :location/ <i>location-id</i>
Task ARN	arn:aws:datsync: <i>region:account-id</i> :task/ <i>task-id</i>
Task execution ARN	arn:aws:datsync: <i>region:account-id</i> :task/ <i>task-id</i> /execution/ <i>exec-id</i>

To grant permissions for specific API operations, such as creating a task, DataSync defines a set of actions that you can specify in a permissions policy. An API operation can require permissions for more than one action. For a list of all the DataSync API actions and the resources that they apply to, see [DataSync API permissions: Actions and resources](#).

## Understanding resource ownership

A *resource owner* is the AWS account that created the resource. That is, the resource owner is the AWS account of the *principal entity* (for example, an IAM role) which authenticates the request that creates the resource. The following examples illustrate how this behavior works:

- If you use the root account credentials of your AWS account to create a task, your AWS account is the owner of the resource (in DataSync, the resource is the task).



- If you create an IAM roles in your AWS account and grant permissions to the CreateTask action to that user, the user can create a task. However, your AWS account, to which the user belongs, owns the task resource.
- If you create an IAM role in your AWS account with permissions to create a task, anyone who can assume the role can create a task. Your AWS account, to which the role belongs, owns the task resource.

## Managing access to resources

A permissions policy describes who has access to what. The following section explains the available options for creating permissions policies.

### Note

This section discusses using IAM in the context of DataSync. It doesn't provide detailed information about the IAM service. For complete IAM documentation, see [What is IAM?](#) in the *IAM User Guide*. For information about IAM policy syntax and descriptions, see [AWS Identity and Access Management policy reference](#) in the *IAM User Guide*.

Policies attached to an IAM identity are referred to as *identity-based* policies (IAM policies) and policies attached to a resource are referred to as *resource-based* policies. DataSync supports only identity-based policies (IAM policies).

## Topics

- [Identity-based policies](#)
- [Resource-based policies](#)

## Identity-based policies

You can manage DataSync resource access with IAM policies. These policies can help an AWS account administrator do the following with DataSync:

- **Grant permissions to create and manage DataSync resources** – Create an IAM policy that allows an IAM role in your AWS account to create and manage DataSync resources, such as agents, locations, and tasks.

- **Grant permissions to a role in another AWS account or an AWS service** – Create an IAM policy that grants permissions to an IAM role in a different AWS account or an AWS service. For example:
  1. The Account A administrator creates an IAM role and attaches a permissions policy to the role that grants permissions on resources in Account A.
  2. The Account A administrator attaches a trust policy to the role that identifies Account B as the principal who can assume the role.

To grant an AWS service permissions to assume the role, the Account A administrator can specify an AWS service as the principal in the trust policy.

3. The Account B administrator can then delegate permissions to assume the role to any users in Account B. This allows anyone using the role in Account B to create or access resources in Account A.

For more information about using IAM to delegate permissions, see [Access management](#) in the *IAM User Guide*.

The following example policy grants permissions to all `List*` actions on all resources. This action is a read-only action and doesn't allow resource modification.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "AllowAllListActionsOnAllResources",
      "Effect": "Allow",
      "Action": [
        "datasync:List*"
      ],
      "Resource": "*"
    }
  ]
}
```

For more information about using identity-based policies with DataSync, see [AWS managed policies](#) and [customer managed policies](#). For more information about IAM identities, see the [IAM User Guide](#).

## Resource-based policies

Other services, such as Amazon S3, support resource-based permissions policies. For example, you can attach a policy to an Amazon S3 bucket to manage access permissions to that bucket. However, DataSync doesn't support resource-based policies.

## Specifying policy elements: Actions, effects, resources, and principals

For each DataSync resource (see [DataSync API permissions: Actions and resources](#)), the service defines a set of API operations (see [Actions](#)). To grant permissions for these API operations, DataSync defines a set of actions that you can specify in a policy. For example, for the DataSync resource, the following actions are defined: `CreateTask`, `DeleteTask`, and `DescribeTask`. Performing an API operation can require permissions for more than one action.

The following are the most basic policy elements:

- **Resource** – In a policy, you use an Amazon Resource Name (ARN) to identify the resource to which the policy applies. For DataSync resources, you can use the wildcard character ( `*` ) in IAM policies. For more information, see [DataSync resources and operations](#).
- **Action** – You use action keywords to identify resource operations that you want to allow or deny. For example, depending on the specified Effect element, the `datasync:CreateTask` permission allows or denies the user permissions to perform the DataSync `CreateTask` operation.
- **Effect** – You specify the effect when the user requests the specific action—this effect can be either `Allow` or `Deny`. If you don't explicitly grant access to (`Allow`) a resource, access is implicitly denied. You can also explicitly deny access to a resource, which you might do to make sure that a user cannot access it, even if a different policy grants that user access. For more information, see [Authorization](#) in the *IAM User Guide*.
- **Principal** – In identity-based policies (IAM policies), the user that the policy is attached to is the implicit principal. For resource-based policies, you specify the user, account, service, or other entity that you want to receive permissions (applies to resource-based policies only). DataSync doesn't support resource-based policies.

To learn more about IAM policy syntax and descriptions, see [AWS Identity and Access Management policy reference](#) in the *IAM User Guide*.

For a table showing all of the DataSync API actions, see [DataSync API permissions: Actions and resources](#).

## Specifying conditions in a policy

When you grant permissions, you can use the IAM policy language to specify the conditions when a policy should take effect when granting permissions. For example, you might want a policy to be applied only after a specific date. For more information about specifying conditions in policy language, see [Condition](#) in the *IAM User Guide*.

To express conditions, you use predefined condition keys. There are no condition keys specific to DataSync. However, there are AWS wide condition keys that you can use as appropriate. For a complete list of AWS wide keys, see [Available keys](#) in the *IAM User Guide*.

## AWS managed policies for AWS DataSync

To add permissions to users, groups, and roles, it's easier to use AWS managed policies than to write policies yourself. It takes time and expertise to [create IAM customer managed policies](#) that provide your team with only the permissions they need. To get started quickly, you can use our AWS managed policies. These policies cover common use cases and are available in your AWS account. For more information about AWS managed policies, see [AWS managed policies](#) in the *IAM User Guide*.

AWS services maintain and update AWS managed policies. You can't change the permissions in AWS managed policies. Services occasionally add additional permissions to an AWS managed policy to support new features. This type of update affects all identities (users, groups, and roles) where the policy is attached. Services are most likely to update an AWS managed policy when a new feature is launched or when new operations become available. Services do not remove permissions from an AWS managed policy, so policy updates won't break your existing permissions.

Additionally, AWS supports managed policies for job functions that span multiple services. For example, the `ReadOnlyAccess` AWS managed policy provides read-only access to all AWS services and resources. When a service launches a new feature, AWS adds read-only permissions for new operations and resources. For a list and descriptions of job function policies, see [AWS managed policies for job functions](#) in the *IAM User Guide*.

## AWS managed policy: `AWSDataSyncReadOnlyAccess`

You can attach the `AWSDataSyncReadOnlyAccess` policy to your IAM identities.

This policy grants read-only permissions for DataSync.

```
{
  "Version": "2012-10-17",
  "Statement": [{
    "Sid": "DataSyncReadOnlyAccessPermissions",
    "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 managed policy: `AWSDataSyncFullAccess`

You can attach the `AWSDataSyncFullAccess` policy to your IAM identities.

This policy grants administrative permissions for DataSync and is required for AWS Management Console access to the service. `AWSDataSyncFullAccess` provides full access to DataSync API operations and the operations that describe related resources (such as Amazon S3 buckets and Amazon EFS file systems). The policy also grants permissions for Amazon CloudWatch, including creating log groups and creating or updating a resource policy.

```
{
  "Version": "2012-10-17",
```

```

"Statement": [{
  "Sid": "DataSyncFullAccessPermissions",
  "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:ListBucketVersions",
    "s3-outposts:ListAccessPoints",
    "s3-outposts:ListRegionalBuckets"
  ],
  "Resource": "*"
},
{
  "Sid": "DataSyncPassRolePermissions",
  "Effect": "Allow",
  "Action": [
    "iam:PassRole"
  ],
  "Resource": "*",
  "Condition": {
    "StringEquals": {
      "iam:PassedToService": [
        "datasync.amazonaws.com"
      ]
    }
  }
}

```



Change	Description	Date
DataSync started tracking changes	DataSync started tracking changes for its AWS managed policies.	March 1, 2021

## IAM customer managed policies for AWS DataSync

In addition to AWS managed policies, you also can create your own identity-based policies for AWS DataSync and attach them to the AWS Identity and Access Management (IAM) identities that require those permissions. These are known as *customer managed policies*, which are standalone policies that you administer in your own AWS account.

### Important

Before you begin, we recommend that you learn about the basic concepts and options for managing access to your DataSync resources. For more information, see [Access management for AWS DataSync](#).

When creating a customer managed policy, you include statements about DataSync operations that can be used on certain AWS resources. The following example policy has two statements (note the Action and Resource elements in each statement):

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "AllowsSpecifiedActionsOnAllTasks",
      "Effect": "Allow",
      "Action": [
        "datasync:DescribeTask",
      ],
      "Resource": "arn:aws:datasync:us-east-2:111222333444:task/*"
    },
    {
      "Sid": "ListAllTasks",
      "Effect": "Allow",
      "Action": [
```



```
        "datasync:ListTasks"
      ],
      "Resource": "*"
    },
  }
}
```

The policy's statements do the following:

- The first statement grants permissions to perform the `datasync:DescribeTask` action on certain transfer task resources by specifying an Amazon Resource Name (ARN) with a wildcard character (\*).
- The second statement grants permissions to perform the `datasync:ListTasks` action on all tasks by specifying just a wildcard character (\*).

## Examples of customer managed policies

The following example customer managed policies grant permissions for various DataSync operations. The policies work if you're using the AWS Command Line Interface (AWS CLI) or an AWS SDK. To use these policies in the console, you must also use the managed policy `AWSDataSyncFullAccess`.

### Topics

- [Example 1: Create a trust relationship that allows DataSync to access your Amazon S3 bucket](#)
- [Example 2: Allow DataSync to read and write to your Amazon S3 bucket](#)
- [Example 3: Allow DataSync to upload logs to CloudWatch log groups](#)

### Example 1: Create a trust relationship that allows DataSync to access your Amazon S3 bucket

The following is an example of a trust policy that allows DataSync to assume an IAM role. This role allows DataSync to access an Amazon S3 bucket. To prevent the [cross-service confused deputy problem](#), we recommend using the [aws:SourceArn](#) and [aws:SourceAccount](#) global condition context keys in the policy.

```
{
  "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:*"
      }
    }
  }
]
}

```

## Example 2: Allow DataSync to read and write to your Amazon S3 bucket

The following example policy grants DataSync the minimum permissions to read and write data to an S3 bucket that's used as a destination location.

```

{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Action": [
        "s3:GetBucketLocation",
        "s3:ListBucket",
        "s3:ListBucketMultipartUploads"
      ],
      "Effect": "Allow",
      "Resource": "arn:aws:s3:::bucket-name"
    },
    {
      "Action": [
        "s3:AbortMultipartUpload",
        "s3:DeleteObject",
        "s3:GetObject",
        "s3:GetObjectTagging",
        "s3:GetObjectVersion",
        "s3:GetObjectVersionTagging",
        "s3:ListMultipartUploadParts",
        "s3:PutObject",

```

```
        "s3:PutObjectTagging"
      ],
      "Effect": "Allow",
      "Resource": "arn:aws:s3:::bucket-name/*"
    }
  ]
}
```

### Example 3: Allow DataSync to upload logs to CloudWatch log groups

DataSync requires permissions to be able to upload logs to your Amazon CloudWatch log groups. You can use CloudWatch log groups to monitor and debug your tasks.

For an example of an IAM policy that grants such permissions, see [Allowing DataSync to upload logs to CloudWatch log groups](#).

## Using service-linked roles for AWS DataSync

AWS DataSync uses AWS Identity and Access Management (IAM) [service-linked roles](#). A service-linked role is a unique type of IAM role that is linked directly to DataSync. Service-linked roles are predefined by DataSync and include all the permissions that the service requires to call other AWS services on your behalf.

A service-linked role makes setting up DataSync easier because you don't have to manually add the necessary permissions. DataSync defines the permissions of its service-linked roles, and unless defined otherwise, only DataSync can assume its roles. The defined permissions include the trust policy and the permissions policy, and that permissions policy cannot be attached to any other IAM entity.

You can delete a service-linked role only after first deleting their related resources. This protects your DataSync resources because you can't inadvertently remove permission to access the resources.

For information about other services that support service-linked roles, see [AWS Services That Work with IAM](#) and look for the services that have **Yes** in the **Service-linked roles** column. Choose a **Yes** with a link to view the service-linked role documentation for that service.

### Service-linked role permissions for DataSync

DataSync uses the service-linked role named **AWSServiceRoleForDataSyncDiscovery** – Service-linked role used by DataSync Discovery to enable integration with other AWS services.

The `AWSServiceRoleForDataSyncDiscovery` service-linked role trusts the following services to assume the role:

- `discovery-datasync.amazonaws.com`

The role permissions policy named `AWSDataSyncDiscoveryServiceRolePolicy` allows DataSync to complete the following actions on the specified resources:

```
{
  "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:*"  
      ]  
    }  
  ]  
}
```

You must configure permissions to allow an IAM entity (such as a role) to create, edit, or delete a service-linked role. For more information, see [Service-linked role permissions](#) in the *IAM User Guide*.

## Creating a service-linked role for DataSync

You don't need to manually create a service-linked role. When you add a storage system in the AWS Management Console, the AWS CLI, or the AWS API, DataSync creates the service-linked role for you.

If you delete this service-linked role, and then need to create it again, you can use the same process to recreate the role in your account. When you add a storage system, DataSync creates the service-linked role for you again.

You can also use the IAM console to create a service-linked role with the **DataSync Discovery** use case. In the AWS CLI or the AWS API, create a service-linked role with the `discovery-datasync.amazonaws.com` service name. For more information, see [Creating a service-linked role](#) in the *IAM User Guide*. If you delete this service-linked role, you can use this same process to create the role again.

## Editing a service-linked role for DataSync

DataSync does not allow you to edit the `AWSServiceRoleForDataSyncDiscovery` service-linked role. This includes the name of the role because various entities might reference it. However, you can edit the description of the role using IAM. For more information, see [Editing a service-linked role](#) in the *IAM User Guide*.

## Deleting a service-linked role for DataSync

If you no longer need to use a feature or service that requires a service-linked role, we recommend that you delete that role. That way you don't have an unused entity that is not actively monitored or maintained. However, you must clean up the resources for your service-linked role before you can manually delete it.

**Note**

If the DataSync service is using the role when you try to delete the resources, then the deletion might fail. If that happens, wait for a few minutes and try the operation again.

**To delete DataSync resources used by the `AWSServiceRoleForDataSyncDiscovery` role**

1. [Remove the on-premises storage systems](#) that you're using with DataSync Discovery.
2. Delete the service-linked role using IAM.

Use the IAM console, the AWS CLI, or the AWS API to delete the `AWSServiceRoleForDataSyncDiscovery` service-linked role. For more information, see [Deleting a service-linked role](#) in the *IAM User Guide*.

**Supported regions for DataSync service-linked roles**

DataSync supports using service-linked roles in all of the regions where the service is available. For more information, see [DataSync regions and endpoints](#).

**Granting permission to tag AWS DataSync resources during creation**

Some resource-creating AWS DataSync API actions enable you to specify tags when you create the resource. You can use resource tags to implement attribute-based access control (ABAC). For more information, see [What is ABAC for AWS?](#) in the *IAM User Guide*.

To enable users to tag resources on creation, they must have permissions to use the action that creates the resource (such as `datasync:CreateAgent` or `datasync:CreateTask`). If tags are specified in the resource-creating action, users must also have explicit permissions to use the `datasync:TagResource` action.

The `datasync:TagResource` action is only evaluated if tags are applied during the resource-creating action. Therefore, a user that has permissions to create a resource (assuming there are no tagging conditions) doesn't require permissions to use the `datasync:TagResource` action if no tags are specified in the request.

However, if the user attempts to create a resource with tags, the request fails if the user doesn't have permissions to use the `datasync:TagResource` action.

## Example IAM policy statements

Use the following example IAM policy statements to grant TagResource permissions to users creating DataSync resources.

The following statement allows users to tag a DataSync agent when they create the agent.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": "datasync:TagResource",
      "Resource": "arn:aws:datasync:region:account-id:agent/*"
    }
  ]
}
```

The following statement allows users to tag a DataSync location when they create the location.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": "datasync:TagResource",
      "Resource": "arn:aws:datasync:region:account-id:location/*"
    }
  ]
}
```

The following statement allows users to tag a DataSync task when they create the task.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": "datasync:TagResource",
      "Resource": "arn:aws:datasync:region:account-id:task/*"
    }
  ]
}
```

```
}
```

## Cross-service confused deputy prevention

The confused deputy problem is a security issue where an entity that doesn't have permission to perform an action can coerce a more-privileged entity to perform the action. In AWS, cross-service impersonation can result in the confused deputy problem. Cross-service impersonation can occur when one service (the *calling service*) calls another service (the *called service*). The calling service can be manipulated to use its permissions to act on another customer's resources in a way it should not otherwise have permission to access. To prevent this, AWS provides tools that help you protect your data for all services with service principals that have been given access to resources in your account.

We recommend using the [aws:SourceArn](#) and [aws:SourceAccount](#) global condition context keys in resource policies to limit the permissions that AWS DataSync gives another service to the resource. If you use both global condition context keys and the `aws:SourceArn` value contains the account ID, the `aws:SourceAccount` value and the account in the `aws:SourceArn` value must use the same account ID when used in the same policy statement. Use `aws:SourceArn` if you want only one resource to be associated with the cross-service access. Use `aws:SourceAccount` if you want any resource in that account to be associated with the cross-service use.

The value of `aws:SourceArn` must include the DataSync location ARN with which DataSync is allowed to assume the IAM role.

The most effective way to protect against the confused deputy problem is to use the `aws:SourceArn` key with the full ARN of the resource. If you don't know the full ARN or if you're specifying multiple resources, use wildcard characters (\*) for the unknown portions. Here are some examples of how to do this for DataSync:

- To limit the trust policy to an existing DataSync location, include the full location ARN in the policy. DataSync will assume the IAM role only when dealing with that particular location.
- When creating an Amazon S3 location for DataSync, you don't know the location's ARN. In these scenarios, use the following format for the `aws:SourceArn` key: `arn:aws:datsync:us-east-2:123456789012:*`. This format validates the partition (aws), account ID, and Region.

The following full example shows how you can use the `aws:SourceArn` and `aws:SourceAccount` global condition context keys in a trust policy to prevent the confused deputy problem with DataSync.



```
{
  "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:*"
        }
      }
    }
  ]
}
```

For more example policies that show how you can use the `aws:SourceArn` and `aws:SourceAccount` global condition context keys with DataSync, see the following topics:

- [Create a trust relationship that allows DataSync to access your Amazon S3 bucket](#)
- [Configure an IAM role to access your Amazon S3 bucket](#)

## DataSync API permissions: Actions and resources

When creating AWS Identity and Access Management (IAM) policies, this page can help you understand the relationship between AWS DataSync API operations, the corresponding actions that you can grant permissions to perform, and the AWS resources for which you can grant the permissions.

In general, here's how you add DataSync permissions to your policy:

- Specify an action in the `Action` element. The value includes a `datasync:` prefix and the API operation name. For example, `datasync:CreateTask`.
- Specify an AWS resource related to the action in the `Resource` element.

You can also use AWS condition keys in your DataSync policies. For a complete list of AWS keys, see [Available keys](#) in the *IAM User Guide*.

For a list of DataSync resources and their Amazon Resource Name (ARN) formats, see [DataSync resources and operations](#).

## DataSync API operations and corresponding actions

### [AddStorageSystem](#)

**Action:** `datasync:AddStorageSystem`

**Resource:** None

**Actions:**

- `kms:Decrypt`
- `iam:CreateServiceLinkedRole`

**Resource:** \*

**Action:** `secretsmanager:CreateSecret`

**Resource:** `arn:aws:secretsmanager:region:account-id:secret:datasync!*`

### [CancelTaskExecution](#)

**Action:** `datasync:CancelTaskExecution`

**Resource:** `arn:aws:datasync:region:account-id:task/task-id/execution/exec-id`

### [CreateAgent](#)

**Action:** `datasync:CreateAgent`

**Resource:** None

### [CreateLocationAzureBlob](#)

**Action:** `dataSync:CreateLocationAzureBlob`

**Resource:** `arn:aws:datasync:region:account-id:agent/agent-id`

### [CreateLocationEfs](#)

**Action:** datasync:CreateLocationEfs

**Resource:** None

### [CreateLocationFsxLustre](#)

**Action:** datasync:CreateLocationFsxLustre

**Resource:** None

### [CreateLocationFsxOntap](#)

**Action:** datasync:CreateLocationFsxOntap

**Resource:** None

### [CreateLocationFsxOpenZfs](#)

**Action:** datasync:CreateLocationFsxOpenZfs

**Resource:** None

### [CreateLocationFsxWindows](#)

**Action:** datasync:CreateLocationFsxWindows

**Resource:** None

### [CreateLocationHdfs](#)

**Action:** dataSync:CreateLocationHdfs

**Resource:** arn:aws:datasync:*region*:*account-id*:agent/*agent-id*

### [CreateLocationNfs](#)

**Action:** datasync:CreateLocationNfs

**Resource:** arn:aws:datasync:*region*:*account-id*:agent/*agent-id*

### [CreateLocationObjectStorage](#)

**Action:** dataSync:CreateLocationObjectStorage

**Resource:** arn:aws:datasync:*region*:*account-id*:agent/*agent-id*

## [CreateLocationS3](#)

**Action:** `datasync:CreateLocationS3`

**Resource:** `arn:aws:datasync:region:account-id:agent/agent-id` (only for Amazon S3 on Outposts)

## [CreateLocationSmb](#)

**Action:** `datasync:CreateLocationSmb`

**Resource:** `arn:aws:datasync:region:account-id:agent/agent-id`

## [CreateTask](#)

**Action:** `datasync:CreateTask`

**Resources:**

- `arn:aws:datasync:region:account-id:location/source-location-id`
- `arn:aws:datasync:region:account-id:location/destination-location-id`

## [DeleteAgent](#)

**Action:** `datasync>DeleteAgent`

**Resource:** `arn:aws:datasync:region:account-id:agent/agent-id`

## [DeleteLocation](#)

**Action:** `datasync>DeleteLocation`

**Resource:** `arn:aws:datasync:region:account-id:location/location-id`

## [DeleteTask](#)

**Action:** `datasync>DeleteTask`

**Resource:** `arn:aws:datasync:region:account-id:task/task-id`

## [DescribeAgent](#)

**Action:** `datasync:DescribeAgent`

**Resource:** `arn:aws:datasync:region:account-id:agent/agent-id`

### [DescribeDiscoveryJob](#)

**Action:** datasync:DescribeDiscoveryJob

**Resource:** arn:aws:datasync:*region*:*account-id*:system/*storage-system-id*/job/*discovery-job-id*

### [DescribeLocationAzureBlob](#)

**Action:** datasync:DescribeLocationAzureBlob

**Resource:** arn:aws:datasync:*region*:*account-id*:location/*location-id*

### [DescribeLocationEfs](#)

**Action:** datasync:DescribeLocationEfs

**Resource:** arn:aws:datasync:*region*:*account-id*:location/*location-id*

### [DescribeLocationFsxLustre](#)

**Action:** datasync:DescribeLocationFsxLustre

**Resource:** arn:aws:datasync:*region*:*account-id*:location/*location-id*

### [DescribeLocationFsxOntap](#)

**Action:** datasync:DescribeLocationFsxOntap

**Resource:** arn:aws:datasync:*region*:*account-id*:location/*location-id*

### [DescribeLocationFsxOpenZfs](#)

**Action:** datasync:DescribeLocationFsxOpenZfs

**Resource:** arn:aws:datasync:*region*:*account-id*:location/*location-id*

### [DescribeLocationFsxWindows](#)

**Action:** datasync:DescribeLocationFsxWindows

**Resource:** arn:aws:datasync:*region*:*account-id*:location/*location-id*

### [DescribeLocationHdfs](#)

**Action:** datasync:DescribeLocationHdfs

**Resource:** arn:aws:datsync:*region*:*account-id*:location/*location-id*

#### [DescribeLocationNfs](#)

**Action:** datsync:DescribeLocationNfs

**Resource:** arn:aws:datsync:*region*:*account-id*:location/*location-id*

#### [DescribeLocationObjectStorage](#)

**Action:** datsync:DescribeLocationObjectStorage

**Resource:** arn:aws:datsync:*region*:*account-id*:location/*location-id*

#### [DescribeLocationS3](#)

**Action:** datsync:DescribeLocationS3

**Resource:** arn:aws:datsync:*region*:*account-id*:location/*location-id*

#### [DescribeLocationSmb](#)

**Action:** datsync:DescribeLocationSmb

**Resource:** arn:aws:datsync:*region*:*account-id*:location/*location-id*

#### [DescribeStorageSystem](#)

**Action:** datsync:DescribeStorageSystem

**Resource:** arn:aws:datsync:*region*:*account-id*:system/*storage-system-id*

**Action:** secretsmanager:DescribeSecret

**Resource:** arn:aws:secretsmanager:*region*:*account-id*:secret:datsync!\*

#### [DescribeStorageSystemResourceMetrics](#)

**Action:** datsync:DescribeStorageSystemResourceMetrics

**Resource:** arn:aws:datsync:*region*:*account-id*:system/*storage-system-id*/  
job/*discovery-job-id*

#### [DescribeStorageSystemResources](#)

**Action:** datsync:DescribeStorageSystemResources

**Resource:** arn:aws:datsync:*region*:*account-id*:system/*storage-system-id*/job/*discovery-job-id*

### [DescribeTask](#)

**Action:** datsync:DescribeTask

**Resource:** arn:aws:datsync:*region*:*account-id*:task/*task-id*

### [DescribeTaskExecution](#)

**Action:** datsync:DescribeTaskExecution

**Resource:** arn:aws:datsync:*region*:*account-id*:task/*task-id*/execution/*exec-id*

### [GenerateRecommendations](#)

**Action:** datsync:GenerateRecommendations

**Resource:** arn:aws:datsync:*region*:*account-id*:system/*storage-system-id*/job/*discovery-job-id*

### [ListAgents](#)

**Action:** datsync:ListAgents

**Resource:** None

### [ListDiscoveryJobs](#)

**Action:** datsync:ListDiscoveryJobs

**Resource:** arn:aws:datsync:*region*:*account-id*:system/*storage-system-id*

### [ListLocations](#)

**Action:** datsync:ListLocations

**Resource:** None

### [ListTagsForResource](#)

**Action:** datsync:ListTagsForResource

**Resources:**

- `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

**Action:** `datsync:ListTaskExecutions`

**Resource:** `arn:aws:datsync:region:account-id:task/task-id`

### ListTasks

**Action:** `datsync:ListTasks`

**Resource:** None

### RemoveStorageSystem

**Action:** `datsync:RemoveStorageSystem`

**Resource:** `arn:aws:datsync:region:account-id:system/storage-system-id`

**Action:** `secretsmanager>DeleteSecret`

**Resource:** `arn:aws:secretsmanager:region:account-id:secret:datsync!*`

### StartDiscoveryJob

**Action:** `datsync:StartDiscoveryJob`

**Resource:** `arn:aws:datsync:region:account-id:system/storage-system-id/job/discovery-job-id`

### StopDiscoveryJob

**Action:** `datsync:StopDiscoveryJob`

**Resource:** `arn:aws:datsync:region:account-id:system/storage-system-id/job/discovery-job-id`

### StartTaskExecution

**Action:** `datsync:StartTaskExecution`

**Resource:** `arn:aws:datsync:region:account-id:task/task-id`



## [TagResource](#)

**Action:** `datasync:TagResource`

**Resources:**

- `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](#)

**Action:** `datasync:UntagResource`

**Resources:**

- `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](#)

**Action:** `datasync:UpdateAgent`

**Resource:** `arn:aws:datasync:region:account-id:agent/agent-id`

## [UpdateDiscoveryJob](#)

**Action:** `datasync:UpdateDiscoveryJob`

**Resource:** `arn:aws:datasync:region:account-id:system/storage-system-id/job/discovery-job-id`

## [UpdateLocationAzureBlob](#)

**Action:** `datasync:UpdateLocationAzureBlob`

**Resources:**

- `arn:aws:datasync:region:account-id:agent/agent-id`
- `arn:aws:datasync:region:account-id:location/location-id`

## [UpdateLocationHdfs](#)

**Action:** `datasync:UpdateLocationHdfs`

**Resources:**

- `arn:aws:datsync:region:account-id:agent/agent-id`
- `arn:aws:datsync:region:account-id:location/location-id`

[UpdateLocationNfs](#)

**Action:** `datsync:UpdateLocationNfs`

**Resource:** `arn:aws:datsync:region:account-id:location/location-id`

[UpdateLocationObjectStorage](#)

**Action:** `datsync:UpdateLocationObjectStorage`

**Resources:**

- `arn:aws:datsync:region:account-id:agent/agent-id`
- `arn:aws:datsync:region:account-id:location/location-id`

[UpdateLocationSmb](#)

**Action:** `datsync:UpdateLocationSmb`

**Resources:**

- `arn:aws:datsync:region:account-id:agent/agent-id`
- `arn:aws:datsync:region:account-id:location/location-id`

[UpdateStorageSystem](#)

**Action:** `datsync:UpdateStorageSystem`

**Resources:**

- `arn:aws:datsync:region:account-id:agent/agent-id`
- `arn:aws:datsync:region:account-id:system/storage-system-id`

[UpdateTask](#)

**Action:** `datsync:UpdateTask`

**Resource:** `arn:aws:datsync:region:account-id:task/task-id`

[UpdateTaskExecution](#)

**Action:** `datsync:UpdateTaskExecution`

**Resource:** arn:aws:datsync:*region*:*account-id*:task/*task-id*/execution/*exec-id*

## Compliance validation for AWS DataSync

To learn whether an AWS service is within the scope of specific compliance programs, see [AWS services in Scope by Compliance Program](#) and choose the compliance program that you are interested in. For general information, see [AWS Compliance Programs](#).

You can download third-party audit reports using AWS Artifact. For more information, see [Downloading Reports in AWS Artifact](#).

Your compliance responsibility when using AWS services is determined by the sensitivity of your data, your company's compliance objectives, and applicable laws and regulations. AWS provides the following resources to help with compliance:

- [Security and Compliance Quick Start Guides](#) – These deployment guides discuss architectural considerations and provide steps for deploying baseline environments on AWS that are security and compliance focused.
- [Architecting for HIPAA Security and Compliance on Amazon Web Services](#) – This whitepaper describes how companies can use AWS to create HIPAA-eligible applications.

### Note

Not all AWS services are HIPAA eligible. For more information, see the [HIPAA Eligible Services Reference](#).

- [AWS Compliance Resources](#) – This collection of workbooks and guides might apply to your industry and location.
- [AWS Customer Compliance Guides](#) – Understand the shared responsibility model through the lens of compliance. The guides summarize the best practices for securing AWS services and map the guidance to security controls across multiple frameworks (including National Institute of Standards and Technology (NIST), Payment Card Industry Security Standards Council (PCI), and International Organization for Standardization (ISO)).
- [Evaluating Resources with Rules](#) in the *AWS Config Developer Guide* – The AWS Config service assesses how well your resource configurations comply with internal practices, industry guidelines, and regulations.

- [AWS Security Hub](#) – This AWS service provides a comprehensive view of your security state within AWS. Security Hub uses security controls to evaluate your AWS resources and to check your compliance against security industry standards and best practices. For a list of supported services and controls, see [Security Hub controls reference](#).
- [Amazon GuardDuty](#) – This AWS service detects potential threats to your AWS accounts, workloads, containers, and data by monitoring your environment for suspicious and malicious activities. GuardDuty can help you address various compliance requirements, like PCI DSS, by meeting intrusion detection requirements mandated by certain compliance frameworks.
- [AWS Audit Manager](#) – This AWS service helps you continuously audit your AWS usage to simplify how you manage risk and compliance with regulations and industry standards.

## Resilience in AWS DataSync

The AWS global infrastructure is built around AWS Regions and Availability Zones. AWS Regions provide multiple physically separated and isolated Availability Zones, which are connected with low-latency, high-throughput, and highly redundant networking. With Availability Zones, you can design and operate applications and databases that automatically fail over between Availability Zones without interruption. Availability Zones are more highly available, fault tolerant, and scalable than traditional single or multiple data center infrastructures.

### Note

If an Availability Zone you're migrating data to or from does fail while you're running a DataSync task, the task also will fail.

For more information about AWS Regions and Availability Zones, see [AWS global infrastructure](#).

## Infrastructure security in AWS DataSync

As a managed service, AWS DataSync is protected by AWS global network security. For information about AWS security services and how AWS protects infrastructure, see [AWS Cloud Security](#). To design your AWS environment using the best practices for infrastructure security, see [Infrastructure Protection](#) in *Security Pillar AWS Well-Architected Framework*.

You use AWS published API calls to access DataSync through the network. Clients must support the following:

- Transport Layer Security (TLS). We require TLS 1.2 and recommend TLS 1.3.
- Cipher suites with perfect forward secrecy (PFS) such as DHE (Ephemeral Diffie-Hellman) or ECDHE (Elliptic Curve Ephemeral Diffie-Hellman). Most modern systems such as Java 7 and later support these modes.

Additionally, requests must be signed by using an access key ID and a secret access key that is associated with an IAM principal. Or you can use the [AWS Security Token Service](#) (AWS STS) to generate temporary security credentials to sign requests.

# AWS DataSync quotas

Find out about quotas and limits when working with AWS DataSync.

## Storage system, file, and object limits

The following table describes the limits that DataSync has when working with storage systems, files, and objects.

Description	Limit
Maximum total file path length	4,096 bytes
Maximum file path component (file name, directory, or subdirectory) length	255 bytes
Maximum length of Windows domain	253 characters
Maximum length of server hostname	255 characters
Maximum Amazon S3 object name length	1,024 UTF-8 characters

## DataSync task quotas

The following table describes the quotas for DataSync tasks in a specific AWS account and AWS Region.

Resource	Quota	Can you increase the quota?
Maximum number of tasks you can create	100	Yes
Maximum number of source and destination files, objects, or directories that DataSync can work with per task execution between	50 million	Yes

Resource	Quota	Can you increase the quota?
<p>on-premises, self-managed, or other cloud storage and AWS storage services</p> <p>For more information, see <a href="#">How DataSync transfers files, objects, and directories</a></p>	<p><b>⚠ Important</b></p> <p>Remember the following about this quota:</p> <ul style="list-style-type: none"> <li>• If you transfer Amazon S3 objects with prefixes, the prefixes are treated as directories and count towards the quota. For example, DataSync would consider <code>s3://bucket/foo/bar.txt</code> as two directories (<code>./</code> and <code>./foo/</code>) and one object (<code>bar.txt</code>).</li> <li>• If your task is working with more than 20 million files, objects, or directories, make sure that you allocate a minimum of 64 GB of RAM to your DataSync agent. For more information, see <a href="#">agent requirements</a></li> </ul>	<p><b>💡 Tip</b></p> <p>Instead of requesting an increase, you can create tasks that focus on specific directories using include and exclude filters. For more information, see <a href="#">filtering the data transfer</a></p>

Resource	Quota	Can you increase the quota?
	<a href="#">Limits for DataSync transfers.</a>	<a href="#">Increased by DataSync.</a>



Resource	Quota	Can you increase the quota?
<p>Maximum number of source and destination files, objects, or directories that DataSync can work with per task execution between AWS storage services</p> <p>For more information, see <a href="#">How DataSync transfers files, objects, and directories</a></p>	<p>25 million</p> <div data-bbox="829 401 1269 1146" style="border: 1px solid #f08080; padding: 10px; margin-top: 10px;"> <p><b>⚠ Important</b></p> <p>If you transfer Amazon S3 objects with prefixes, the prefixes are treated as directories and count towards the quota. For example, DataSync would consider <code>s3://bucket/foo/bar.txt</code> as two directories (<code>./</code> and <code>./foo/</code>) and one object (<code>bar.txt</code>).</p> </div>	<p>Yes</p> <div data-bbox="1307 401 1520 1862" style="border: 1px solid #add8e6; padding: 10px; margin-top: 10px;"> <p><b>💡 Tip</b></p> <p>Instead of requesting an increase, you can create tasks that focus on specific directories using include and exclude filters. For more information, see <a href="#">filtering the data transferr</a></p> </div>

Resource	Quota	Can you increase the quota?
		<a href="#">ed by DataSync.</a>
Maximum number of files, objects, or directories per task execution when running DataSync on an AWS Snowcone device	200,000	No
Maximum throughput per task	10 Gbps (for transfers that use a DataSync agent)	No
	5 Gbps (for transfers that don't use a DataSync agent)	No
Maximum number of characters you can include in a task filter	102,400 characters	No
	<div data-bbox="829 978 1268 1388" style="border: 1px solid #add8e6; border-radius: 10px; padding: 10px;"> <p><b>Note</b></p> <p>If you're using the DataSync console, this limit includes all the characters combined in your include and exclude patterns.</p> </div>	
Maximum number of queued executions for a task	50	No
Maximum number of days a task execution's history is retained	30	No

## DataSync Discovery quotas

The following table describes the quotas for DataSync Discovery in a specific AWS account and AWS Region.

Resource	Quota	Can you increase the quota?
Maximum number of storage systems you can use with DataSync Discovery	10	No
Maximum number of storage systems a DataSync agent can access at a time	4	No

## Request a quota increase

You can request an increase for some DataSync quotas. Increases aren't granted right away and might take a couple of days to take effect.

### To request a quota increase

1. Open the Service Quotas console at <https://console.aws.amazon.com/servicequotas/>.
2. In the navigation pane, choose **AWS services** and then choose **AWS DataSync**.
3. Choose the quota that you want to increase, then choose **Request increase at account-level**.
4. Enter the total amount that you want the quota to be, then choose **Request**.

If you need to increase a different quota, fill out a separate request.

# Troubleshooting AWS DataSync issues

Use the following information to troubleshoot AWS DataSync issues and errors.

## Topics

- [Troubleshooting issues with AWS DataSync agents](#)
- [Troubleshooting issues with AWS DataSync transfers](#)
- [My Amazon S3 storage costs are higher than expected with AWS DataSync](#)

## Troubleshooting issues with AWS DataSync agents

The following topics describe issues common to AWS DataSync agents and how you can resolve them.

### How do I connect to an Amazon EC2 agent's local console?

To connect to an Amazon EC2 agent's local console, you must use SSH. Make sure that your EC2 instance's security group allows access with SSH (TCP port 22).

In a terminal, run the following ssh command to connect to the instance:

```
ssh -i /path/key-pair-name.pem instance-user-name@instance-public-ip-address
```

- For */path/key-pair-name*, specify the path and file name (.pem) of the private key required to connect to your instance.
- For *instance-user-name*, specify admin.
- For *instance-public-ip-address*, specify the public IP address of your instance.

### What does the Failed to retrieve agent activation key error mean?

When activating your DataSync agent, the agent connects to the service endpoint that you specify to request an activation key. This error likely means that your network security settings are blocking the connection.

#### Action to take

If you're using a virtual private cloud (VPC) service endpoint, verify that your security group settings allow your agent to connect to the VPC endpoint. For information about required ports, see [Network requirements for VPC service endpoints](#).

If you're using a public Federal Information Processing Standard (FIPS) endpoint, check that your firewall and router settings allow your agent to connect to the endpoint. For information, see [Network requirements for public or FIPS service endpoints](#).

## I still can't activate an agent by using a VPC service endpoint

If you're still having issues activating a DataSync agent with a VPC service endpoint, see [Getting help with your agent from AWS Support](#).

## What do I do if my agent is offline?

Your DataSync agent can be offline for a few reasons, but you might be able to get it back online. Before you delete the agent and create a new one, go through the following checklist to help you understand what might have happened.

- **Contact your backup team** – If your agent is offline because its virtual machine (VM) was restored from a snapshot or backup, you might need to [replace the agent](#).
- **Check if the agent's VM or Amazon EC2 instance is off** – Depending on the type of agent that you're using, try turning the VM or EC2 instance back on if it's off. Once it's on again, [test your agent's network connectivity](#) to AWS.
- **Verify your agent meets the minimum hardware requirements** – Your agent might be offline because its VM or EC2 instance configuration was accidentally changed since the agent was activated. For example, if your VM no longer has the minimum required memory or space, the agent might appear as offline. For more information, see [Requirements for AWS DataSync agents](#).
- **Wait for agent-related software updates to finish** – Your agent might go offline briefly following [software updates provided by AWS](#). If you believe this is why the agent is offline, wait a short period then check if the agent is back online.
- **Check your VPC service endpoint settings** – If your offline agent is using a public service endpoint and also in the same VPC where you created a VPC service endpoint for DataSync, you might need to disable [private DNS support](#) for that VPC endpoint.

If none of these seem to be the reason that the agent is offline, you likely need to [replace the agent](#).

## I don't know what's going on with my agent. Can someone help me?

If you're having issues with your DataSync agent that you can't solve, see [Getting help with your agent from AWS Support](#).

## Troubleshooting issues with AWS DataSync transfers

The following topics describe issues common to AWS DataSync locations and tasks and how you can resolve them.

### How do I configure DataSync to use a specific NFS or SMB version to mount my file share?

For locations that support Network File System (NFS) or Server Message Block (SMB), DataSync by default chooses the protocol version for you. You can also specify the version yourself by using the DataSync console or API.

#### Action to take (DataSync console)

When creating your NFS or SMB location, configure the protocol version that you want DataSync to use. For more information, see [Configuring AWS DataSync transfers with an NFS file server](#) or [Configuring AWS DataSync transfers with an SMB file server](#).

#### Action to take (DataSync API)

When creating or updating your NFS or SMB location, specify the `Version` parameter. For example, see [CreateLocationNfs](#) or [CreateLocationSmb](#).

The following example AWS CLI command creates an NFS location that DataSync mounts by using NFS version 4.0.

```
$ 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"
```

The following example AWS CLI command creates an SMB location that DataSync mounts by using SMB version 3.

```
$ 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"
```

## Error: Invalid SyncOption value. Option: TransferMode,PreserveDeletedFiles, Value: ALL,REMOVE.

This error occurs when you're creating or editing your DataSync task and you select the **Transfer all data** option and deselect the **Keep deleted files** option. When you transfer all data, DataSync doesn't scan your destination location and doesn't know what to delete.

## My task keeps failing with an EniNotFound error

This error occurs if you delete one of your task's network interfaces in your virtual private cloud (VPC). If your task is scheduled or queued, the task will fail if it's missing a [network interface required to transfer your data](#).

### Actions to take

You have the following options to work around this issue:

- Manually restart the task. When you do this, DataSync will create any missing network interfaces it needs to run the task.
- If you need to clean up resources in your VPC, make sure that you don't delete network interfaces related to a DataSync task that you're still using.

To see the network interfaces allocated to your task, do one of the following:

- Use the [DescribeTask](#) operation. You can view the network interfaces in the SourceNetworkInterfaceArns and DestinationNetworkInterfaceArns response elements.
- In the Amazon EC2 console, search for your task ID (such as task-f012345678abcdef0) to find its network interfaces.
- Consider not running your tasks automatically. This could include disabling task queuing or scheduling (through DataSync or custom automation).

## My task failed with an NFS permissions denied error

You can get a "permissions denied" error message if you configure your NFS file server with `root_squash` or `all_squash` and your files don't all have read access.

### Action to take

To fix this issue, configure your NFS export with `no_root_squash` or make sure that the permissions for all of the files that you want to transfer allow read access for all users.

For DataSync to access directories, you must also enable all-execute access. To make sure that the directory can be mounted, first connect to any computer that has the same network configuration as your agent. Then run the following CLI command:

```
mount -t nfs -o nfsvers=<your-nfs-server-version> <your-nfs-server-name>:<nfs-export-path-you-specified> <new-test-folder-on-your-computer>
```

If the issue still isn't resolved, contact [AWS Support Center](#).

## My task failed with an NFS mount error

You might see the following error when running a DataSync task that involves an NFS file server location:

```
Task failed to access location loc-1111222233334444a: x40016: mount.nfs: Connection timed out
```

### Actions to take

Do the following until the error is resolved.

1. Make sure that the NFS file server and export that you specify in your DataSync location are valid. If they aren't, delete your location and task, then create a new location and task that use a valid NFS file server and export. For more information, see [Using the DataSync console](#).
2. Check your firewall configuration between your agent and NFS file server. For more information, see [Network requirements for on-premises, self-managed, other cloud, and edge storage](#).
3. Make sure that your agent can access the NFS file server and mount the export. For more information, see [Providing DataSync access to NFS file servers](#).
4. If you still see the error, open a support channel with AWS Support. For more information, see [Getting help with your agent from AWS Support](#).



## My task failed with an Amazon EFS mount error

You might see the following error when running a DataSync task that involves an Amazon EFS location:

Task failed to access location loc-1111222233334444a: x40016: Failed to connect to EFS mount target with IP: 10.10.1.0.

This can happen if the Amazon EFS file system's mount path that you configure with your location gets updated or deleted. DataSync isn't aware of these changes in the file system.

### Action to take

Delete your location and task and [create a new Amazon EFS location](#) with the new mount path.

## File ownership isn't maintained with NFS transfer

After your transfer, you might notice that the files in your DataSync destination location have different user IDs (UIDs) or group IDs (GIDs) than the same files in your source location. For example, the files in your destination might have a UID of 65534, 99, or nobody.

This can happen if a file system involved in your transfer uses NFS version 4 ID mapping, a feature that DataSync doesn't support.

### Action to take

You have a couple options to work around this issue:

- Create a new location for the file system that uses NFS version 3 instead of version 4.
- Disable NFS version 4 ID mapping on the file system.

Retry the transfer. Either option should resolve the issue.

## My task failed with a Cannot allocate memory error

When your DataSync task fails with a Cannot allocate memory error, it can mean a few different things.

### Action to take

Try the following until you no longer see the issue:

- If your transfer involves an agent, make sure that the agent meets the [virtual machine \(VM\)](#) or [Amazon EC2 instance](#) requirements.
- Split your transfer into multiple tasks by using [filters](#). It's possible that you're trying to transfer more files or objects than what [one DataSync task can handle](#).
- If you still see the issue, [contact AWS Support](#).

## My task failed with an input/output error

You can get an input/output error message if your storage system fails I/O requests from the DataSync agent. Common reasons for this include a server disk failure, changes to your firewall configuration, or a network router failure.

If the error involves an NFS server or Hadoop Distributed File System (HDFS) cluster, use the following steps to resolve the error.

### Action to take (NFS)

First, check your NFS server's logs and metrics to determine if the problem started on the NFS server. If yes, resolve that issue.

Next, check that your network configuration hasn't changed. To check if the NFS server is configured correctly and that DataSync can access it, do the following:

1. Set up another NFS client on the same network subnet as the agent.
2. Mount your share on that client.
3. Validate that the client can read and write to the share successfully.

### Action to take (HDFS)

Make sure that your HDFS cluster allows the agent to communicate with the cluster's NameNode and DataNode ports. In most clusters, you can find the port numbers that the cluster uses in the following configuration files.

1. To find the NameNode port, look in the `core-site.xml` file under the `fs.default` or `fs.default.name` property (depending on the Hadoop distribution).

2. To find the DataNode port, look in the `hdfs-site.xml` file under the `dfs.datanode.address` property.

## My task execution has a launching status but nothing seems to be happening

Your DataSync task can get stuck with a **Launching** status typically because the agent is powered off or has lost network connectivity.

### Action to take

Make sure that your agent's status is **ONLINE**. If the agent is **OFFLINE**, make sure it's powered on.

If the agent is powered on and the task is still **Launching**, then there's likely a network connection problem between your agent and AWS. For information about how to test network connectivity, see [Testing your agent's connection to AWS](#).

If you're still having this issue, see [Getting help with your agent from AWS Support](#).

## My task execution seems stuck in the preparing status

The time your DataSync transfer task has the **Preparing** status depends on the amount of data in your transfer source and destination and the performance of those storage systems.

When a task starts, DataSync performs a recursive directory listing to discover all files, objects, directories, and metadata in your source and destination. DataSync uses these listings to identify differences between storage systems and determine what to copy. This process can take a few minutes or even a few hours.

### Action to take

You shouldn't have to do anything. Continue to wait for the task status to change to **Transferring**. If the status still doesn't change, contact [AWS Support Center](#).

## How long does it take DataSync to verify a task I've run?

By default, DataSync verifies data integrity at the end of a transfer. How long verification takes depends on a number of factors. The number of files or objects, the total amount of data in the source and destination storage systems, and the performance of these systems affect how long

verification takes. Verification includes an SHA256 checksum on all file content and an exact comparison of all file metadata.

### Action to take

You shouldn't have to do anything. If task status still doesn't change to **Success** or **Error**, contact [AWS Support Center](#).

## My task fails when transferring from a Google Cloud Storage bucket

Because DataSync communicates with Google Cloud Storage by using the Amazon S3 API, there's a limitation that might cause your DataSync transfer to fail if you try to copy object tags. The following message related to the issue appears in your CloudWatch logs:

```
[WARN] Failed to read metadata for file /your-bucket/your-object: S3 Get Object Tagging Failed: proceeding without tagging
```

To prevent this, deselect the **Copy object tags** option when configuring your transfer task settings.

## My task fails with an Unable to list Azure Blobs on the volume root error

If your DataSync transfer task fails with an `Unable to list Azure Blobs on the volume root` error, there might be an issue with your shared access signature (SAS) token or your Azure storage account's network.

### Actions to take

Try the following and run your task again until you fix the issue:

- Make sure that your [SAS token](#) has the right permissions to access your Microsoft Azure Blob Storage.
- If you're running your DataSync agent in Azure, configure your storage account to allow access from the virtual network where your agent resides.
- If you're running your agent on Amazon EC2, configure your Azure storage firewall to allow access from the agent's public IP address.

For information on how to configure your Azure storage account's network, see the [Azure Blob Storage documentation](#).

## My task's start and end times don't match up with what's in the logs

Your task execution's start and end times that you see in the DataSync console may differ between timestamps you see elsewhere related to your transfer. This is because the console doesn't take into account the time a task execution spends in the launching or queueing states.

For example, your Amazon CloudWatch logs can indicate that your task execution ended later than what's displayed in the DataSync console. You may notice a similar discrepancy in the following areas:

- Logs for the file system or object storage system involved in your transfer
- The last modified date on an Amazon S3 object that DataSync wrote to
- Network traffic coming from the DataSync agent
- Amazon EventBridge events

## Error: SyncTaskDeletedByUser

You may see this error unexpectedly when automating some DataSync workflows. For example, maybe you have a script that's deleting your task before a task execution has finished or is in [queue](#).

To fix this issue, reconfigure your automation so that these types of actions don't overlap.

## Error: NoMem

The set of data you're trying to transfer may be too large for DataSync. If you see this error, contact [AWS Support Center](#).

## Error: FsS3UnableToConnectToEndpoint

DataSync can't connect to your [Amazon S3 location](#). This could mean the location's S3 bucket isn't reachable or the location isn't configured correctly.

Do the following until you resolve the issue:

- Check if DataSync can [access your S3 bucket](#).
- Make your sure location is configured correctly by using the DataSync console or [DescribeLocationS3](#) operation.

## Error: FsS3HeadBucketFailed

DataSync can't access the S3 bucket that you're transferring to or from. Check if DataSync has permission to access the bucket by using the Amazon S3 [HeadBucket](#) operation. If you need to adjust your permissions, see [Providing DataSync access to S3 buckets](#).

## Error: FsAzureBlobVolRootListBlobsFailed

The shared access signature (SAS) token that DataSync uses to access your Microsoft Azure Blob Storage doesn't have the List permission.

To resolve the issue, [update your location](#) with a token that has the List permission and try running your task again.

## Error: SrcLocHitAccess

DataSync can't access your source location. Check that DataSync has permission to access the location and try running your task again.

## Error: SyncTaskErrorLocationNotAdded

DataSync can't access your location. Check that DataSync has permission to access the location and try running your task again.

## Task report errors

You might run into one of the following errors while trying to monitor your DataSync transfer with a task report.

Error message	Workaround
File path exceeds the maximum length of 4,096 characters. Cannot write to Task Report	N/A (DataSync can't transfer a file with a path that exceeds 4,096 bytes)  For more information, see <a href="#">Storage system, file, and object limits</a> .
Failed to upload Task Report(s) to S3 due to an invalid bucket or IAM role	Check that the <a href="#">DataSync IAM role</a> has the right permissions to upload a task report to your S3 bucket.

Error message	Workaround
Execution error occurred prior to generating any Task Reports	Check your <a href="#">CloudWatch logs</a> to identify why your task execution failed.

## Task with Amazon S3 fails with HeadObject or GetObjectTagging error

If you're transferring objects with specific version IDs from an S3 bucket, you might see an error related to HeadObject or GetObjectTagging. For example, here's an error related to GetObjectTagging:

```
[WARN] Failed to read metadata for file /picture1.png (versionId: 111111): S3 Get Object Tagging Failed
[ERROR] S3 Exception: op=GetObjectTagging photos/picture1.png, code=403, type=15, exception=AccessDenied,
msg=Access Denied req-hdrs: content-type=application/xml, x-amz-api-version=2006-03-01
rsp-hdrs: content-type=application/xml,
date=Wed, 07 Feb 2024 20:16:14 GMT, server=AmazonS3, transfer-encoding=chunked,
x-amz-id-2=I0WQ4fDEXAMPLEQM+ey7N9WgVhSnQ6JEXAMPLEZb7hSQDASK+Jd1vEXAMPLERa3Km, x-amz-request-id=79104EXAMPLEB723
```

If you see either of these errors, validate that the IAM role that DataSync uses to access your S3 source location has the following permissions:

- s3:GetObjectVersion
- s3:GetObjectVersionTagging

If you need to update your role with these permissions, see [Creating an IAM role for DataSync to access your Amazon S3 location](#).

## My Amazon S3 storage costs are higher than expected with AWS DataSync

If your Amazon S3 storage costs are higher than you thought they would be following an AWS DataSync transfer, it might be due to one or more of the following reasons:

- When transferring to or from S3 buckets, you incur costs related to S3 API requests made by DataSync.
- DataSync uses the Amazon S3 multipart upload feature to upload objects to S3 buckets. This approach can result in unexpected storage charges for uploads that don't complete successfully.
- Object versioning might be enabled on your S3 bucket. Object versioning results in Amazon S3 storing multiple copies of objects that have the same name.

## Actions to take

In these cases, you can take the following steps:

- Make sure you understand how DataSync uses S3 requests and how they might be affecting your storage costs. For more information, see [Evaluating S3 request costs when using DataSync](#).
- If the issue's related to multipart uploads, configure a policy for multipart uploads for your S3 bucket to clean up incomplete multipart uploads to reduce storage cost. For more information, see the AWS blog post [S3 Lifecycle Management Update - Support for Multipart Uploads and Delete Markers](#).
- If the issue's related to object versioning, disable object versioning on your S3 bucket.
- If you need more help, contact [AWS Support Center](#).



# AWS DataSync tutorials

These tutorials walk you through some real-world scenarios with AWS DataSync.

## Topics

- [Tutorial: Transferring data from on-premises storage to Amazon S3 across AWS accounts](#)
- [Tutorial: Transferring data between Amazon S3 buckets across AWS accounts](#)

## Tutorial: Transferring data from on-premises storage to Amazon S3 across AWS accounts

When using AWS DataSync with on-premises storage, you typically transfer data to an AWS storage service that belongs to the same AWS account as your DataSync agent. There are situations, however, where you might need to transfer data to an Amazon S3 bucket that's associated with a different account.

### Important

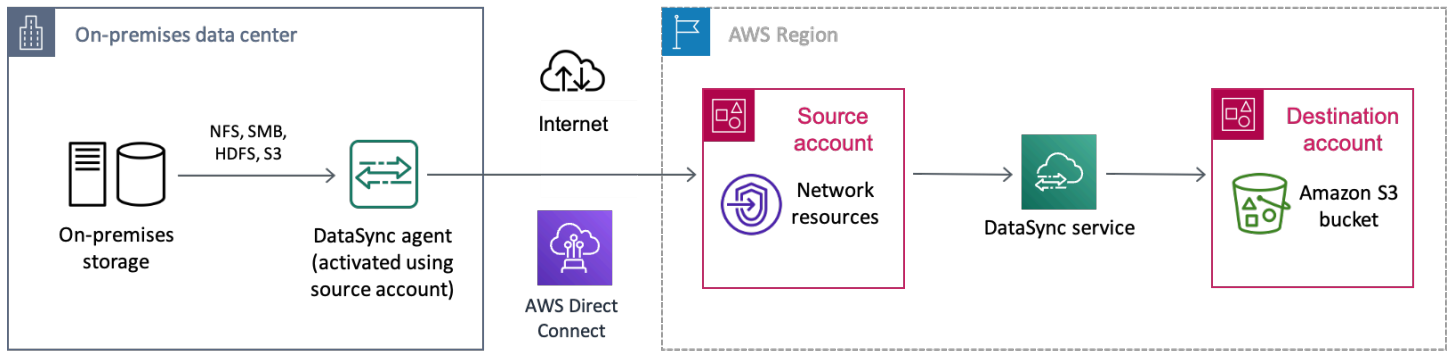
Transferring data across AWS accounts by using the methods in this tutorial works only when Amazon S3 is one of the DataSync transfer locations.

## Overview

It's not uncommon to need to transfer data between different AWS accounts, especially if you have separate teams managing your organization's resources. Here's what a cross-account transfer using DataSync can look like:

- **Source account:** The AWS account for managing network resources. This is the account that you activate your DataSync agent with.
- **Destination account:** The AWS account for managing the S3 bucket that you need to transfer data to.

The following diagram illustrates this kind of scenario.



## Prerequisite: Required source account permissions

For your source AWS account, there are two sets of permissions to consider with this kind of cross-account transfer:

- *User permissions* that allow a user to work with DataSync (this might be you or your storage administrator). These permissions let you create DataSync locations and tasks.
- *DataSync service permissions* that allow DataSync to transfer data to your destination account bucket.

### User permissions

In your source account, add at least the following permissions to an IAM role for creating your DataSync locations and task. For information on how to add permissions to a role, see [creating](#) or [modifying](#) an IAM role.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "SourceUserRolePermissions",
      "Effect": "Allow",
      "Action": [
        "datasync:CreateLocationS3",
        "datasync:CreateTask",
        "datasync:DescribeLocation*",
        "datasync:DescribeTaskExecution",
        "datasync:ListLocations",
        "datasync:ListTaskExecutions",
        "datasync:DescribeTask",
        "datasync:CancelTaskExecution",

```

```

        "datasync:ListTasks",
        "datasync:StartTaskExecution",
        "iam:CreateRole",
        "iam:CreatePolicy",
        "iam:AttachRolePolicy",
        "iam:ListRoles",
        "s3:GetBucketLocation",
        "s3:ListAllMyBuckets"
    ],
    "Resource": "*"
},
{
    "Effect": "Allow",
    "Action": [
        "iam:PassRole"
    ],
    "Resource": "*",
    "Condition": {
        "StringEquals": {
            "iam:PassedToService": [
                "datasync.amazonaws.com"
            ]
        }
    }
}
]
}

```

### Tip

To set up your *user permissions*, consider using [AWSDataSyncFullAccess](#). This is an AWS managed policy that provides a user full access to DataSync and minimal access to its dependencies.

## DataSync service permissions

The DataSync service needs the following permissions in your source account to transfer data to your destination account bucket.

Later in this tutorial, you add these permissions when [creating an IAM role](#) for DataSync. You also specify this role (*source-datasync-role*) in your [destination bucket policy](#) and when [creating your DataSync destination location](#).

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

## Prerequisite: Required destination account permissions

In your destination account, your *user permissions* must allow you to update your destination bucket's policy and disable its access control lists (ACLs). For more information on these specific permissions, see the [Amazon S3 User Guide](#).

## Step 1: In your source account, create a DataSync agent

To get started, you must create a DataSync agent that can read from your on-premises storage system and communicate with the DataSync service. This process includes deploying an agent in your on-premises storage environment and activating the agent in your source AWS account.

### Note

The steps in this tutorial apply to any type of agent and service endpoint that you use.

### To create a DataSync agent

1. [Deploy a DataSync agent](#) in your on-premises storage environment.
2. [Choose a service endpoint](#) that the agent will use to communicate with AWS.
3. [Activate your agent](#) in your source account.

## Step 2: In your source account, create a DataSync IAM role for destination bucket access

In your source account, you need an IAM role that gives DataSync the permissions to transfer data to your destination account bucket.

Since you're transferring across accounts, you must create the role manually. (DataSync can create this role for you in the console when transferring in the same account.)

### Create the DataSync IAM role

Create an IAM role with DataSync as the trusted entity.

#### To create the IAM role

1. Log in to the AWS Management Console with your source account.
2. Open the IAM console at <https://console.aws.amazon.com/iam/>.
3. In the left navigation pane, under **Access management**, choose **Roles**, and then choose **Create role**.
4. On the **Select trusted entity** page, for **Trusted entity type**, choose **AWS service**.
5. For **Use case**, choose **DataSync** in the dropdown list and select **DataSync**. Choose **Next**.

6. On the **Add permissions** page, choose **Next**.
7. Give your role a name and choose **Create role**.

For more information, see [Creating a role for an AWS service \(console\)](#) in the *IAM User Guide*.

## Add permissions to the DataSync IAM role

The IAM role that you just created needs the permissions that allow DataSync to transfer data to the S3 bucket in your destination account.

### To add permissions to your IAM role

1. On the **Roles** page of the IAM console, search for the role that you just created and choose its name.
2. On the role's details page, choose the **Permissions** tab. Choose **Add permissions** then **Create inline policy**.
3. Choose the **JSON** tab and do the following:
  - a. Paste the following JSON into the policy editor:

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

```

    ],
    "Effect": "Allow",
    "Resource": "arn:aws:s3:::destination-bucket/*"
  }
]
}

```

- b. Replace each instance of *destination-bucket* with the name of the S3 bucket in your destination account.
4. Choose **Next**. Give your policy a name and choose **Create policy**.

### Step 3: In your destination account, update your S3 bucket policy

In your destination account, modify the destination S3 bucket policy to include the [DataSync IAM role](#) that you created in your source account.

**Before you begin:** Make sure that you have the [required permissions for your destination account](#).

#### To update the destination S3 bucket policy

1. In the AWS Management Console, switch to your destination account.
2. Open the Amazon S3 console at <https://console.aws.amazon.com/s3/>.
3. In the left navigation pane, choose **Buckets**.
4. In the **Buckets** list, choose the S3 bucket that you're transferring data to.
5. On the bucket's detail page, choose the **Permissions** tab.
6. Under **Bucket policy**, choose **Edit** and do the following to modify your S3 bucket policy:
  - a. Update what's in the editor to include the following policy statements:

```

{
  "Version": "2008-10-17",
  "Statement": [
    {
      "Sid": "DataSyncCreateS3LocationAndTaskAccess",
      "Effect": "Allow",
      "Principal": {
        "AWS": "arn:aws:iam::source-account:role/source-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:::destination-bucket",
        "arn:aws:s3:::destination-bucket/*"
    ]
}
]
}

```

- b. Replace each instance of *source-account* with the AWS account ID for your source account.
  - c. Replace *source-datasync-role* with the [IAM role that you created for DataSync in your source account](#).
  - d. Replace each instance of *destination-bucket* with the name of the S3 bucket in your destination account.
7. Choose **Save changes**.

## Step 4: In your destination account, disable ACLs for your S3 bucket

It's important that all the data that you copy to the S3 bucket belongs to your destination account. To ensure that this account owns the data, disable the bucket's access control lists (ACLs). For more information, see [Controlling ownership of objects and disabling ACLs for your bucket](#) in the *Amazon S3 User Guide*.

### To disable ACLs for your destination bucket

1. While still logged in to the S3 console with your destination account, choose the S3 bucket that you're transferring data to.
2. On the bucket's detail page, choose the **Permissions** tab.
3. Under **Object Ownership**, choose **Edit**.



4. If it isn't already selected, choose the **ACLs disabled (recommended)** option.
5. Choose **Save changes**.

## Step 5: In your source account, create a DataSync source location for your on-premises storage

In your source account, create a [DataSync source location](#) for the on-premises storage system that you're transferring data from. This location uses the [agent that you activated](#) in your source account.

## Step 6: In your source account, create a DataSync destination location for your S3 bucket

While still in your source account, create a location for the S3 bucket that you're transferring data to.

**Before you begin:** Make sure that you have the [required permissions for your source account](#).

Since you can't create cross-account locations by using the DataSync console interface, these instructions require that you run a `create-location-s3` command to create your destination location. We recommend running the command by using AWS CloudShell, a browser-based, pre-authenticated shell that you launch directly from the console. CloudShell allows you to run AWS CLI commands like `create-location-s3` without downloading or installing command line tools.

### Note

To complete the following steps by using a command line tool other than CloudShell, make sure that your [AWS CLI profile](#) uses the same IAM role that includes the [required user permissions](#) to use DataSync in your source account.

### To create a DataSync destination location by using CloudShell

1. While still in your source account, do one of the following to launch CloudShell from the console:
  - Choose the CloudShell icon on the console navigation bar. It's located to the right of the search box.

- Use the search box on the console navigation bar to search for **CloudShell** and then choose the **CloudShell** option.
2. Copy the following command:

```
aws datasync create-location-s3 \  
  --s3-bucket-arn arn:aws:s3:::destination-bucket \  
  --s3-config '{  
    "BucketAccessRoleArn": "arn:aws:iam::source-user-account:role/source-datasync-  
role"  
  }'
```

3. Replace *destination-bucket* with the name of the S3 bucket in your destination account.
4. Replace *source-user-account* with the AWS account ID for your source account.
5. Replace *source-datasync-role* with the [DataSync IAM role](#) that you created in your source account.
6. Run the command in CloudShell.

If the command returns a DataSync location ARN similar to this, you successfully created the location:

```
{  
  "LocationArn": "arn:aws:datasync:us-east-2:123456789012:location/loc-  
  abcdef01234567890"  
}
```

7. In the left navigation pane, expand **Data transfer**, then choose **Locations**.

From your source account, you can see the S3 location that you just created for your destination account bucket.

## Step 7: In your source account, create and start your DataSync task

Before starting a DataSync task to transfer your data, let's recap what you've done so far:

- In your source account, you created your DataSync agent. The agent can read from your on-premises storage system and communicate with the DataSync service.
- In your source account, you created an IAM role that allows DataSync to transfer data to the S3 bucket in your destination account.

- In your destination account, you configured your S3 bucket so that DataSync can transfer data to it.
- In your source account, you created the DataSync source and destination locations for your transfer.

## To create and start the DataSync task

1. While still using the DataSync console in your source account, expand **Data transfer** in the left navigation pane, then choose **Tasks** and **Create task**.
2. On the **Configure source location** page, choose **Choose an existing location**. Choose the source location that you're copying data from (your on-premises storage) then **Next**.
3. On the **Configure destination location** page, choose **Choose an existing location**. Choose the destination location that you're copying data to (the S3 bucket in your destination account) then **Next**.
4. On the **Configure settings** page, give the task a name. As needed, configure additional settings, such as specifying an Amazon CloudWatch log group. Choose **Next**.
5. On the **Review** page, review your settings and choose **Create task**.
6. On the task's details page, choose **Start**, and then choose one of the following:
  - To run the task without modification, choose **Start with defaults**.
  - To modify the task before running it, choose **Start with overriding options**.

When your task finishes, check the S3 bucket in your destination account. You should see the data that moved from your source location.

## Related resources

For more information about what you did in this tutorial, see the following topics:

- [Creating a role for an AWS service \(console\)](#)
- [Modifying a role trust policy \(console\)](#)
- [Adding a bucket policy by using the Amazon S3 console](#)
- [Create an S3 location with the AWS CLI](#)

# Tutorial: Transferring data between Amazon S3 buckets across AWS accounts

With AWS DataSync, you can transfer data between Amazon S3 buckets that belong to different AWS accounts.

## Important

Transferring data across AWS accounts using the methods in this tutorial works only with Amazon S3. Additionally, this tutorial can help you transfer data between S3 buckets that are also in different AWS Regions (unless you're working with one or more [opt-in Regions](#)).

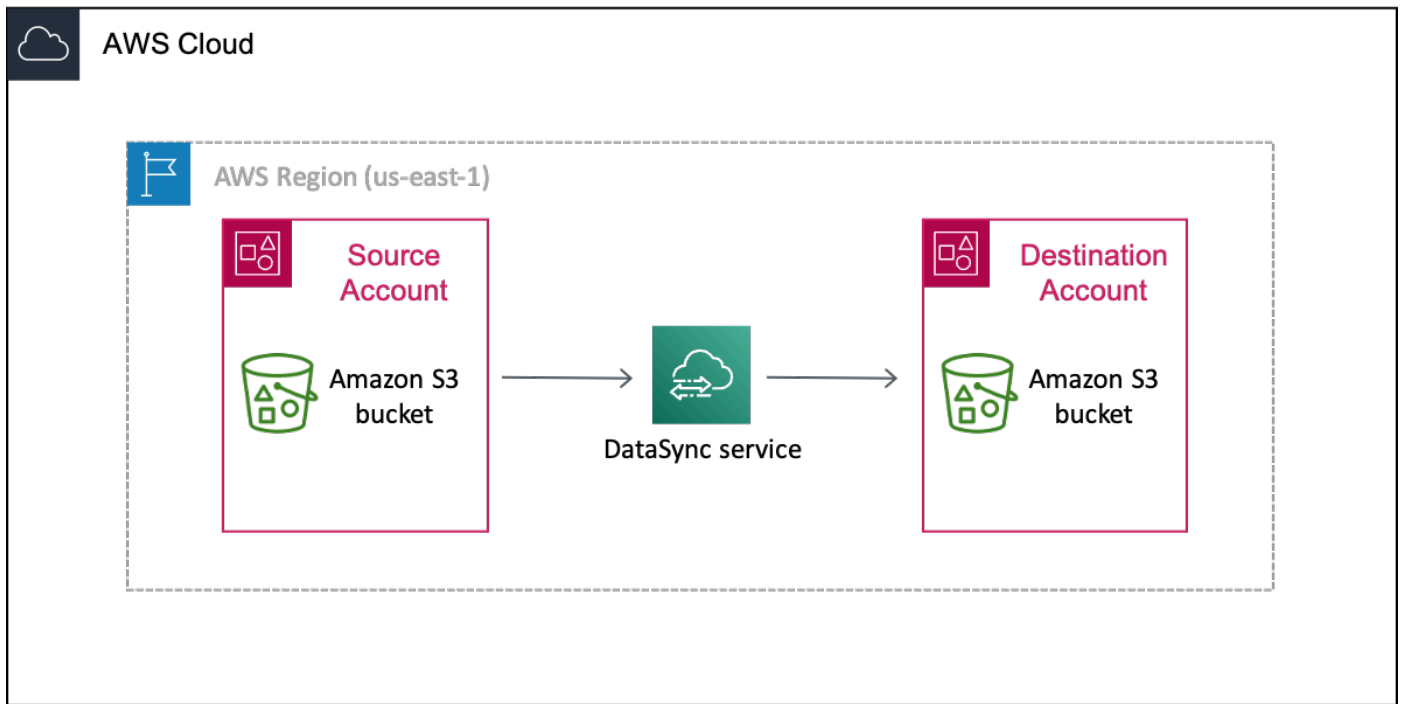
## Overview

It's not uncommon to transfer data between AWS accounts, especially if you have separate teams managing your organization's resources. Here's what a cross-account transfer using DataSync can look like:

- **Source account:** The AWS account for managing the S3 bucket that you need to transfer data from.
- **Destination account:** The AWS account for managing the S3 bucket that you need to transfer data to.

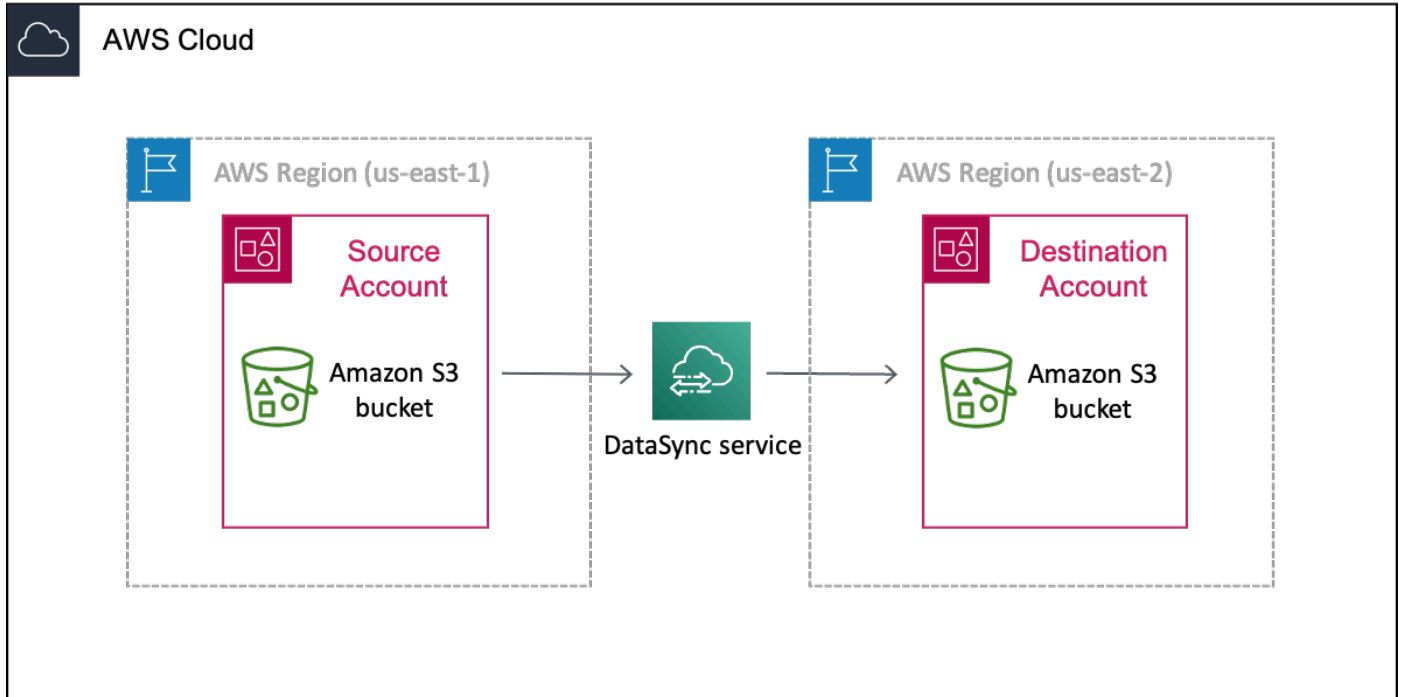
### Transfers across accounts

The following diagram illustrates a scenario where you transfer data from an S3 bucket to another S3 bucket that's in a different AWS account.



### Transfers across accounts and Regions

The following diagram illustrates a scenario where you transfer data from an S3 bucket to another S3 bucket that's in a different AWS account and Region.



## Prerequisite: Required source account permissions

For your source AWS account, there are two sets of permissions to consider with this kind of cross-account transfer:

- *User permissions* that allow a user to work with DataSync (this might be you or your storage administrator). These permissions let you create DataSync locations and tasks.
- *DataSync service permissions* that allow DataSync to transfer data to your destination account bucket.

### User permissions for your source account

In your source account, add at least the following permissions to an IAM role for creating your DataSync locations and task. For information on how to add permissions to a role, see [creating](#) or [modifying](#) an IAM role.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "SourceUserRolePermissions",
      "Effect": "Allow",
      "Action": [
        "datasync:CreateLocationS3",
        "datasync:CreateTask",
        "datasync:DescribeLocation*",
        "datasync:DescribeTaskExecution",
        "datasync:ListLocations",
        "datasync:ListTaskExecutions",
        "datasync:DescribeTask",
        "datasync:CancelTaskExecution",
        "datasync:ListTasks",
        "datasync:StartTaskExecution",
        "iam:CreateRole",
        "iam:CreatePolicy",
        "iam:AttachRolePolicy",
        "iam:ListRoles",
        "s3:GetBucketLocation",
        "s3:ListAllMyBuckets"
      ],
      "Resource": "*"
    }
  ]
}
```

```

    },
    {
      "Effect": "Allow",
      "Action": [
        "iam:PassRole"
      ],
      "Resource": "*",
      "Condition": {
        "StringEquals": {
          "iam:PassedToService": [
            "datasync.amazonaws.com"
          ]
        }
      }
    }
  ]
}

```

### Tip

To set up your *user permissions*, consider using [AWSDataSyncFullAccess](#). This is an AWS managed policy that provides a user full access to DataSync and minimal access to its dependencies.

## DataSync service permissions for your source account

The DataSync service needs the following permissions in your source account to transfer data to your destination account bucket.

Later in this tutorial, you add these permissions when [creating an IAM role](#) for DataSync. You also specify this role (*source-datasync-role*) in your [destination bucket policy](#) and when [creating your DataSync destination location](#).

```

{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Action": [
        "s3:GetBucketLocation",
        "s3:ListBucket",

```

```
    "s3:ListBucketMultipartUploads"
  ],
  "Effect": "Allow",
  "Resource": "arn:aws:s3:::destination-bucket"
},
{
  "Action": [
    "s3:AbortMultipartUpload",
    "s3:DeleteObject",
    "s3:GetObject",
    "s3:ListMultipartUploadParts",
    "s3:PutObject",
    "s3:GetObjectTagging",
    "s3:PutObjectTagging"
  ],
  "Effect": "Allow",
  "Resource": "arn:aws:s3:::destination-bucket/*"
}
]
```

## Prerequisite: Required destination account permissions

In your destination account, your *user permissions* must allow you to update your destination bucket's policy and disable its access control lists (ACLs). For more information on these specific permissions, see the [Amazon S3 User Guide](#).

## Step 1: In your source account, create a DataSync IAM role for destination bucket access

In your source AWS account, you need an IAM role that gives DataSync the permissions to transfer data to your destination account bucket.

Since you're transferring across accounts, you must create the role manually. (DataSync can create this role for you in the console when transferring in the same account.)

### Create the DataSync IAM role

Create an IAM role with DataSync as the trusted entity.

1. Log in to the AWS Management Console with your source account.



2. Open the IAM console at <https://console.aws.amazon.com/iam/>.
3. In the left navigation pane, under **Access management**, choose **Roles**, and then choose **Create role**.
4. On the **Select trusted entity** page, for **Trusted entity type**, choose **AWS service**.
5. For **Use case**, choose **DataSync** in the dropdown list and select **DataSync**. Choose **Next**.
6. On the **Add permissions** page, choose **Next**.
7. Give your role a name and choose **Create role**.

For more information, see [Creating a role for an AWS service \(console\)](#) in the *IAM User Guide*.

### Add permissions to the DataSync IAM role

The IAM role that you just created needs the permissions that allow DataSync to transfer data to the S3 bucket in your destination account.

1. On the **Roles** page of the IAM console, search for the role that you just created and choose its name.
2. On the role's details page, choose the **Permissions** tab. Choose **Add permissions** then **Create inline policy**.
3. Choose the **JSON** tab and do the following:
  - a. Paste the following JSON into the policy editor:

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Action": [
        "s3:GetBucketLocation",
        "s3:ListBucket",
        "s3:ListBucketMultipartUploads"
      ],
      "Effect": "Allow",
      "Resource": "arn:aws:s3:::destination-bucket"
    },
    {
      "Action": [
        "s3:AbortMultipartUpload",
        "s3:DeleteObject",
```

```

        "s3:GetObject",
        "s3:ListMultipartUploadParts",
        "s3:PutObject",
        "s3:GetObjectTagging",
        "s3:PutObjectTagging"
    ],
    "Effect": "Allow",
    "Resource": "arn:aws:s3:::destination-bucket/*"
}
]
}

```

- b. Replace each instance of *destination-bucket* with the name of the S3 bucket in your destination account.
4. Choose **Next**. Give your policy a name and choose **Create policy**.

## Step 2: In your destination account, update your S3 bucket policy

In your destination account, modify the destination S3 bucket policy to include the [DataSync IAM role](#) that you created in your source account.

**Before you begin:** Make sure that you have the [required permissions for your destination account](#).

### Update your destination S3 bucket policy

1. In the AWS Management Console, switch to your destination account.
2. Open the Amazon S3 console at <https://console.aws.amazon.com/s3/>.
3. In the left navigation pane, choose **Buckets**.
4. In the **Buckets** list, choose the S3 bucket that you're transferring data to.
5. On the bucket's detail page, choose the **Permissions** tab.
6. Under **Bucket policy**, choose **Edit** and do the following to modify your S3 bucket policy:
  - a. Update what's in the editor to include the following policy statements:

```

{
  "Version": "2008-10-17",
  "Statement": [
    {
      "Sid": "DataSyncCreateS3LocationAndTaskAccess",

```

```

    "Effect": "Allow",
    "Principal": {
      "AWS": "arn:aws:iam::source-account:role/source-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::destination-bucket",
      "arn:aws:s3::destination-bucket/*"
    ]
  }
]
}

```

- b. Replace each instance of *source-account* with the AWS account ID for your source account.
- c. Replace *source-datasync-role* with the [IAM role that you created for DataSync in your source account](#).
- d. Replace each instance of *destination-bucket* with the name of the S3 bucket in your destination account.

7. Choose **Save changes**.

### Step 3: In your destination account, disable ACLs for your S3 bucket

It's important that all the data that you transfer to the S3 bucket belongs to your destination account. To ensure that this account owns the data, disable the bucket's access control lists (ACLs). For more information, see [Controlling ownership of objects and disabling ACLs for your bucket](#) in the *Amazon S3 User Guide*.

**Before you begin:** Make sure that you have the [required permissions for your destination account](#).

## Disable your destination S3 bucket ACLs

1. While still logged in to the S3 console with your destination account, choose the S3 bucket that you're transferring data to.
2. On the bucket's detail page, choose the **Permissions** tab.
3. Under **Object Ownership**, choose **Edit**.
4. If it isn't already selected, choose the **ACLs disabled (recommended)** option.
5. Choose **Save changes**.

## Step 4: In your source account, create your DataSync locations

In your source account, create the DataSync locations for your source and destination S3 buckets.

**Before you begin:** Make sure that you have the [required permissions for your source account](#).

### Create your DataSync source location

- In your source account, create a [location](#) for the S3 bucket that you're transferring data from.

### Create your DataSync destination location

While still in your source account, create a location for the S3 bucket that you're transferring data to.

Since you can't create cross-account locations by using the DataSync console interface, these instructions require that you run a `create-location-s3` command to create your destination location. We recommend running the command by using AWS CloudShell, a browser-based, pre-authenticated shell that you launch directly from the console. CloudShell allows you to run AWS CLI commands like `create-location-s3` without downloading or installing command line tools.

#### Note

To complete the following steps by using a command line tool other than CloudShell, make sure that your [AWS CLI profile](#) uses the same IAM role that includes the [required user permissions](#) to use DataSync in your source account.

## To create a DataSync destination location by using CloudShell

1. While still in your source account, do one of the following to launch CloudShell from the console:
  - Choose the CloudShell icon on the console navigation bar. It's located to the right of the search box.
  - Use the search box on the console navigation bar to search for **CloudShell** and then choose the **CloudShell** option.
2. Copy the following `create-location-s3` command:

```
aws datasync create-location-s3 \  
  --s3-bucket-arn arn:aws:s3:::destination-bucket \  
  --region destination-bucket-region \  
  --s3-config '{  
    "BucketAccessRoleArn": "arn:aws:iam::source-account-id:role/source-datasync-  
role"  
  }'
```

3. Replace *destination-bucket* with the name of the S3 bucket in your destination account.
4. If your destination bucket is in a different Region than your source bucket, replace *destination-bucket-region* with the Region where the destination bucket resides (for example, *us-east-2*). Remove this option if your buckets are in the same Region.
5. Replace *source-account-id* with the source AWS account ID.
6. Replace *source-datasync-role* with the [DataSync IAM role](#) that you created in your source account.
7. Run the command in CloudShell.

If the command returns a DataSync location ARN similar to this, you successfully created the location:

```
{  
  "LocationArn": "arn:aws:datasync:us-east-2:123456789012:location/loc-  
abcdef01234567890"  
}
```

8. In the left navigation pane, expand **Data transfer**, then choose **Locations**.
9. If you created the location in a different Region, choose that Region in the navigation pane.

From your source account, you can see the S3 location that you just created for your destination account bucket.

## Step 5: In your source account, create and start your DataSync task

Before starting a DataSync task to transfer your data, let's recap what you've done so far:

- In your source account, you created an IAM role that allows DataSync to transfer data to the S3 bucket in your destination account.
- In your destination account, you configured your S3 bucket so that DataSync can transfer data to it.
- In your source account, you created the DataSync source and destination locations for your transfer.

### Create and start your DataSync task

1. While still using the DataSync console in your source account, expand **Data transfer** in the left navigation pane, then choose **Tasks** and **Create task**.
2. If the bucket in your destination account is in a different Region than the bucket in your source account, choose the destination bucket's Region in the top navigation pane.

#### Important

To avoid a network connection error, you must create your DataSync task in the same Region as the destination location.

3. On the **Configure source location** page, do the following:
  - a. Select **Choose an existing location**.
  - b. (For transfers across Regions) In the **Region** dropdown, choose the Region where the source bucket resides.
  - c. For **Existing locations**, choose the source location for the S3 bucket that you're transferring data from, then choose **Next**.
4. On the **Configure destination location** page, do the following:
  - a. Select **Choose an existing location**.

- b. For **Existing locations**, choose the destination location for the S3 bucket that you're transferring data to, then choose **Next**.
5. On the **Configure settings** page, give the task a name. As needed, configure additional settings, such as specifying an Amazon CloudWatch log group. Choose **Next**.
6. On the **Review** page, review your settings and choose **Create task**.
7. On the task's details page, choose **Start**, and then choose one of the following:
  - To run the task without modification, choose **Start with defaults**.
  - To modify the task before running it, choose **Start with overriding options**.

When your task finishes, check the S3 bucket in your destination account. You should see the data that moved from your source account bucket.

## Troubleshooting

Refer to the following information if you run into issues trying to complete your cross-account transfer.

### Connection errors

When transferring between S3 buckets in different AWS accounts and Regions, you might get a network connection error when starting your DataSync task. To resolve this, create a task in the same Region as your destination location and try running that task.

## Related: Cross-account transfers with S3 buckets using server-side encryption

If you're trying to do this transfer with S3 buckets using server-side encryption, see the [AWS Storage Blog](#) for instructions.

# Additional resources for AWS DataSync

In this section, you can find additional information about and resources for AWS DataSync.

## Topics

- [Other use cases for AWS DataSync](#)
- [Open-source components for AWS DataSync](#)

## Other use cases for AWS DataSync

Learn about about some less-common use cases with AWS DataSync that are not common to most users.

### Transferring data in opposite directions

Transferring data in opposite directions allows for workflows where the active application moves between locations. AWS DataSync doesn't support workflows where multiple active applications write to both locations at the same time. Use the steps in the following procedure to configure DataSync to transfer data in opposite directions.

#### To configure DataSync to data transfers in opposite directions

1. Create a location and name it **Location A**.
2. Create a second location and name it **Location B**.
3. Create a task, name it **Task A-B**, and then configure **Location A** as the source location and **Location B** as the destination location.
4. Create a second task, name it **Task B-A**, and then configure **Location B** as the source location and **Location A** as the destination location.
5. To update **Location B** with data from **Location A**, run **Task A-B**.

To update **Location A** with data from **Location B**, run **Task B-A**.

Don't run these two tasks concurrently. DataSync can transfer files in opposite directions periodically. However, it doesn't support workflows where multiple active applications write to both **Location A** and **Location B** simultaneously.



## Using multiple DataSync tasks to transfer to the same S3 bucket

In certain use cases, you might want different tasks to write to the same Amazon S3 bucket. In this case, you create different folders in the S3 bucket for each of the task. This approach prevents file name conflicts between the tasks, and also means that you can set different permissions for each of folders.

For example, you might have three tasks: `task1`, `task2`, and `task3` write to an S3 bucket named `MyBucket`.

You create three folders in the bucket:

```
s3://MyBucket/task1
```

```
s3://MyBucket/task2
```

```
s3://MyBucket/task3
```

For each task, you choose the folder in `MyBucket` that corresponds to the task as the destination, and set different permissions for each of the three folders.

## Allowing DataSync to access a restricted S3 bucket

In some cases, you might want to limit access to your Amazon S3 bucket. You can edit the S3 bucket policy so that DataSync can still access the bucket when you run a task.

### To allow DataSync to access a restricted S3 bucket

1. Copy the following sample policy.

```
{
  "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. In the sample policy, replace these values:

- *bucket-name*: The name of the S3 bucket that you're restricting access to.
- *datasync-role-id*: The ID of the IAM role that DataSync accesses the S3 bucket with. Run the following AWS CLI command to get the IAM role ID:

```
aws iam get-role --role-name datasync-iam-role-name
```

In the output, look for the RoleId value:

```
"RoleId": "ANPAJ2UCCR6DPCEXAMPLE"
```

- *your-role-id*: The ID of the IAM role that you create the DataSync location for the S3 bucket with. Run the following command to get the IAM role ID:

```
aws iam get-role --role-name your-iam-role-name
```

In the output, look for the UserId value:

```
"RoleId": "AIDACKCEVSQ6C2EXAMPLE"
```

3. Add this policy to your S3 bucket policy. For more information, see how to [edit a bucket policy](#) in the *Amazon S3 User Guide*.

Once you've updated the S3 bucket policy, you must add additional IAM roles or users to the policy for those who need to access the S3 bucket.

## Open-source components for AWS DataSync

To view the open-source components used by AWS DataSync, download the following link:

- [datasync-open-source-components.zip](#)

# Using the AWS Command Line Interface with AWS DataSync

In this section, you can find examples of using the AWS Command Line Interface (AWS CLI) commands for AWS DataSync. You can use these commands to create an agent, create source and destination locations, and run a task.

Before you begin, we recommend reading [How AWS DataSync works](#) to understand the components and terms used in DataSync and how the service works. We also recommend reading [IAM customer managed policies for AWS DataSync](#) to understand the AWS Identity and Access Management (IAM) permissions that DataSync requires.

Before you use AWS CLI commands, install the AWS CLI. For information about how to install the AWS CLI, see [Installing the AWS Command Line Interface](#) in the *AWS Command Line Interface User Guide*. After you install the AWS CLI, you can use the `help` command to see the DataSync operations and the parameters associated with them.

To see the available operations, enter the following command.

```
aws datasync help
```

To see the parameters associated with a specific operation, enter the following command.

```
aws datasync operation help
```

For more information about the AWS CLI, see [What is the AWS Command Line Interface?](#)

## Topics

- [Creating an AWS DataSync agent with the AWS CLI](#)
- [Creating AWS DataSync locations with the AWS CLI](#)
- [Creating an AWS DataSync task with the AWS CLI](#)
- [Starting an AWS DataSync task with the AWS CLI](#)
- [Filtering AWS DataSync resources](#)

For information about supported AWS Regions and endpoints, see [AWS DataSync endpoints and quotas](#).

For information about DataSync Amazon Resource Name (ARN) values, see [Amazon Resource Names](#).

## Creating an AWS DataSync agent with the AWS CLI

To access your self-managed storage, you first deploy and activate an AWS DataSync agent. The activation process associates your agent with your AWS account. An agent isn't required when transferring between AWS storage services within the same AWS account. To set up a data transfer between two AWS services, see [Creating AWS DataSync locations with the AWS CLI](#).

A DataSync agent can transfer data through public service endpoints, Federal Information Processing Standard (FIPS) endpoints, and Amazon VPC endpoints.

### Note

When you configure your agent to use Amazon VPC endpoints, the data transferred between your agent and the DataSync service doesn't cross the public internet and doesn't require public IP addresses. For end-to-end instructions for this configuration, see [Using a VPC service endpoint with DataSync](#).

### To create an agent to read from a Network File System (NFS), Server Message Block (SMB), Hadoop Distributed File System (HDFS), or self-managed object storage source location

1. Download the current DataSync .ova image or launch the current DataSync Amazon Machine Image (AMI) based on Amazon EC2 from the AWS DataSync console. For information about how to get the .ova image or Amazon EC2 AMI, see [Deploying your AWS DataSync agent](#). For information about hardware requirements and recommended Amazon EC2 instance types, see [Virtual machine requirements](#).

### Important

If you are deploying your agent on Amazon EC2, deploy the agent so that it doesn't require network traffic between Availability Zones (to avoid charges for such traffic).

- To access your Amazon EFS or Amazon FSx for Windows File Server file system, deploy the agent in an Availability Zone that has a mount target to your file system.
- For self-managed file systems, deploy the agent in the Availability Zone where your file system resides.

To learn more about data-transfer prices for all AWS Regions, see [Amazon EC2 On-Demand pricing](#).

2. Make sure that you satisfy the network-connectivity requirements for the agent. For information about network requirements, see [AWS DataSync network requirements](#).
3. Deploy the .ova image in your hypervisor, power on the hypervisor, and note the agent's IP address. Make sure that you can reach the agent on port 80. You can use the following command to check.

```
nc -vz agent-ip-address 80
```

### Note

The .ova default credentials are login **admin**, password **password**. You can change the password on the virtual machine (VM) local console. You don't need to log in to the VM for basic DataSync functionality. Logging in is required mainly for troubleshooting, network-specific settings, and so on.

You log in to the agent VM local console by using your VM's hypervisor client. For information about how to use the VM local console, see [Working with your AWS DataSync agent's local console](#).

4. Send an HTTP/1.1 GET request to the agent to get the activation key. You can do this by using standard Unix tools:
  - To activate an agent by using a public service endpoint, use the following command.

```
curl "http://agent-ip-address?gatewayType=SYNC&activationRegion=aws-region&no_redirect"
```

- To activate an agent by using a virtual private cloud (VPC) endpoint, use the IP address of the VPC endpoint. Use the following command.

```
curl "http://agent-ip-address?gatewayType=SYNC&activationRegion=aws-region&privateLinkEndpoint=IP address of VPC endpoint&endpointType=PRIVATE_LINK&no_redirect"
```

To find the correct IP address, open the Amazon VPC console at <https://console.aws.amazon.com/vpc/> and choose **Endpoints** from the navigation pane at left. Choose the DataSync endpoint, and check **Subnets list** to find the private IP address that corresponds to the subnet that you chose for your VPC endpoint setup.

For more information about VPC endpoint configuration, see step 5 in [Configuring your DataSync agent to use a VPC service endpoint](#).

- To activate an agent using a Federal Information Processing Standard (FIPS) endpoint, specify `endpointType=FIPS`. Also, the `activationRegion` value must be set to an AWS Region within the United States. To activate a FIPS endpoint, use the following command.

```
curl "http://agent-IP-address?gatewayType=SYNC&activationRegion=US-based-aws-region&endpointType=FIPS&no_redirect"
```

This command returns an activation key similar to the one following.

```
F0EFT-7FPPR-GG7MC-3I9R3-27DOH
```

5. After you have the activation key, do one of the following:
  - To activate your agent using a public endpoint or FIPS endpoint, use the following command.

```
aws datasync create-agent \  
  --agent-name agent-name-you-specify \  
  --activation-key obtained-activation-key
```

- To activate your agent using a VPC endpoint, use the following command.

```
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
```

In this command, use the following arguments:

- *vpc-endpoint-id* – The AWS endpoint that the agent connects to. To find the endpoint ID, open the Amazon VPC console at <https://console.aws.amazon.com/vpc/>, and choose **Endpoints** from the navigation pane on the left. Copy the **Endpoint ID** value of the DataSync endpoint. For more information about VPC endpoint configuration, see step 5 in [Configuring your DataSync agent to use a VPC service endpoint](#).
- *security-group-arn* – The Amazon Resource Names (ARNs) of the security groups to use for the task's endpoint.

This is the security group that you created in step 3 of [Configuring your DataSync agent to use a VPC service endpoint](#).

- *subnet-arns* – The ARNs of the subnets where the task endpoints for the agent are created.

This is the subnet that you chose in step 1 of [Configuring your DataSync agent to use a VPC service endpoint](#).

These commands return the ARN of the agent that you just activated. The ARN is similar to the one following.

```
{
  "AgentArn": "arn:aws:datsync:us-east-1:111222333444:agent/
agent-0b0addbeef44baca3"
}
```

 **Note**

After you choose a service endpoint, you can't change it later.

After you activate the agent, it closes port 80 and the port is no longer accessible. If you can't connect to the agent after you have activated it, verify that the activation was successful by using the following command:

```
aws datsync list-agents
```



**Note**

Make sure that you are using the same AWS credentials throughout the whole process. Don't switch between multiple terminals where you are authenticated with different AWS credentials.

## Creating AWS DataSync locations with the AWS CLI

Each AWS DataSync task is made up of a pair of locations in a transfer. The source location defines the storage system or service that you want to read data from. The destination location defines the storage system or service that you want to write data to.

### Creating an SMB location

An SMB location defines a file system on an SMB server that can be read from or written to. You can also create an SMB location by using the console. For more information, see [Configuring AWS DataSync transfers with an SMB file server](#).

#### To create an SMB location by using the CLI

- Use the following command to create an SMB source location.

```
aws datasync 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 datasync-agent-arns \  
  --subdirectory smb-export-path
```

The *smb-export-path* that you provide for the `--subdirectory` parameter should be a path that's exported by the SMB server. Specify the path by using forward slashes; for example, `/path/to/folder`. Other SMB clients in your network should be able to access this path.

DataSync automatically chooses the SMB version that it uses to read from an SMB location. To specify an SMB version, use the optional `Version` parameter in the [SmbMountOptions](#) API operation.

This command returns the Amazon Resource Name (ARN) of the SMB location, similar to the ARN shown following.

```
{
  "LocationArn": "arn:aws:datasync:us-east-1:111222333444:location/
loc-0f01451b140b2af49"
}
```

## Creating an HDFS location

An HDFS location defines a file system on a Hadoop cluster that can be read from or written to. You can also create an HDFS location by using the AWS Management Console. For more information, see [Configuring AWS DataSync transfers with HDFS](#).

### To create an HDFS location by using the AWS CLI

- Use the following command to create an HDFS location. In the following example, replace each *user input placeholder* with your own information.

```
aws datasync create-location-hdfs --name-nodes [{"Hostname":"host1", "Port": 8020}]
\
  --authentication-type "SIMPLE|KERBEROS" \
  --agent-arns [arn:aws:datasync:us-east-1:123456789012:agent/
agent-01234567890example] \
  --subdirectory "/path/to/my/data"
```

The following parameters are required in the `create-location-hdfs` command:

- `name-nodes` – Specifies the hostname or IP address of the NameNode in the Hadoop cluster and the TCP port that the NameNode is listening on.
- `authentication-type` – The type of authentication to use when connecting to the Hadoop cluster. Specify `SIMPLE` or `KERBEROS`.

If you use `SIMPLE` authentication, use the `--simple-user` parameter to specify the user name of the user. If you use `KERBEROS` authentication, use the `--kerberos-principal`, `--kerberos-keytab`, and `--kerberos-krb5-conf` parameters. For more information, see [create-location-hdfs](#).

- `agent-arns` – The ARNs of the DataSync agents to use for the HDFS location.

The preceding the command returns the location ARN, similar to the following:

```
{
  "arn:aws:datsync:us-east-1:123456789012:location/loc-01234567890example"
}
```

## Creating an object storage location

A *location* is a DataSync endpoint that represents an object storage system hosted on-premises or by another cloud provider (for example, a Google Cloud Storage bucket).

For more information about object storage locations, including compatibility requirements, see [Configuring AWS DataSync transfers with an object storage system](#).

### To create an object storage location by using the AWS CLI

1. Copy the following `create-location-object-storage` command:

```
aws datsync create-location-object-storage \
  --server-hostname object-storage-server.example.com \
  --bucket-name your-bucket \
  --agent-arns arn:aws:datsync:us-east-1:123456789012:agent/
agent-01234567890deadfb
```

2. Specify the following required parameters in the command:
  - `--server-hostname` – Specify the domain name or IP address of your object storage server.
  - `--bucket-name` – Specify the name of the bucket on your object storage server that you're transferring to or from.
  - `--agent-arns` – Specify the DataSync agents that you want to connect to your object storage server.
3. (Optional) Add any of the following parameters to the command:
  - `--server-port` – Specifies the port that your object storage server accepts inbound network traffic on (for example, port 443).

- `--server-protocol` – Specifies the protocol (HTTP or HTTPS) which your object storage server uses to communicate.
- `--access-key` – Specifies the access key (for example, a user name) if credentials are required to authenticate with the object storage server.
- `--secret-key` – Specifies the secret key (for example, a password) if credentials are required to authenticate with the object storage server.
- `--server-certificate` – Specifies a certificate chain for DataSync to authenticate with your object storage system if the system uses a private or self-signed certificate authority (CA). You must specify a single `.pem` file with a full certificate chain (for example, `file:///home/user/.ssh/object_storage_certificates.pem`).

The certificate chain might include:

- The object storage system's certificate
- All intermediate certificates (if there are any)
- The root certificate of the signing CA

You can concatenate your certificates into a `.pem` file (which can be up to 32768 bytes before base64 encoding). The following example `cat` command creates an `object_storage_certificates.pem` file that includes three certificates:

```
cat object_server_certificate.pem intermediate_certificate.pem ca_root_certificate.pem  
> object_storage_certificates.pem
```

- `--subdirectory` – Specifies the object prefix for your object storage server.

DataSync only copies objects with this prefix.

- `--tags` – Specifies the key-value pair that represents a tag that you want to add to the location resource.

Tags can help you manage, filter, and search for your resources. We recommend creating a name tag for your location.

#### 4. Run the `create-location-object-storage` command.

You get a response that shows you the location ARN that you just created.

```
{
```

```
"LocationArn": "arn:aws:datsync:us-east-1:123456789012:location/loc-01234567890abcdef"
}
```

## Creating an Amazon EFS location

A *location* is the endpoint for an Amazon EFS file system that can be read from or written to. You can also create this kind of location by using the console. For more information, see [Configuring AWS DataSync transfers with Amazon EFS](#).

### To create an Amazon EFS location by using the AWS CLI

1. If you don't have an Amazon EFS file system, create one. For information about how to create an EFS file system, see [Getting started with Amazon Elastic File System](#) in the *Amazon Elastic File System User Guide*.
2. Identify a subnet that has at least one mount target for that file system. You can see all the mount targets and the subnets associated with an EFS file system by using the `describe-mount-targets` command.

```
aws efs describe-mount-targets \
  --region aws-region \
  --file-system-id file-system-id
```

#### Note

The AWS Region that you specify is the one where your target S3 bucket or EFS file system is located.

This command returns information about the target similar to the information shown following.

```
{
  "MountTargets": [
    {
      "OwnerId": "111222333444",
      "MountTargetId": "fsmt-22334a10",
      "FileSystemId": "fs-123456ab",
```

```

        "SubnetId": "subnet-f12a0e34",
        "LifecycleState": "available",
        "IpAddress": "11.222.0.123",
        "NetworkInterfaceId": "eni-1234a044"
    }
]
}

```

3. Specify an Amazon EC2 security group that can access the mount target. You can run the following command to find out the security group of the mount target.

```

aws efs describe-mount-target-security-groups \
  --region aws-region \
  --mount-target-id mount-target-id

```

The security group that you provide must be able to communicate with the security group on the mount target in the subnet specified.

The relationship between security group M on the mount target and security group S, which you provide for DataSync to use at this stage, is as follows:

- Security group M, which you associate with the mount target, must allow inbound access for the TCP protocol on the NFS port (2049) from security group S.

You can enable an inbound connection either by its IP address (CIDR range) or its security group.

- Security group S, which you provide to DataSync to access Amazon EFS, should have a rule that enables outbound connections to the NFS port. It enables outbound connections on one of the file system's mount targets.

You can enable outbound connections either by IP address (CIDR range) or security group.

For information about security groups and mount targets, see [Security groups for Amazon EC2 instances and mount targets](#) in the *Amazon Elastic File System User Guide*.

4. Create the location. To create the EFS location, you need the ARNs for your Amazon EC2 subnet, Amazon EC2 security group, and a file system. Because the DataSync API accepts fully qualified ARNs, you can construct these ARNs. For information about how to construct ARNs for different services, see [Amazon Resource Names \(ARNs\)](#) in the *AWS General Reference*.

Use the following command to create an EFS location.

```
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

The AWS Region that you specify is the one where your target S3 bucket or EFS file system is located.

The command returns a location ARN similar to the one shown following.

```
{
  "LocationArn": "arn:aws:datasync:us-west-2:111222333444:location/
loc-07db7abfc326c50fb"
}
```

## Creating an Amazon FSx for Windows File Server location

A *location* is an endpoint for an FSx for Windows File Server that you can read from or write to.

You can also create an FSx for Windows File Server location by using the console. For more information, see [Configuring AWS DataSync transfers with Amazon FSx for Windows File Server](#).

### To create an FSx for Windows File Server location by using the AWS CLI

- Use the following command to create an Amazon FSx location.

```
aws datasync 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
```

In the `create-location-fsx-windows` command, specify the following:

- `fsx-file-system-arn` – The fully qualified Amazon Resource Name (ARN) of the file system that you want to read from or write to.

The DataSync API accepts fully qualified ARNs, and you can construct these ARNs. For information about how to construct ARNs for different services, see [Amazon Resource Names \(ARNs\)](#) in the *AWS General Reference*.

- `security-group-arns` – The ARN of an Amazon EC2 security group that can be applied to the [network interfaces](#) of the file system's preferred subnet.
- The AWS Region – The Region that you specify is the one where your target Amazon FSx file system is located.

The preceding command returns a location ARN similar to the one shown following.

```
{
  "LocationArn": "arn:aws:datsync:us-west-2:111222333444:location/
loc-07db7abfc326c50fb"
}
```

## Creating an Amazon FSx for Lustre location

A *location* is an endpoint for an FSx for Lustre file system that you can read or write to.

You can also create an FSx for Lustre location by using the console. For more information, see [Configuring AWS DataSync transfers with Amazon FSx for Lustre](#).

### To create an FSx for Lustre location by using the AWS CLI

- Use the following command to create an FSx for Lustre location.

```
aws datasync create-location-fsx-lustre \
  --fsx-file-system-arn arn:aws:fsx:region:account-id:file-system:filesystem-id \
  --security-group-arns arn:aws:ec2:region:account-id:security-group/group-id
```

The following parameters are required in the `create-location-fsx-lustre` command.

- `fsx-file-system-arn` – The fully qualified Amazon Resource Name (ARN) of the file system that you want to read from or write to.



- `security-group-arns` – The ARN of an Amazon EC2 security group to apply to the [network interfaces](#) of the file system's preferred subnet.

The preceding command returns a location ARN similar to the following.

```
{
  "LocationArn": "arn:aws:datsync:us-west-2:111222333444:location/
loc-07sb7abfc326c50fb"
}
```

## Creating an Amazon FSx for OpenZFS location

A *location* is an endpoint for an FSx for OpenZFS file system that DataSync can access for a transfer. You also can create an [FSx for OpenZFS location in the console](#).

### To create an FSx for OpenZFS location by using the AWS CLI

1. Copy the following command:

```
$ aws datsync 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. Specify the following required options in the command:
  - For `fsx-filesystem-arn`, specify the location file system's fully qualified Amazon Resource Name (ARN). This includes the AWS Region where your file system resides, your AWS account, and the file system ID.
  - For `security-group-arns`, specify the ARN of the Amazon EC2 security group that provides access to the [network interfaces](#) of your FSx for OpenZFS file system's preferred subnet. This includes the AWS Region where your Amazon EC2 instance resides, your AWS account, and the security group ID.  
  
For more information about security groups, see [File System Access Control with Amazon VPC](#) in the *Amazon FSx for OpenZFS User Guide*.
  - For `protocol`, specify the protocol that DataSync uses to access your file system. (DataSync currently supports only NFS.)
3. Run the command. You get a response showing the location that you just created.

```
{
  "LocationArn": "arn:aws:datsync:us-west-2:123456789012:location/loc-
  abcdef01234567890"
}
```

## Creating an Amazon FSx for NetApp ONTAP location

A *location* is an endpoint for an FSx for ONTAP file system that DataSync can access for a transfer. You also can create an [FSx for ONTAP location in the console](#).

### To create an FSx for ONTAP location by using the AWS CLI

1. Copy the following command:

```
$ aws datsync 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. Specify the following required options in the command:
  - For `storage-virtual-machine-arn`, specify the fully qualified Amazon Resource Name (ARN) of a storage virtual machine (SVM) in your file system where you want to copy data to or from.

This ARN includes the AWS Region where your file system resides, your AWS account, and the file system and SVM IDs.

- For `security-group-arns`, specify the ARNs of the Amazon EC2 security groups that provide access to the [network interfaces](#) of your file system's preferred subnet.

This includes the AWS Region where your Amazon EC2 instance resides, your AWS account, and your security group IDs. You can specify up to five security group ARNs.

For more information about security groups, see [File System Access Control with Amazon VPC](#) in the *Amazon FSx for NetApp ONTAP User Guide*.

- For `protocol`, configure the protocol that DataSync uses to access your file system's SVM.
  - For NFS, you can use the default configuration:

```
--protocol NFS={}
```

- For SMB, you must specify a user name and password that can access the SVM:

```
--protocol SMB={User=smb-user, Password=smb-password}
```

3. Run the command.

You get a response that shows the location that you just created.

```
{
  "LocationArn": "arn:aws:datasync:us-west-2:123456789012:location/loc-
  abcdef01234567890"
}
```

## Creating an AWS DataSync task with the AWS CLI

Once you create your source and destination locations, you can create your AWS DataSync task.

### Important

If you're planning to transfer data to or from an Amazon S3 location, review [how DataSync can affect your S3 request charges](#) and the [DataSync pricing page](#) before you begin.

### To create a task by using the AWS CLI

1. Create an Amazon CloudWatch Logs log group by using the following command.

```
aws logs create-log-group \
  --log-group-name your-log-group
```

2. Attach an IAM resource policy to your log group. For instructions on how to attach the policy, see [Allowing DataSync to upload logs to CloudWatch log groups](#).
3. Create a task by using the following command.

```
aws datasync create-task \
  --source-location-arn 'arn:aws:datasync: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
```

This command returns the Amazon Resource Name (ARN) for a task, similar to the one shown following.

```
{
  "TaskArn": "arn:aws:datsync:us-east-1:111222333444:task/
task-08de6e6697796f026"
}
```

**When creating a task that transfers data between AWS services in different Regions,** and the other location must be specified in a different Region (for example, to transfer data between us-east-1 and us-east-2), use DataSync in one of the Regions and create a task by using the following command.

```
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
```

Your task is created with the default configuration options. If you want to configure different options as part of your task creation, add the `--options` parameter to your `create-task` command. The following example shows how to specify different options. For a description of these options, see [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
```

When you create a task, you can configure the task to include or exclude specific files, folders, and objects. For more information, see [Specifying what AWS DataSync transfers by using filters](#). You can also schedule when you want your task to run. For more information, see [Scheduling your AWS DataSync task](#).

## Starting an AWS DataSync task with the AWS CLI

You can begin transferring your data with AWS DataSync by using the AWS CLI.

To start your DataSync task, you just need to specify the Amazon Resource Name (ARN) of the task you want to run. Here's an example `start-task-execution` command:

```
aws datasync start-task-execution \  
--task-arn 'arn:aws:datasync:region:account-id:task/task-id'
```

You can modify the task's settings for a specific task execution, as shown in the example following. For a description of these options, see [the section called "Options"](#). You also configure a specific task run to focus on specific files, folders, and objects to transfer. For more information, see [Specifying what AWS DataSync transfers by using filters](#).

The following example starts a task with a few settings that are different than the task's default settings:

```
aws datasync start-task-execution \  
--override-options VerifyMode=NONE,OverwriteMode=NEVER,PosixPermissions=NONE
```

The command returns an ARN for your task execution similar to the following example:

```
{  
  "TaskExecutionArn": "arn:aws:datasync:us-east-1:209870788375:task/  
task-08de6e6697796f026/execution/exec-04ce9d516d69bd52f"  
}
```

**Note**

Each agent can run a single task at a time.

## Filtering AWS DataSync resources

You can filter your AWS DataSync locations and tasks by using the `ListLocations` and `ListTasks` API operations in the AWS CLI. For example, retrieve a list of your most recent tasks.

### Parameters for filtering

You can use API filters to narrow down the list of resources returned by `ListTasks` and `ListLocations`. For example, to retrieve all of your Amazon S3 locations, you can use `ListLocations` with the filter name `LocationType S3` and operator `Equals`.

To filter API results, you must specify a filter name, operator, and value.

- **Name** – The name of the filter that's being used. Each API call supports a list of filters that are available for it (for example, `LocationType` for `ListLocations`).
- **Values** – The values that you want to filter for. For example, you might want to display only Amazon S3 locations.
- **Operator** – The operator that's used to compare filter values (for example, `Equals` or `Contains`).

The following table lists the available operators.

Operator	Key types
<code>Equals</code>	String, Number
<code>NotEquals</code>	String, Number
<code>LessThan</code>	Number
<code>LessThanOrEqualTo</code>	Number
<code>GreaterThan</code>	Number

Operator	Key types
GreaterThanOrEqualTo	Number
In	String
Contains	String
NotContains	String
BeginsWith	String

## Filtering by location

ListLocations supports the following filter names:

- **LocationType** – Filters on the location type:
  - SMB
  - NFS
  - HDFS
  - OBJECT\_STORAGE
  - S3
  - OUTPOST\_S3
  - FSX\_WINDOWS
  - FSX\_LUSTRE
  - FSX\_OPENZFS\_NFS
  - FSX\_ONTAP\_NFS
  - FSX\_ONTAP\_SMB
- **LocationUri** – Filters on the uniform resource identifier (URI) assigned to the location, as returned by the DescribeLocation\* API call (for example, `s3://bucket-name/your-prefix` for Amazon S3 locations).
- **CreationTime** – Filters on the time that the location was created. The input format is yyyy-MM-dd:mm:ss in Coordinated Universal Time (UTC).

The following AWS CLI example lists all locations of type Amazon S3 that have a location URI starting with the string "s3://DOC-EXAMPLE-BUCKET" and that were created at or after 2019-12-15 17:15:20 UTC.

```
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}]
```

This command returns output similar to the following.

```
{
  "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/"
    }
  ]
}
```

## Filtering by task

ListTasks supports the following filter names.

- **LocationId** – Filters on both source and destination locations on Amazon Resource Name (ARN) values.
- **CreationTime** – Filters on the time that the task was created. The input format is yyyy-MM-dd:mm:ss in UTC.

The following AWS CLI example shows the syntax when filtering on LocationId.

```
aws datasync list-tasks \
  --filters Name=LocationId,Values=arn:aws:datasync:us-east-1:your-account-id:location/your-location-id,Operator=Contains
```



The output of this command looks similar to the following.

```
{
  "Tasks": [
    {
      "TaskArn": "arn:aws:datsync:us-east-1:your-account-id:task/your-task-id",
      "Status": "AVAILABLE",
      "Name": "DOC-EXAMPLE-BUCKET"
    }
  ]
}
```

# AWS DataSync API

In addition to the AWS Management Console and AWS CLI, you can use the AWS DataSync API to configure and manage DataSync with the [AWS SDKs](#).

## Topics

- [Actions](#)
- [Data Types](#)
- [Common Errors](#)
- [Common Parameters](#)

## Actions

The following actions are supported:

- [AddStorageSystem](#)
- [CancelTaskExecution](#)
- [CreateAgent](#)
- [CreateLocationAzureBlob](#)
- [CreateLocationEfs](#)
- [CreateLocationFsxLustre](#)
- [CreateLocationFsxOntap](#)
- [CreateLocationFsxOpenZfs](#)
- [CreateLocationFsxWindows](#)
- [CreateLocationHdfs](#)
- [CreateLocationNfs](#)
- [CreateLocationObjectStorage](#)
- [CreateLocationS3](#)
- [CreateLocationSmb](#)
- [CreateTask](#)
- [DeleteAgent](#)
- [DeleteLocation](#)

- [DeleteTask](#)
- [DescribeAgent](#)
- [DescribeDiscoveryJob](#)
- [DescribeLocationAzureBlob](#)
- [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](#)
- [UpdateLocationAzureBlob](#)
- [UpdateLocationHdfs](#)
- [UpdateLocationNfs](#)
- [UpdateLocationObjectStorage](#)
- [UpdateLocationSmb](#)
- [UpdateStorageSystem](#)
- [UpdateTask](#)
- [UpdateTaskExecution](#)

## AddStorageSystem

Creates an AWS resource for an on-premises storage system that you want DataSync Discovery to collect information about.

### Request Syntax

```
{
  "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"
    }
  ]
}
```

### Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

#### AgentArns

Specifies the Amazon Resource Name (ARN) of the DataSync agent that connects to and reads from your on-premises storage system's management interface. You can only specify one ARN.

Type: Array of strings

Array Members: Fixed number of 1 item.

Length Constraints: Maximum length of 128.

Pattern: `^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}$`

Required: Yes

### ClientToken

Specifies a client token to make sure requests with this API operation are idempotent. If you don't specify a client token, DataSync generates one for you automatically.

Type: String

Pattern: `[a-f0-9]{8}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{12}`

Required: Yes

### CloudWatchLogGroupArn

Specifies the ARN of the Amazon CloudWatch log group for monitoring and logging discovery job events.

Type: String

Length Constraints: Maximum length of 562.

Pattern: `^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):logs:[a-z\-\0-9]+:[0-9]{12}:log-group:([^\:]*)(:\:)*$`

Required: No

### Credentials

Specifies the user name and password for accessing your on-premises storage system's management interface.

Type: [Credentials](#) object

Required: Yes

### Name

Specifies a familiar name for your on-premises storage system.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 256.

Pattern: `^\p{L}\p{M}\p{N}\s+=. _:@\|- ]+$`

Required: No

### ServerConfiguration

Specifies the server name and network port required to connect with the management interface of your on-premises storage system.

Type: [DiscoveryServerConfiguration](#) object

Required: Yes

### SystemType

Specifies the type of on-premises storage system that you want DataSync Discovery to collect information about.

#### Note

DataSync Discovery currently supports NetApp Fabric-Attached Storage (FAS) and All Flash FAS (AFF) systems running ONTAP 9.7 or later.

Type: String

Valid Values: NetAppONTAP

Required: Yes

### Tags

Specifies labels that help you categorize, filter, and search for your AWS resources. We recommend creating at least a name tag for your on-premises storage system.

Type: Array of [TagListEntry](#) objects

Array Members: Minimum number of 0 items. Maximum number of 50 items.

Required: No

## Response Syntax

```
{
```

```
"StorageSystemArn": "string"
}
```

## Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

### StorageSystemArn

The ARN of the on-premises storage system that you can use with DataSync Discovery.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^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}$`

## Errors

For information about the errors that are common to all actions, see [Common Errors](#).

### InternalException

This exception is thrown when an error occurs in the AWS DataSync service.

HTTP Status Code: 500

### InvalidRequestException

This exception is thrown when the client submits a malformed request.

HTTP Status Code: 400

## Examples

### Sample Request

The following example adds an on-premises storage system to 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"
  }
}
```

## Sample Response

A response returns the ARN of the on-premises storage system that you just added to DataSync Discovery.

```
{
  "StorageSystemArn": "arn:aws:datasync:us-east-1:111222333444:system/storage-system-
abcdef01234567890"
}
```

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)

- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

## CancelTaskExecution

Stops an AWS DataSync task execution that's in progress. The transfer of some files are abruptly interrupted. File contents that're transferred to the destination might be incomplete or inconsistent with the source files.

However, if you start a new task execution using the same task and allow it to finish, file content on the destination will be complete and consistent. This applies to other unexpected failures that interrupt a task execution. In all of these cases, DataSync successfully completes the transfer when you start the next task execution.

### Request Syntax

```
{  
  "TaskExecutionArn": "string"  
}
```

### Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

#### [TaskExecutionArn](#)

The Amazon Resource Name (ARN) of the task execution to stop.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^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}$`

Required: Yes

### Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

## Errors

For information about the errors that are common to all actions, see [Common Errors](#).

### InternalException

This exception is thrown when an error occurs in the AWS DataSync service.

HTTP Status Code: 500

### InvalidRequestException

This exception is thrown when the client submits a malformed request.

HTTP Status Code: 400

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

## CreateAgent

Activates an AWS DataSync agent that you've deployed in your storage environment. The activation process associates the agent with your AWS account.

If you haven't deployed an agent yet, see the following topics to learn more:

- [Agent requirements](#)
- [Create an agent](#)

### Note

If you're transferring between AWS storage services, you don't need a DataSync agent.

## Request Syntax

```
{
  "ActivationKey": "string",
  "AgentName": "string",
  "SecurityGroupArns": [ "string" ],
  "SubnetArns": [ "string" ],
  "Tags": [
    {
      "Key": "string",
      "Value": "string"
    }
  ],
  "VpcEndpointId": "string"
}
```

## Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

### ActivationKey

Specifies your DataSync agent's activation key. If you don't have an activation key, see [Activate your agent](#).

Type: String

Length Constraints: Maximum length of 29.

Pattern: `[A-Z0-9]{5}(-[A-Z0-9]{5}){4}`

Required: Yes

### AgentName

Specifies a name for your agent. You can see this name in the DataSync console.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 256.

Pattern: `^[a-zA-Z0-9\s+=. _:@/-]+$`

Required: No

### SecurityGroupArns

Specifies the Amazon Resource Name (ARN) of the security group that protects your task's [network interfaces](#) when [using a virtual private cloud \(VPC\) endpoint](#). You can only specify one ARN.

Type: Array of strings

Array Members: Fixed number of 1 item.

Length Constraints: Maximum length of 128.

Pattern: `^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]+$`

Required: No

### SubnetArns

Specifies the ARN of the subnet where you want to run your DataSync task when using a VPC endpoint. This is the subnet where DataSync creates and manages the [network interfaces](#) for your transfer. You can only specify one ARN.

Type: Array of strings

Array Members: Fixed number of 1 item.

Length Constraints: Maximum length of 128.

Pattern: `^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):ec2:[a-z\-\0-9]*:[0-9]{12}:subnet/.*$`

Required: No

## Tags

Specifies labels that help you categorize, filter, and search for your AWS resources. We recommend creating at least one tag for your agent.

Type: Array of [TagListEntry](#) objects

Array Members: Minimum number of 0 items. Maximum number of 50 items.

Required: No

## VpcEndpointId

Specifies the ID of the VPC endpoint that you want your agent to connect to. For example, a VPC endpoint ID looks like `vpce-01234d5aff67890e1`.

### Important

The VPC endpoint you use must include the DataSync service name (for example, `com.amazonaws.us-east-2.datasync`).

Type: String

Pattern: `^vpce-[0-9a-f]{17}$`

Required: No

## Response Syntax

```
{
  "AgentArn": "string"
}
```

## Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

### AgentArn

The ARN of the agent that you just activated. Use the [ListAgents](#) operation to return a list of agents in your AWS account and AWS Region.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^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}$`

## Errors

For information about the errors that are common to all actions, see [Common Errors](#).

### **InternalException**

This exception is thrown when an error occurs in the AWS DataSync service.

HTTP Status Code: 500

### **InvalidRequestException**

This exception is thrown when the client submits a malformed request.

HTTP Status Code: 400

## Examples

### **Sample Request**

The following example activates a DataSync agent and associates it with your AWS account.

```
{
  "ActivationKey": "AAAAA-1AAAA-BB1CC-33333-EEEEEE",
```



```
"AgentName": "MyAgent",
"Tags": [{
  "Key": "Job",
  "Value": "TransferJob-1"
}]
}
```

## Sample Response

The response returns the ARN of the activated agent.

```
{
  "AgentArn": "arn:aws:datsync:us-east-2:111222333444:agent/agent-0b0addbeef44baca3"
}
```

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

## CreateLocationAzureBlob

Creates a transfer *location* for a Microsoft Azure Blob Storage container. AWS DataSync can use this location as a transfer source or destination.

Before you begin, make sure you know [how DataSync accesses Azure Blob Storage](#) and works with [access tiers](#) and [blob types](#). You also need a [DataSync agent](#) that can connect to your container.

### Request Syntax

```
{
  "AccessTier": "string",
  "AgentArns": [ "string" ],
  "AuthenticationType": "string",
  "BlobType": "string",
  "ContainerUrl": "string",
  "SasConfiguration": {
    "Token": "string"
  },
  "Subdirectory": "string",
  "Tags": [
    {
      "Key": "string",
      "Value": "string"
    }
  ]
}
```

### Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

#### [AccessTier](#)

Specifies the access tier that you want your objects or files transferred into. This only applies when using the location as a transfer destination. For more information, see [Access tiers](#).

Type: String

Valid Values: HOT | COOL | ARCHIVE

Required: No

### AgentArns

Specifies the Amazon Resource Name (ARN) of the DataSync agent that can connect with your Azure Blob Storage container.

You can specify more than one agent. For more information, see [Using multiple agents for your transfer](#).

Type: Array of strings

Array Members: Minimum number of 1 item. Maximum number of 4 items.

Length Constraints: Maximum length of 128.

Pattern: `^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}$`

Required: Yes

### AuthenticationType

Specifies the authentication method DataSync uses to access your Azure Blob Storage. DataSync can access blob storage using a shared access signature (SAS).

Type: String

Valid Values: SAS

Required: Yes

### BlobType

Specifies the type of blob that you want your objects or files to be when transferring them into Azure Blob Storage. Currently, DataSync only supports moving data into Azure Blob Storage as block blobs. For more information on blob types, see the [Azure Blob Storage documentation](#).

Type: String

Valid Values: BLOCK

Required: No

### ContainerUrl

Specifies the URL of the Azure Blob Storage container involved in your transfer.

Type: String

Length Constraints: Maximum length of 325.

Pattern: `^https://[A-Za-z0-9](\\.|-)?[A-Za-z0-9]{0,252}/[a-z0-9](-?[a-z0-9]){2,62}$`

Required: Yes

### SasConfiguration

Specifies the SAS configuration that allows DataSync to access your Azure Blob Storage.

Type: [AzureBlobSasConfiguration](#) object

Required: No

### Subdirectory

Specifies path segments if you want to limit your transfer to a virtual directory in your container (for example, /my/images).

Type: String

Length Constraints: Maximum length of 1024.

Pattern: `^[\\p{L}\\p{M}\\p{Z}\\p{S}\\p{N}\\p{P}\\p{C}]*$`

Required: No

### Tags

Specifies labels that help you categorize, filter, and search for your AWS resources. We recommend creating at least a name tag for your transfer location.

Type: Array of [TagListEntry](#) objects

Array Members: Minimum number of 0 items. Maximum number of 50 items.

Required: No

## Response Syntax

```
{
```

```
"LocationArn": "string"  
}
```

## Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

### LocationArn

The ARN of the Azure Blob Storage transfer location that you created.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^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}$`

## Errors

For information about the errors that are common to all actions, see [Common Errors](#).

### **InternalException**

This exception is thrown when an error occurs in the AWS DataSync service.

HTTP Status Code: 500

### **InvalidRequestException**

This exception is thrown when the client submits a malformed request.

HTTP Status Code: 400

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)

- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

## CreateLocationEfs

Creates a transfer *location* for an Amazon EFS file system. AWS DataSync can use this location as a source or destination for transferring data.

Before you begin, make sure that you understand how DataSync [accesses Amazon EFS file systems](#).

### Request Syntax

```
{
  "AccessPointArn": "string",
  "Ec2Config": {
    "SecurityGroupArns": [ "string" ],
    "SubnetArn": "string"
  },
  "EfsFilesystemArn": "string",
  "FileSystemAccessRoleArn": "string",
  "InTransitEncryption": "string",
  "Subdirectory": "string",
  "Tags": [
    {
      "Key": "string",
      "Value": "string"
    }
  ]
}
```

### Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

#### [AccessPointArn](#)

Specifies the Amazon Resource Name (ARN) of the access point that DataSync uses to access the Amazon EFS file system.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^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}$`

Required: No

### Ec2Config

Specifies the subnet and security groups DataSync uses to access your Amazon EFS file system.

Type: [Ec2Config](#) object

Required: Yes

### EfsFilesystemArn

Specifies the ARN for the Amazon EFS file system.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):elasticfilesystem:[a-z\-\0-9]*:[0-9]{12}:file-system/fs-.*$`

Required: Yes

### FileSystemAccessRoleArn

Specifies an AWS Identity and Access Management (IAM) role that DataSync assumes when mounting the Amazon EFS file system.

Type: String

Length Constraints: Maximum length of 2048.

Pattern: `^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):iam::[0-9]{12}:role/.*$`

Required: No

### InTransitEncryption

Specifies whether you want DataSync to use Transport Layer Security (TLS) 1.2 encryption when it copies data to or from the Amazon EFS file system.



If you specify an access point using `AccessPointArn` or an IAM role using `FileSystemAccessRoleArn`, you must set this parameter to `TLS1_2`.

Type: String

Valid Values: NONE | TLS1\_2

Required: No

## Subdirectory

Specifies a mount path for your Amazon EFS file system. This is where DataSync reads or writes data (depending on if this is a source or destination location). By default, DataSync uses the root directory, but you can also include subdirectories.

### Note

You must specify a value with forward slashes (for example, `/path/to/folder`).

Type: String

Length Constraints: Maximum length of 4096.

Pattern: `^[a-zA-Z0-9_\-\\+\.\/\(\)\p{Zs}]*$`

Required: No

## Tags

Specifies the key-value pair that represents a tag that you want to add to the resource. The value can be an empty string. This value helps you manage, filter, and search for your resources. We recommend that you create a name tag for your location.

Type: Array of [TagListEntry](#) objects

Array Members: Minimum number of 0 items. Maximum number of 50 items.

Required: No

## Response Syntax

```
{
```

```
"LocationArn": "string"  
}
```

## Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

### LocationArn

The Amazon Resource Name (ARN) of the Amazon EFS file system location that you create.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^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}$`

## Errors

For information about the errors that are common to all actions, see [Common Errors](#).

### **InternalException**

This exception is thrown when an error occurs in the AWS DataSync service.

HTTP Status Code: 500

### **InvalidRequestException**

This exception is thrown when the client submits a malformed request.

HTTP Status Code: 400

## Examples

### **Sample Request**

The following example creates a location for an Amazon EFS file system.

```
{
  "Ec2Config": {
    "SubnetArn": "arn:aws:ec2:us-east-2:111222333444:subnet/
subnet-1234567890abcdef1",
    "SecurityGroupArns": [
      "arn:aws:ec2:us-east-2:111222333444: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"
  }]
}
```

### Sample Request: Creating a location for a restricted Amazon EFS file system

The following example creates a location for an Amazon EFS file system with restricted access. In this kind of scenario, you might have to specify values for `AccessPointArn`, `FileSystemAccessRoleArn`, and `InTransitEncryption` in your request.

```
{
  "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"  
  }]  
}
```

## Sample Response

A response returns the location ARN of the Amazon EFS file system.

```
{  
  "LocationArn": "arn:aws:datsync:us-east-2:111222333444:location/  
loc-12abcdef012345678"  
}
```

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

## CreateLocationFsxLustre

Creates a transfer *location* for an Amazon FSx for Lustre file system. AWS DataSync can use this location as a source or destination for transferring data.

Before you begin, make sure that you understand how DataSync [accesses FSx for Lustre file systems](#).

### Request Syntax

```
{
  "FsxFilesystemArn": "string",
  "SecurityGroupArns": [ "string" ],
  "Subdirectory": "string",
  "Tags": [
    {
      "Key": "string",
      "Value": "string"
    }
  ]
}
```

### Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

#### [FsxFilesystemArn](#)

The Amazon Resource Name (ARN) for the FSx for Lustre file system.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):fsx:[a-z\-\0-9]*:[0-9]{12}:file-system/fs-.*$`

Required: Yes

## SecurityGroupArns

The Amazon Resource Names (ARNs) of the security groups that are used to configure the FSx for Lustre file system.

Type: Array of strings

Array Members: Minimum number of 1 item. Maximum number of 5 items.

Length Constraints: Maximum length of 128.

Pattern: `^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]+$`

Required: Yes

## Subdirectory

A subdirectory in the location's path. This subdirectory in the FSx for Lustre file system is used to read data from the FSx for Lustre source location or write data to the FSx for Lustre destination.

Type: String

Length Constraints: Maximum length of 4096.

Pattern: `^[a-zA-Z0-9_\-\+\.\^(\)\$\p{Zs}]+$`

Required: No

## Tags

The key-value pair that represents a tag that you want to add to the resource. The value can be an empty string. This value helps you manage, filter, and search for your resources. We recommend that you create a name tag for your location.

Type: Array of [TagListEntry](#) objects

Array Members: Minimum number of 0 items. Maximum number of 50 items.

Required: No

## Response Syntax

```
{
```

```
"LocationArn": "string"  
}
```

## Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

### LocationArn

The Amazon Resource Name (ARN) of the FSx for Lustre file system location that's created.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^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}$`

## Errors

For information about the errors that are common to all actions, see [Common Errors](#).

### **InternalException**

This exception is thrown when an error occurs in the AWS DataSync service.

HTTP Status Code: 500

### **InvalidRequestException**

This exception is thrown when the client submits a malformed request.

HTTP Status Code: 400

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)

- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)



## CreateLocationFsxOntap

Creates a transfer *location* for an Amazon FSx for NetApp ONTAP file system. AWS DataSync can use this location as a source or destination for transferring data.

Before you begin, make sure that you understand how DataSync [accesses FSx for ONTAP file systems](#).

### Request Syntax

```
{
  "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"
    }
  ]
}
```

### Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

## Protocol

Specifies the data transfer protocol that AWS DataSync uses to access your Amazon FSx file system.

Type: [FsxProtocol](#) object

Required: Yes

## SecurityGroupArns

Specifies the Amazon EC2 security groups that provide access to your file system's preferred subnet.

The security groups must allow outbound traffic on the following ports (depending on the protocol you use):

- **Network File System (NFS):** TCP ports 111, 635, and 2049
- **Server Message Block (SMB):** TCP port 445

Your file system's security groups must also allow inbound traffic on the same ports.

Type: Array of strings

Array Members: Minimum number of 1 item. Maximum number of 5 items.

Length Constraints: Maximum length of 128.

Pattern: `^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]+$`

Required: Yes

## StorageVirtualMachineArn

Specifies the ARN of the storage virtual machine (SVM) in your file system where you want to copy data to or from.

Type: String

Length Constraints: Maximum length of 162.

Pattern: `^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,}$`

Required: Yes

## Subdirectory

Specifies a path to the file share in the SVM where you'll copy your data.

You can specify a junction path (also known as a mount point), qtree path (for NFS file shares), or share name (for SMB file shares). For example, your mount path might be `/vol1`, `/vol1/tree1`, or `/share1`.

### Note

Don't specify a junction path in the SVM's root volume. For more information, see [Managing FSx for ONTAP storage virtual machines](#) in the *Amazon FSx for NetApp ONTAP User Guide*.

Type: String

Length Constraints: Maximum length of 255.

Pattern: `^[^\u0000\u0085\u2028\u2029\r\n]{1,255}$`

Required: No

## Tags

Specifies labels that help you categorize, filter, and search for your AWS resources. We recommend creating at least a name tag for your location.

Type: Array of [TagListEntry](#) objects

Array Members: Minimum number of 0 items. Maximum number of 50 items.

Required: No

## Response Syntax

```
{
  "LocationArn": "string"
}
```

## Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

### LocationArn

Specifies the ARN of the FSx for ONTAP file system location that you create.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^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}$`

## Errors

For information about the errors that are common to all actions, see [Common Errors](#).

### **InternalException**

This exception is thrown when an error occurs in the AWS DataSync service.

HTTP Status Code: 500

### **InvalidRequestException**

This exception is thrown when the client submits a malformed request.

HTTP Status Code: 400

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)

- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

## CreateLocationFsxOpenZfs

Creates a transfer *location* for an Amazon FSx for OpenZFS file system. AWS DataSync can use this location as a source or destination for transferring data.

Before you begin, make sure that you understand how DataSync [accesses FSx for OpenZFS file systems](#).

### Note

Request parameters related to SMB aren't supported with the CreateLocationFsxOpenZfs operation.

## Request Syntax

```
{
  "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"
    }
  ]
}
```

```
}
```

## Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

### FsxFilesystemArn

The Amazon Resource Name (ARN) of the FSx for OpenZFS file system.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):fsx:[a-z\-\0-9]*:[0-9]{12}:file-system/fs-.*$`

Required: Yes

### Protocol

The type of protocol that AWS DataSync uses to access your file system.

Type: [FsxProtocol](#) object

Required: Yes

### SecurityGroupArns

The ARNs of the security groups that are used to configure the FSx for OpenZFS file system.

Type: Array of strings

Array Members: Minimum number of 1 item. Maximum number of 5 items.

Length Constraints: Maximum length of 128.

Pattern: `^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]+$`

Required: Yes

## Subdirectory

A subdirectory in the location's path that must begin with /fsx. DataSync uses this subdirectory to read or write data (depending on whether the file system is a source or destination location).

Type: String

Length Constraints: Maximum length of 4096.

Pattern: `^[^\u0000\u0085\u2028\u2029\r\n]{1,4096}$`

Required: No

## Tags

The key-value pair that represents a tag that you want to add to the resource. The value can be an empty string. This value helps you manage, filter, and search for your resources. We recommend that you create a name tag for your location.

Type: Array of [TagListEntry](#) objects

Array Members: Minimum number of 0 items. Maximum number of 50 items.

Required: No

## Response Syntax

```
{
  "LocationArn": "string"
}
```

## Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

### LocationArn

The ARN of the FSx for OpenZFS file system location that you created.



Type: String

Length Constraints: Maximum length of 128.

Pattern: `^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}$`

## Errors

For information about the errors that are common to all actions, see [Common Errors](#).

### InternalException

This exception is thrown when an error occurs in the AWS DataSync service.

HTTP Status Code: 500

### InvalidRequestException

This exception is thrown when the client submits a malformed request.

HTTP Status Code: 400

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

## CreateLocationFsxWindows

Creates a transfer *location* for an Amazon FSx for Windows File Server file system. AWS DataSync can use this location as a source or destination for transferring data.

Before you begin, make sure that you understand how DataSync [accesses FSx for Windows File Server file systems](#).

### Request Syntax

```
{
  "Domain": "string",
  "FsxFilesystemArn": "string",
  "Password": "string",
  "SecurityGroupArns": [ "string" ],
  "Subdirectory": "string",
  "Tags": [
    {
      "Key": "string",
      "Value": "string"
    }
  ],
  "User": "string"
}
```

### Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

#### Domain

Specifies the name of the Microsoft Active Directory domain that the FSx for Windows File Server file system belongs to.

If you have multiple Active Directory domains in your environment, configuring this parameter makes sure that DataSync connects to the right file system.

Type: String

Length Constraints: Maximum length of 253.

Pattern: `^[A-Za-z0-9](\.|-+)?[A-Za-z0-9]{0,252}$`

Required: No

### FsxFilesystemArn

Specifies the Amazon Resource Name (ARN) for the FSx for Windows File Server file system.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):fsx:[a-z\-\0-9]*:[0-9]{12}:file-system/fs-.*$`

Required: Yes

### Password

Specifies the password of the user with the permissions to mount and access the files, folders, and file metadata in your FSx for Windows File Server file system.

Type: String

Length Constraints: Maximum length of 104.

Pattern: `^.{0,104}$`

Required: Yes

### SecurityGroupArns

Specifies the ARNs of the security groups that provide access to your file system's preferred subnet.

#### **Note**

If you choose a security group that doesn't allow connections from within itself, do one of the following:

- Configure the security group to allow it to communicate within itself.
- Choose a different security group that can communicate with the mount target's security group.

Type: Array of strings

Array Members: Minimum number of 1 item. Maximum number of 5 items.

Length Constraints: Maximum length of 128.

Pattern: `^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]+$`

Required: Yes

## Subdirectory

Specifies a mount path for your file system using forward slashes. This is where DataSync reads or writes data (depending on if this is a source or destination location).

Type: String

Length Constraints: Maximum length of 4096.

Pattern: `^[a-zA-Z0-9_\-\+\.\^(\)\$\p{Zs}]+$`

Required: No

## Tags

Specifies labels that help you categorize, filter, and search for your AWS resources. We recommend creating at least a name tag for your location.

Type: Array of [TagListEntry](#) objects

Array Members: Minimum number of 0 items. Maximum number of 50 items.

Required: No

## User

Specifies the user with the permissions to mount and access the files, folders, and file metadata in your FSx for Windows File Server file system.

For information about choosing a user with the right level of access for your transfer, see [required permissions](#) for FSx for Windows File Server locations.

Type: String

Length Constraints: Maximum length of 104.

Pattern: `^[^\x5B\x5D\\/:;|=,+*?]{1,104}$`

Required: Yes

## Response Syntax

```
{
  "LocationArn": "string"
}
```

## Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

### LocationArn

The ARN of the FSx for Windows File Server file system location you created.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^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}$`

## Errors

For information about the errors that are common to all actions, see [Common Errors](#).

### InternalException

This exception is thrown when an error occurs in the AWS DataSync service.

HTTP Status Code: 500

### InvalidRequestException

This exception is thrown when the client submits a malformed request.

HTTP Status Code: 400

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

## CreateLocationHdfs

Creates a transfer *location* for a Hadoop Distributed File System (HDFS). AWS DataSync can use this location as a source or destination for transferring data.

Before you begin, make sure that you understand how DataSync [accesses HDFS clusters](#).

### Request Syntax

```
{
  "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"
    }
  ]
}
```

### Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

## AgentArns

The Amazon Resource Names (ARNs) of the agents that are used to connect to the HDFS cluster.

Type: Array of strings

Array Members: Minimum number of 1 item. Maximum number of 4 items.

Length Constraints: Maximum length of 128.

Pattern: `^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}$`

Required: Yes

## AuthenticationType

The type of authentication used to determine the identity of the user.

Type: String

Valid Values: SIMPLE | KERBEROS

Required: Yes

## BlockSize

The size of data blocks to write into the HDFS cluster. The block size must be a multiple of 512 bytes. The default block size is 128 mebibytes (MiB).

Type: Integer

Valid Range: Minimum value of 1048576. Maximum value of 1073741824.

Required: No

## KerberosKeytab

The Kerberos key table (keytab) that contains mappings between the defined Kerberos principal and the encrypted keys. You can load the keytab from a file by providing the file's address. If you're using the AWS CLI, it performs base64 encoding for you. Otherwise, provide the base64-encoded text.

### Note

If KERBEROS is specified for AuthenticationType, this parameter is required.



Type: Base64-encoded binary data object

Length Constraints: Maximum length of 65536.

Required: No

### KerberosKrb5Conf

The `krb5.conf` file that contains the Kerberos configuration information. You can load the `krb5.conf` file by providing the file's address. If you're using the AWS CLI, it performs the base64 encoding for you. Otherwise, provide the base64-encoded text.

#### Note

If KERBEROS is specified for `AuthenticationType`, this parameter is required.

Type: Base64-encoded binary data object

Length Constraints: Maximum length of 131072.

Required: No

### KerberosPrincipal

The Kerberos principal with access to the files and folders on the HDFS cluster.

#### Note

If KERBEROS is specified for `AuthenticationType`, this parameter is required.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 256.

Pattern: `^\.+`

Required: No

### KmsKeyProviderUri

The URI of the HDFS cluster's Key Management Server (KMS).

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `^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$`

Required: No

## NameNodes

The NameNode that manages the HDFS namespace. The NameNode performs operations such as opening, closing, and renaming files and directories. The NameNode contains the information to map blocks of data to the DataNodes. You can use only one NameNode.

Type: Array of [HdfsNameNode](#) objects

Array Members: Minimum number of 1 item.

Required: Yes

## QopConfiguration

The Quality of Protection (QOP) configuration specifies the Remote Procedure Call (RPC) and data transfer protection settings configured on the Hadoop Distributed File System (HDFS) cluster. If `QopConfiguration` isn't specified, `RpcProtection` and `DataTransferProtection` default to `PRIVACY`. If you set `RpcProtection` or `DataTransferProtection`, the other parameter assumes the same value.

Type: [QopConfiguration](#) object

Required: No

## ReplicationFactor

The number of DataNodes to replicate the data to when writing to the HDFS cluster. By default, data is replicated to three DataNodes.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 512.

Required: No

## SimpleUser

The user name used to identify the client on the host operating system.

### **Note**

If SIMPLE is specified for `AuthenticationType`, this parameter is required.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 256.

Pattern: `^[_.A-Za-z0-9][_.A-Za-z0-9]*$`

Required: No

## Subdirectory

A subdirectory in the HDFS cluster. This subdirectory is used to read data from or write data to the HDFS cluster. If the subdirectory isn't specified, it will default to `/`.

Type: String

Length Constraints: Maximum length of 4096.

Pattern: `^[a-zA-Z0-9_-\+\.\^(\)\$\p{Zs}]+$`

Required: No

## Tags

The key-value pair that represents the tag that you want to add to the location. The value can be an empty string. We recommend using tags to name your resources.

Type: Array of [TagListEntry](#) objects

Array Members: Minimum number of 0 items. Maximum number of 50 items.

Required: No

## **Response Syntax**

```
{
```

```
"LocationArn": "string"  
}
```

## Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

### LocationArn

The ARN of the source HDFS cluster location that's created.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^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}$`

## Errors

For information about the errors that are common to all actions, see [Common Errors](#).

### **InternalException**

This exception is thrown when an error occurs in the AWS DataSync service.

HTTP Status Code: 500

### **InvalidRequestException**

This exception is thrown when the client submits a malformed request.

HTTP Status Code: 400

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)

- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

## CreateLocationNfs

Creates a transfer *location* for a Network File System (NFS) file server. AWS DataSync can use this location as a source or destination for transferring data.

Before you begin, make sure that you understand how DataSync [accesses NFS file servers](#).

### Note

If you're copying data to or from an AWS Snowcone device, you can also use `CreateLocationNfs` to create your transfer location. For more information, see [Configuring transfers with Snowcone](#).

## Request Syntax

```
{
  "MountOptions": {
    "Version": "string"
  },
  "OnPremConfig": {
    "AgentArns": [ "string" ]
  },
  "ServerHostname": "string",
  "Subdirectory": "string",
  "Tags": [
    {
      "Key": "string",
      "Value": "string"
    }
  ]
}
```

## Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

### MountOptions

Specifies the options that DataSync can use to mount your NFS file server.

Type: [NfsMountOptions](#) object

Required: No

### [OnPremConfig](#)

Specifies the Amazon Resource Name (ARN) of the DataSync agent that want to connect to your NFS file server.

You can specify more than one agent. For more information, see [Using multiple agents for transfers](#).

Type: [OnPremConfig](#) object

Required: Yes

### [ServerHostname](#)

Specifies the Domain Name System (DNS) name or IP version 4 address of the NFS file server that your DataSync agent connects to.

Type: String

Length Constraints: Maximum length of 255.

Pattern: `^(([a-zA-Z0-9\-\ ]*[a-zA-Z0-9])\.)*([A-Za-z0-9\-\ ]*[A-Za-z0-9]))$`

Required: Yes

### [Subdirectory](#)

Specifies the export path in your NFS file server that you want DataSync to mount.

This path (or a subdirectory of the path) is where DataSync transfers data to or from. For information on configuring an export for DataSync, see [Accessing NFS file servers](#).

Type: String

Length Constraints: Maximum length of 4096.

Pattern: `^[a-zA-Z0-9_\-\+\.\ /(\)\p{Zs}]+$`

Required: Yes

## Tags

Specifies labels that help you categorize, filter, and search for your AWS resources. We recommend creating at least a name tag for your location.

Type: Array of [TagListEntry](#) objects

Array Members: Minimum number of 0 items. Maximum number of 50 items.

Required: No

## Response Syntax

```
{
  "LocationArn": "string"
}
```

## Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

### [LocationArn](#)

The ARN of the transfer location that you created for your NFS file server.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^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}$`

## Errors

For information about the errors that are common to all actions, see [Common Errors](#).

### InternalException

This exception is thrown when an error occurs in the AWS DataSync service.



HTTP Status Code: 500

## InvalidRequestException

This exception is thrown when the client submits a malformed request.

HTTP Status Code: 400

## Examples

### Example

The following example creates a DataSync transfer location for an NFS file server.

### Sample Request

```
{
  "MountOptions": {
    "Version": : "NFS4_0"
  },
  "OnPremConfig": {
    "AgentArn": [ "arn:aws:datsync:us-east-2:111222333444:agent/
agent-0b0addbeef44b3nfs" ]
  },

  "ServerHostname": "MyServer@amazon.com",
  "Subdirectory": "/MyFolder",
  "Tags": [
    {
      "Key": "Name",
      "Value": "FileSystem-1"
    }
  ]
}
```

### Example

The response returns the ARN of the NFS location.

### Sample Response

```
{
```

```
"LocationArn": "arn:aws:datsync:us-east-2:111222333444:location/loc-07db7abfc326c50aa"
}
```

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

## CreateLocationObjectStorage

Creates a transfer *location* for an object storage system. AWS DataSync can use this location as a source or destination for transferring data.

Before you begin, make sure that you understand the [prerequisites](#) for DataSync to work with object storage systems.

### Request Syntax

```
{
  "AccessKey": "string",
  "AgentArns": [ "string" ],
  "BucketName": "string",
  "SecretKey": "string",
  "ServerCertificate": blob,
  "ServerHostname": "string",
  "ServerPort": number,
  "ServerProtocol": "string",
  "Subdirectory": "string",
  "Tags": [
    {
      "Key": "string",
      "Value": "string"
    }
  ]
}
```

### Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

#### [AccessKey](#)

Specifies the access key (for example, a user name) if credentials are required to authenticate with the object storage server.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 200.

Pattern: `^.*$`

Required: No

### AgentArns

Specifies the Amazon Resource Names (ARNs) of the DataSync agents that can securely connect with your location.

Type: Array of strings

Array Members: Minimum number of 1 item. Maximum number of 4 items.

Length Constraints: Maximum length of 128.

Pattern: `^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}$`

Required: Yes

### BucketName

Specifies the name of the object storage bucket involved in the transfer.

Type: String

Length Constraints: Minimum length of 3. Maximum length of 63.

Pattern: `^[a-zA-Z0-9_\\-\\+\\.\\/\\(\\)\\$\\p{Zs}]+$`

Required: Yes

### SecretKey

Specifies the secret key (for example, a password) if credentials are required to authenticate with the object storage server.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 200.

Pattern: `^.*$`

Required: No

## ServerCertificate

Specifies a certificate chain for DataSync to authenticate with your object storage system if the system uses a private or self-signed certificate authority (CA). You must specify a single `.pem` file with a full certificate chain (for example, `file:///home/user/.ssh/object_storage_certificates.pem`).

The certificate chain might include:

- The object storage system's certificate
- All intermediate certificates (if there are any)
- The root certificate of the signing CA

You can concatenate your certificates into a `.pem` file (which can be up to 32768 bytes before base64 encoding). The following example `cat` command creates an `object_storage_certificates.pem` file that includes three certificates:

```
cat object_server_certificate.pem intermediate_certificate.pem  
ca_root_certificate.pem > object_storage_certificates.pem
```

To use this parameter, configure `ServerProtocol` to `HTTPS`.

Type: Base64-encoded binary data object

Length Constraints: Maximum length of 32768.

Required: No

## ServerHostname

Specifies the domain name or IP address of the object storage server. A DataSync agent uses this hostname to mount the object storage server in a network.

Type: String

Length Constraints: Maximum length of 255.

Pattern: `^(([a-zA-Z0-9\-\ ]*[a-zA-Z0-9])\.)*([A-Za-z0-9\-\ ]*[A-Za-z0-9])$`

Required: Yes

## ServerPort

Specifies the port that your object storage server accepts inbound network traffic on (for example, port 443).

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 65536.

Required: No

### ServerProtocol

Specifies the protocol that your object storage server uses to communicate.

Type: String

Valid Values: HTTPS | HTTP

Required: No

### Subdirectory

Specifies the object prefix for your object storage server. If this is a source location, DataSync only copies objects with this prefix. If this is a destination location, DataSync writes all objects with this prefix.

Type: String

Length Constraints: Maximum length of 4096.

Pattern: `^[a-zA-Z0-9_\-\.\/\(\)\p{Zs}]*$`

Required: No

### Tags

Specifies the key-value pair that represents a tag that you want to add to the resource. Tags can help you manage, filter, and search for your resources. We recommend creating a name tag for your location.

Type: Array of [TagListEntry](#) objects

Array Members: Minimum number of 0 items. Maximum number of 50 items.

Required: No

## Response Syntax

```
{
```

```
"LocationArn": "string"  
}
```

## Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

### LocationArn

Specifies the ARN of the object storage system location that you create.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^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}$`

## Errors

For information about the errors that are common to all actions, see [Common Errors](#).

### **InternalException**

This exception is thrown when an error occurs in the AWS DataSync service.

HTTP Status Code: 500

### **InvalidRequestException**

This exception is thrown when the client submits a malformed request.

HTTP Status Code: 400

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)

- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)



## CreateLocationS3

Creates a transfer *location* for an Amazon S3 bucket. AWS DataSync can use this location as a source or destination for transferring data.

### Important

Before you begin, make sure that you read the following topics:

- [Storage class considerations with Amazon S3 locations](#)
- [Evaluating S3 request costs when using DataSync](#)

For more information, see [Configuring transfers with Amazon S3](#).

## Request Syntax

```
{
  "AgentArns": [ "string" ],
  "S3BucketArn": "string",
  "S3Config": {
    "BucketAccessRoleArn": "string"
  },
  "S3StorageClass": "string",
  "Subdirectory": "string",
  "Tags": [
    {
      "Key": "string",
      "Value": "string"
    }
  ]
}
```

## Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

## AgentArns

(Amazon S3 on Outposts only) Specifies the Amazon Resource Name (ARN) of the DataSync agent on your Outpost.

For more information, see [Deploy your DataSync agent on AWS Outposts](#).

Type: Array of strings

Array Members: Minimum number of 1 item. Maximum number of 4 items.

Length Constraints: Maximum length of 128.

Pattern: `^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}$`

Required: No

## S3BucketArn

Specifies the ARN of the S3 bucket that you want to use as a location. (When creating your DataSync task later, you specify whether this location is a transfer source or destination.)

If your S3 bucket is located on an AWS Outposts resource, you must specify an Amazon S3 access point. For more information, see [Managing data access with Amazon S3 access points](#) in the *Amazon S3 User Guide*.

Type: String

Length Constraints: Maximum length of 156.

Pattern: `^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):(s3|s3-outposts):[a-z\\-0-9]*:[0-9]*:.*$`

Required: Yes

## S3Config

Specifies the Amazon Resource Name (ARN) of the AWS Identity and Access Management (IAM) role that DataSync uses to access your S3 bucket.

For more information, see [Accessing S3 buckets](#).

Type: [S3Config](#) object

Required: Yes

### S3StorageClass

Specifies the storage class that you want your objects to use when Amazon S3 is a transfer destination.

For buckets in AWS Regions, the storage class defaults to STANDARD. For buckets on AWS Outposts, the storage class defaults to OUTPOSTS.

For more information, see [Storage class considerations with Amazon S3 transfers](#).

Type: String

Valid Values: STANDARD | STANDARD\_IA | ONEZONE\_IA | INTELLIGENT\_TIERING | GLACIER | DEEP\_ARCHIVE | OUTPOSTS | GLACIER\_INSTANT\_RETRIEVAL

Required: No

### Subdirectory

Specifies a prefix in the S3 bucket that DataSync reads from or writes to (depending on whether the bucket is a source or destination location).

#### Note

DataSync can't transfer objects with a prefix that begins with a slash (/) or includes //, /./, or /../ patterns. For example:

- /photos
- photos//2006/January
- photos/./2006/February
- photos/../2006/March

Type: String

Length Constraints: Maximum length of 4096.

Pattern: `^[a-zA-Z0-9_\-\\+\\.\/\(\)\p{Zs}]*$`

Required: No

## Tags

Specifies labels that help you categorize, filter, and search for your AWS resources. We recommend creating at least a name tag for your transfer location.

Type: Array of [TagListEntry](#) objects

Array Members: Minimum number of 0 items. Maximum number of 50 items.

Required: No

## Response Syntax

```
{
  "LocationArn": "string"
}
```

## Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

### [LocationArn](#)

The ARN of the S3 location that you created.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^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}$`

## Errors

For information about the errors that are common to all actions, see [Common Errors](#).

### InternalException

This exception is thrown when an error occurs in the AWS DataSync service.

HTTP Status Code: 500

## InvalidRequestException

This exception is thrown when the client submits a malformed request.

HTTP Status Code: 400

## Examples

### Step 1. Allow to assume the IAM role required to write to the bucket

The following example shows the simplest policy that grants the required permissions for AWS DataSync to access a destination Amazon S3 bucket, followed by an IAM role to which the `create-location-s3-iam-role` policy has been attached.

```
{
  "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"
        }
      }
    ]
  }
}
```



```
}
```

## Sample Response

```
{  
  "LocationArn": "arn:aws:datsync:us-east-2:111222333444:location/  
loc-07db7abfc326c50s3"  
}
```

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

## CreateLocationSmb

Creates a transfer *location* for a Server Message Block (SMB) file server. AWS DataSync can use this location as a source or destination for transferring data.

Before you begin, make sure that you understand how DataSync [accesses SMB file servers](#).

### Request Syntax

```
{
  "AgentArns": [ "string" ],
  "Domain": "string",
  "MountOptions": {
    "Version": "string"
  },
  "Password": "string",
  "ServerHostname": "string",
  "Subdirectory": "string",
  "Tags": [
    {
      "Key": "string",
      "Value": "string"
    }
  ],
  "User": "string"
}
```

### Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

#### [AgentArns](#)

Specifies the DataSync agent (or agents) which you want to connect to your SMB file server. You specify an agent by using its Amazon Resource Name (ARN).

Type: Array of strings

Array Members: Minimum number of 1 item. Maximum number of 4 items.

Length Constraints: Maximum length of 128.



Pattern: `^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}$`

Required: Yes

### Domain

Specifies the name of the Active Directory domain that your SMB file server belongs to.

If you have multiple Active Directory domains in your environment, configuring this parameter makes sure that DataSync connects to the right file server.

Type: String

Length Constraints: Maximum length of 253.

Pattern: `^[A-Za-z0-9](\\.|-+)?[A-Za-z0-9]{0,252}$`

Required: No

### MountOptions

Specifies the version of the SMB protocol that DataSync uses to access your SMB file server.

Type: [SmbMountOptions](#) object

Required: No

### Password

Specifies the password of the user who can mount your SMB file server and has permission to access the files and folders involved in your transfer.

For more information, see [required permissions](#) for SMB locations.

Type: String

Length Constraints: Maximum length of 104.

Pattern: `^.{0,104}$`

Required: Yes

### ServerHostname

Specifies the Domain Name Service (DNS) name or IP address of the SMB file server that your DataSync agent will mount.

**Note**

You can't specify an IP version 6 (IPv6) address.

Type: String

Length Constraints: Maximum length of 255.

Pattern: `^(([a-zA-Z0-9\-\ ]*[a-zA-Z0-9])\.)*([A-Za-z0-9\-\ ]*[A-Za-z0-9])$`

Required: Yes

**Subdirectory**

Specifies the name of the share exported by your SMB file server where DataSync will read or write data. You can include a subdirectory in the share path (for example, `/path/to/subdirectory`). Make sure that other SMB clients in your network can also mount this path.

To copy all data in the subdirectory, DataSync must be able to mount the SMB share and access all of its data. For more information, see [required permissions](#) for SMB locations.

Type: String

Length Constraints: Maximum length of 4096.

Pattern: `^[a-zA-Z0-9_\-\+\.\^(\)\$\p{Zs}]+$`

Required: Yes

**Tags**

Specifies labels that help you categorize, filter, and search for your AWS resources. We recommend creating at least a name tag for your location.

Type: Array of [TagListEntry](#) objects

Array Members: Minimum number of 0 items. Maximum number of 50 items.

Required: No

**User**

Specifies the user that can mount and access the files, folders, and file metadata in your SMB file server.

For information about choosing a user with the right level of access for your transfer, see [required permissions](#) for SMB locations.

Type: String

Length Constraints: Maximum length of 104.

Pattern: `^[^\x5B\x5D\\\/:;|=,+*?]{1,104}$`

Required: Yes

## Response Syntax

```
{  
  "LocationArn": "string"  
}
```

## Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

### [LocationArn](#)

The ARN of the SMB location that you created.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^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}$`

## Errors

For information about the errors that are common to all actions, see [Common Errors](#).

### InternalException

This exception is thrown when an error occurs in the AWS DataSync service.

HTTP Status Code: 500

## InvalidRequestException

This exception is thrown when the client submits a malformed request.

HTTP Status Code: 400

## Examples

### Sample Request

The following example creates a location for an SMB file server.

```
{
  "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"
}
```

### Sample Response

A response returns the location ARN of your SMB file server.

```
{
  "LocationArn": "arn:aws:datasync:us-east-1:111222333444:location/loc-0f01451b140b2af49"
```

```
}
```

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

## CreateTask

Configures a *task*, which defines where and how AWS DataSync transfers your data.

A task includes a source location, destination location, and transfer options (such as bandwidth limits, scheduling, and more).

### Important

If you're planning to transfer data to or from an Amazon S3 location, review [how DataSync can affect your S3 request charges](#) and the [DataSync pricing page](#) before you begin.

## Request Syntax

```
{
  "CloudWatchLogGroupArn": "string",
  "DestinationLocationArn": "string",
  "Excludes": [
    {
      "FilterType": "string",
      "Value": "string"
    }
  ],
  "Includes": [
    {
      "FilterType": "string",
      "Value": "string"
    }
  ],
  "ManifestConfig": {
    "Action": "string",
    "Format": "string",
    "Source": {
      "S3": {
        "BucketAccessRoleArn": "string",
        "ManifestObjectPath": "string",
        "ManifestObjectVersionId": "string",
        "S3BucketArn": "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",
  "Status": "string"
},
"SourceLocationArn": "string",
"Tags": [
  {
    "Key": "string",
    "Value": "string"
  }
],
"TaskReportConfig": {
  "Destination": {
    "S3": {
      "BucketAccessRoleArn": "string",
      "S3BucketArn": "string",
      "Subdirectory": "string"
    }
  },
  "ObjectVersionIds": "string",
  "OutputType": "string",
  "Overrides": {
    "Deleted": {
      "ReportLevel": "string"
    },
    "Skipped": {

```

```

        "ReportLevel": "string"
    },
    "Transferred": {
        "ReportLevel": "string"
    },
    "Verified": {
        "ReportLevel": "string"
    }
},
"ReportLevel": "string"
}
}

```

## Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

### CloudWatchLogGroupArn

Specifies the Amazon Resource Name (ARN) of an Amazon CloudWatch log group for monitoring your task.

For more information, see [Monitoring DataSync with Amazon CloudWatch](#).

Type: String

Length Constraints: Maximum length of 562.

Pattern: `^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):logs:[a-z\-\0-9]+:[0-9]{12}:log-group:([^\:]*)(:\:)*?$`

Required: No

### DestinationLocationArn

Specifies the ARN of your transfer's destination location.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^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}$`



Required: Yes

### Excludes

Specifies exclude filters that define the files, objects, and folders in your source location that you don't want DataSync to transfer. For more information and examples, see [Specifying what DataSync transfers by using filters](#).

Type: Array of [FilterRule](#) objects

Array Members: Minimum number of 0 items. Maximum number of 1 item.

Required: No

### Includes

Specifies include filters define the files, objects, and folders in your source location that you want DataSync to transfer. For more information and examples, see [Specifying what DataSync transfers by using filters](#).

Type: Array of [FilterRule](#) objects

Array Members: Minimum number of 0 items. Maximum number of 1 item.

Required: No

### ManifestConfig

Configures a manifest, which is a list of files or objects that you want DataSync to transfer. For more information and configuration examples, see [Specifying what DataSync transfers by using a manifest](#).

When using this parameter, your caller identity (the role that you're using DataSync with) must have the `iam:PassRole` permission. The [AWSDataSyncFullAccess](#) policy includes this permission.

Type: [ManifestConfig](#) object

Required: No

### Name

Specifies the name of your task.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 256.

Pattern: `^[a-zA-Z0-9\s+=. _:@/-]+$`

Required: No

## Options

Specifies your task's settings, such as preserving file metadata, verifying data integrity, among other options.

Type: [Options](#) object

Required: No

## Schedule

Specifies a schedule for when you want your task to run. For more information, see [Scheduling your task](#).

Type: [TaskSchedule](#) object

Required: No

## SourceLocationArn

Specifies the ARN of your transfer's source location.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^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}$`

Required: Yes

## Tags

Specifies the tags that you want to apply to your task.

*Tags* are key-value pairs that help you manage, filter, and search for your DataSync resources.

Type: Array of [TagListEntry](#) objects

Array Members: Minimum number of 0 items. Maximum number of 50 items.

Required: No

### **TaskReportConfig**

Specifies how you want to configure a task report, which provides detailed information about your DataSync transfer. For more information, see [Monitoring your DataSync transfers with task reports](#).

When using this parameter, your caller identity (the role that you're using DataSync with) must have the `iam:PassRole` permission. The [AWSDataSyncFullAccess](#) policy includes this permission.

Type: [TaskReportConfig](#) object

Required: No

## **Response Syntax**

```
{
  "TaskArn": "string"
}
```

## **Response Elements**

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

### **TaskArn**

The Amazon Resource Name (ARN) of the task.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^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}$`

## **Errors**

For information about the errors that are common to all actions, see [Common Errors](#).

## InternalException

This exception is thrown when an error occurs in the AWS DataSync service.

HTTP Status Code: 500

## InvalidRequestException

This exception is thrown when the client submits a malformed request.

HTTP Status Code: 400

## Examples

### Sample Request

The following example creates a task.

```
{
  "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:datasync:us-east-2:111222333444:location/loc-07db7abfc326c50fb",
  "Name": "MyTask",
  "SourceLocationArn": "arn:aws:datasync:us-east-2:111222333444:location/loc-0f01451b140b2af49",
  "Tags": [{
    "Key": "Name",
    "Value": "Task-1"
  }]
}
```

## Sample Response

The following response includes the Amazon Resource Name (ARN) of the task.

```
{
  "TaskArn": "arn:aws:datasync:us-east-2:111222333444:task/task-08de6e6697796f026"
}
```

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

## DeleteAgent

Removes an AWS DataSync agent resource from your AWS account.

Keep in mind that this operation (which can't be undone) doesn't remove the agent's virtual machine (VM) or Amazon EC2 instance from your storage environment. For next steps, you can delete the VM or instance from your storage environment or reuse it to [activate a new agent](#).

### Request Syntax

```
{  
  "AgentArn": "string"  
}
```

### Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

#### [AgentArn](#)

The Amazon Resource Name (ARN) of the agent to delete. Use the `ListAgents` operation to return a list of agents for your account and AWS Region.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^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}$`

Required: Yes

### Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

### Errors

For information about the errors that are common to all actions, see [Common Errors](#).

## InternalException

This exception is thrown when an error occurs in the AWS DataSync service.

HTTP Status Code: 500

## InvalidRequestException

This exception is thrown when the client submits a malformed request.

HTTP Status Code: 400

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

## DeleteLocation

Deletes a transfer location resource from AWS DataSync.

### Request Syntax

```
{  
  "LocationArn": "string"  
}
```

### Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

#### LocationArn

The Amazon Resource Name (ARN) of the location to delete.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^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}$`

Required: Yes

### Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

### Errors

For information about the errors that are common to all actions, see [Common Errors](#).

#### InternalException

This exception is thrown when an error occurs in the AWS DataSync service.

HTTP Status Code: 500



## InvalidRequestException

This exception is thrown when the client submits a malformed request.

HTTP Status Code: 400

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

## DeleteTask

Deletes a transfer task resource from AWS DataSync.

### Request Syntax

```
{  
  "TaskArn": "string"  
}
```

### Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

#### TaskArn

Specifies the Amazon Resource Name (ARN) of the task that you want to delete.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^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}$`

Required: Yes

### Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

### Errors

For information about the errors that are common to all actions, see [Common Errors](#).

#### InternalException

This exception is thrown when an error occurs in the AWS DataSync service.

HTTP Status Code: 500

## InvalidRequestException

This exception is thrown when the client submits a malformed request.

HTTP Status Code: 400

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

## DescribeAgent

Returns information about an AWS DataSync agent, such as its name, service endpoint type, and status.

### Request Syntax

```
{  
  "AgentArn": "string"  
}
```

### Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

#### AgentArn

Specifies the Amazon Resource Name (ARN) of the DataSync agent that you want information about.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^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}$`

Required: Yes

### Response Syntax

```
{  
  "AgentArn": "string",  
  "CreationTime": number,  
  "EndpointType": "string",  
  "LastConnectionTime": number,  
  "Name": "string",  
  "Platform": {  
    "Version": "string"  
  }
```

```
  },
  "PrivateLinkConfig": {
    "PrivateLinkEndpoint": "string",
    "SecurityGroupArns": [ "string" ],
    "SubnetArns": [ "string" ],
    "VpcEndpointId": "string"
  },
  "Status": "string"
}
```

## Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

### AgentArn

The ARN of the agent.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^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

The time that the agent was [activated](#).

Type: Timestamp

### EndpointType

The type of [service endpoint](#) that your agent is connected to.

Type: String

Valid Values: PUBLIC | PRIVATE\_LINK | FIPS

### LastConnectionTime

The last time that the agent was communicating with the DataSync service.

Type: Timestamp

## Name

The name of the agent.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 256.

Pattern: `^[a-zA-Z0-9\s+=. _:@/- ]+$`

## Platform

The platform-related details about the agent, such as the version number.

Type: [Platform](#) object

## PrivateLinkConfig

The network configuration that the agent uses when connecting to a [VPC service endpoint](#).

Type: [PrivateLinkConfig](#) object

## Status

The status of the agent.

- If the status is ONLINE, the agent is configured properly and ready to use.
- If the status is OFFLINE, the agent has been out of contact with DataSync for five minutes or longer. This can happen for a few reasons. For more information, see [What do I do if my agent is offline?](#)

Type: String

Valid Values: ONLINE | OFFLINE

## **Errors**

For information about the errors that are common to all actions, see [Common Errors](#).

## **InternalException**

This exception is thrown when an error occurs in the AWS DataSync service.

HTTP Status Code: 500

## InvalidRequestException

This exception is thrown when the client submits a malformed request.

HTTP Status Code: 400

## Examples

### Sample Request

The following example returns information about an agent specified in a request.

```
{
  "AgentArn": "arn:aws:datsync:us-east-2:111122223333:agent/agent-1234567890abcdef0"
}
```

### Sample Response

The following example response describes an agent that uses a public service endpoint.

```
{
  "AgentArn": "arn:aws:datsync:us-east-2:111122223333:agent/
agent-1234567890abcdef0",
  "Name": "Data center migration agent",
  "Status": "ONLINE",
  "LastConnectionTime": "2022-10-17T17:21:35.540000+00:00",
  "CreationTime": "2022-10-05T20:52:29.499000+00:00",
  "EndpointType": "PUBLIC",
  "Platform": {
    "Version": "2"
  }
}
```

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)

- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)



## DescribeDiscoveryJob

Returns information about a DataSync discovery job.

### Request Syntax

```
{  
  "DiscoveryJobArn": "string"  
}
```

### Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

#### DiscoveryJobArn

Specifies the Amazon Resource Name (ARN) of the discovery job that you want information about.

Type: String

Length Constraints: Maximum length of 256.

Pattern: `^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}/job/discovery-job-[a-f0-9]{8}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{12}$`

Required: Yes

### Response Syntax

```
{  
  "CollectionDurationMinutes": number,  
  "DiscoveryJobArn": "string",  
  "JobEndTime": number,  
  "JobStartTime": number,  
  "Status": "string",  
}
```

```
"StorageSystemArn": "string"  
}
```

## Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

### CollectionDurationMinutes

The number of minutes that the discovery job runs.

Type: Integer

Valid Range: Minimum value of 60. Maximum value of 44640.

### DiscoveryJobArn

The ARN of the discovery job.

Type: String

Length Constraints: Maximum length of 256.

Pattern: `^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

The time when the discovery job ended.

Type: Timestamp

### JobStartTime

The time when the discovery job started.

Type: Timestamp

### Status

Indicates the status of a discovery job. For more information, see [Discovery job statuses](#).

Type: String

Valid Values: RUNNING | WARNING | TERMINATED | FAILED | STOPPED | COMPLETED  
| COMPLETED\_WITH\_ISSUES

### StorageSystemArn

The ARN of the on-premises storage system you're running the discovery job on.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^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}$`

## Errors

For information about the errors that are common to all actions, see [Common Errors](#).

### InternalException

This exception is thrown when an error occurs in the AWS DataSync service.

HTTP Status Code: 500

### InvalidRequestException

This exception is thrown when the client submits a malformed request.

HTTP Status Code: 400

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)

- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

## DescribeLocationAzureBlob

Provides details about how an AWS DataSync transfer location for Microsoft Azure Blob Storage is configured.

### Request Syntax

```
{
  "LocationArn": "string"
}
```

### Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

#### LocationArn

Specifies the Amazon Resource Name (ARN) of your Azure Blob Storage transfer location.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^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}$`

Required: Yes

### Response Syntax

```
{
  "AccessTier": "string",
  "AgentArns": [ "string" ],
  "AuthenticationType": "string",
  "BlobType": "string",
  "CreationTime": number,
  "LocationArn": "string",
  "LocationUri": "string"
}
```

## Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

### AccessTier

The access tier that you want your objects or files transferred into. This only applies when using the location as a transfer destination. For more information, see [Access tiers](#).

Type: String

Valid Values: HOT | COOL | ARCHIVE

### AgentArns

The ARNs of the DataSync agents that can connect with your Azure Blob Storage container.

Type: Array of strings

Array Members: Minimum number of 1 item. Maximum number of 4 items.

Length Constraints: Maximum length of 128.

Pattern: `^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

The authentication method DataSync uses to access your Azure Blob Storage. DataSync can access blob storage using a shared access signature (SAS).

Type: String

Valid Values: SAS

### BlobType

The type of blob that you want your objects or files to be when transferring them into Azure Blob Storage. Currently, DataSync only supports moving data into Azure Blob Storage as block blobs. For more information on blob types, see the [Azure Blob Storage documentation](#).

Type: String

Valid Values: BLOCK

### CreationTime

The time that your Azure Blob Storage transfer location was created.

Type: Timestamp

### LocationArn

The ARN of your Azure Blob Storage transfer location.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^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

The URL of the Azure Blob Storage container involved in your transfer.

Type: String

Length Constraints: Maximum length of 4360.

Pattern: `^(efs|nfs|s3|smb|hdfs|fsx[a-z0-9-]+)://[a-zA-Z0-9.:/\-]+$`

## Errors

For information about the errors that are common to all actions, see [Common Errors](#).

### InternalException

This exception is thrown when an error occurs in the AWS DataSync service.

HTTP Status Code: 500

### InvalidRequestException

This exception is thrown when the client submits a malformed request.

HTTP Status Code: 400

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)



## DescribeLocationEfs

Provides details about how an AWS DataSync transfer location for an Amazon EFS file system is configured.

### Request Syntax

```
{  
  "LocationArn": "string"  
}
```

### Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

#### LocationArn

The Amazon Resource Name (ARN) of the Amazon EFS file system location that you want information about.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^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}$`

Required: Yes

### Response Syntax

```
{  
  "AccessPointArn": "string",  
  "CreationTime": number,  
  "Ec2Config": {  
    "SecurityGroupArns": [ "string" ],  
    "SubnetArn": "string"  
  },  
  "FileSystemAccessRoleArn": "string",
```

```
"InTransitEncryption": "string",  
"LocationArn": "string",  
"LocationUri": "string"  
}
```

## Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

### AccessPointArn

The ARN of the access point that DataSync uses to access the Amazon EFS file system.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^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

The time that the location was created.

Type: Timestamp

### Ec2Config

The subnet and security groups that AWS DataSync uses to access your Amazon EFS file system.

Type: [Ec2Config](#) object

### FileSystemAccessRoleArn

The AWS Identity and Access Management (IAM) role that DataSync assumes when mounting the Amazon EFS file system.

Type: String

Length Constraints: Maximum length of 2048.

Pattern: `^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):iam::[0-9]{12}:role/.*$`

## InTransitEncryption

Describes whether DataSync uses Transport Layer Security (TLS) encryption when copying data to or from the Amazon EFS file system.

Type: String

Valid Values: NONE | TLS1\_2

## LocationArn

The ARN of the Amazon EFS file system location.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^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

The URL of the Amazon EFS file system location.

Type: String

Length Constraints: Maximum length of 4360.

Pattern: `^(efs|nfs|s3|smb|hdfs|fsx[a-z0-9-]+)://[a-zA-Z0-9.:\-]+$`

## **Errors**

For information about the errors that are common to all actions, see [Common Errors](#).

### **InternalException**

This exception is thrown when an error occurs in the AWS DataSync service.

HTTP Status Code: 500

### **InvalidRequestException**

This exception is thrown when the client submits a malformed request.

HTTP Status Code: 400

## Examples

### Sample Request

The following example shows how to get information about a specific Amazon EFS file system location.

```
{
  "LocationArn": "arn:aws:datasync:us-east-2:111222333444:location/
loc-12abcdef012345678"
}
```

### Sample Response

The following example returns location details about an Amazon EFS file system.

```
{
  "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/"
}
```

### Sample Response: Describing a location for a restricted Amazon EFS file system

The following example returns location details about an Amazon EFS file system with restricted access, including the `AccessPointArn`, `FileSystemAccessRoleArn`, and `InTransitEncryption` elements.

```
{
  "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"
}]
}
```

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

## DescribeLocationFsxLustre

Provides details about how an AWS DataSync transfer location for an Amazon FSx for Lustre file system is configured.

### Request Syntax

```
{  
  "LocationArn": "string"  
}
```

### Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

#### LocationArn

The Amazon Resource Name (ARN) of the FSx for Lustre location to describe.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^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}$`

Required: Yes

### Response Syntax

```
{  
  "CreationTime": number,  
  "LocationArn": "string",  
  "LocationUri": "string",  
  "SecurityGroupArns": [ "string" ]  
}
```

### Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

### CreationTime

The time that the FSx for Lustre location was created.

Type: Timestamp

### LocationArn

The Amazon Resource Name (ARN) of the FSx for Lustre location that was described.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^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

The URI of the FSx for Lustre location that was described.

Type: String

Length Constraints: Maximum length of 4360.

Pattern: `^(efs|nfs|s3|smb|hdfs|fsx[a-z0-9-]+)://[a-zA-Z0-9.:\-]+$`

### SecurityGroupArns

The Amazon Resource Names (ARNs) of the security groups that are configured for the FSx for Lustre file system.

Type: Array of strings

Array Members: Minimum number of 1 item. Maximum number of 5 items.

Length Constraints: Maximum length of 128.

Pattern: `^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]+$`

## Errors

For information about the errors that are common to all actions, see [Common Errors](#).

## InternalException

This exception is thrown when an error occurs in the AWS DataSync service.

HTTP Status Code: 500

## InvalidRequestException

This exception is thrown when the client submits a malformed request.

HTTP Status Code: 400

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)



## DescribeLocationFsxOntap

Provides details about how an AWS DataSync transfer location for an Amazon FSx for NetApp ONTAP file system is configured.

### Note

If your location uses SMB, the DescribeLocationFsxOntap operation doesn't actually return a Password.

## Request Syntax

```
{
  "LocationArn": "string"
}
```

## Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

### LocationArn

Specifies the Amazon Resource Name (ARN) of the FSx for ONTAP file system location that you want information about.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^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}$`

Required: Yes

## Response Syntax

```
{
  "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"
  }

```

## Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

### CreationTime

The time that the location was created.

Type: Timestamp

### FsxFilesystemArn

The ARN of the FSx for ONTAP file system.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):fsx:[a-z\-\0-9]*:[0-9]{12}:file-system/fs-.*$`

## LocationArn

The ARN of the FSx for ONTAP file system location.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^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

The uniform resource identifier (URI) of the FSx for ONTAP file system location.

Type: String

Length Constraints: Maximum length of 4360.

Pattern: `^(efs|nfs|s3|smb|hdfs|fsx[a-z0-9-]+)://[a-zA-Z0-9.:/\-\_]+$`

## Protocol

Specifies the data transfer protocol that AWS DataSync uses to access your Amazon FSx file system.

Type: [FsxProtocol](#) object

## SecurityGroupArns

The security groups that DataSync uses to access your FSx for ONTAP file system.

Type: Array of strings

Array Members: Minimum number of 1 item. Maximum number of 5 items.

Length Constraints: Maximum length of 128.

Pattern: `^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

The ARN of the storage virtual machine (SVM) on your FSx for ONTAP file system where you're copying data to or from.

Type: String

Length Constraints: Maximum length of 162.

Pattern: `^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,}$`

## Errors

For information about the errors that are common to all actions, see [Common Errors](#).

### InternalException

This exception is thrown when an error occurs in the AWS DataSync service.

HTTP Status Code: 500

### InvalidRequestException

This exception is thrown when the client submits a malformed request.

HTTP Status Code: 400

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

## DescribeLocationFsxOpenZfs

Provides details about how an AWS DataSync transfer location for an Amazon FSx for OpenZFS file system is configured.

### Note

Response elements related to SMB aren't supported with the DescribeLocationFsxOpenZfs operation.

## Request Syntax

```
{
  "LocationArn": "string"
}
```

## Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

### LocationArn

The Amazon Resource Name (ARN) of the FSx for OpenZFS location to describe.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^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}$`

Required: Yes

## Response Syntax

```
{
```

```

"CreationTime": number,
"LocationArn": "string",
"LocationUri": "string",
"Protocol": {
  "NFS": {
    "MountOptions": {
      "Version": "string"
    }
  },
  "SMB": {
    "Domain": "string",
    "MountOptions": {
      "Version": "string"
    },
    "Password": "string",
    "User": "string"
  }
},
"SecurityGroupArns": [ "string" ]
}

```

## Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

### CreationTime

The time that the FSx for OpenZFS location was created.

Type: Timestamp

### LocationArn

The ARN of the FSx for OpenZFS location that was described.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^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}$`

## LocationUri

The uniform resource identifier (URI) of the FSx for OpenZFS location that was described.

Example: `fsxz://us-west-2.fs-1234567890abcdef02/fsx/folderA/folder`

Type: String

Length Constraints: Maximum length of 4360.

Pattern: `^(efs|nfs|s3|smb|hdfs|fsx[a-z0-9-]+)://[a-zA-Z0-9.\:/\-\-]+$`

## Protocol

The type of protocol that AWS DataSync uses to access your file system.

Type: [FsxProtocol](#) object

## SecurityGroupArns

The ARNs of the security groups that are configured for the FSx for OpenZFS file system.

Type: Array of strings

Array Members: Minimum number of 1 item. Maximum number of 5 items.

Length Constraints: Maximum length of 128.

Pattern: `^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]+$`

## Errors

For information about the errors that are common to all actions, see [Common Errors](#).

### InternalException

This exception is thrown when an error occurs in the AWS DataSync service.

HTTP Status Code: 500

### InvalidRequestException

This exception is thrown when the client submits a malformed request.

HTTP Status Code: 400

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)



## DescribeLocationFsxWindows

Provides details about how an AWS DataSync transfer location for an Amazon FSx for Windows File Server file system is configured.

### Request Syntax

```
{  
  "LocationArn": "string"  
}
```

### Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

#### LocationArn

Specifies the Amazon Resource Name (ARN) of the FSx for Windows File Server location.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^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}$`

Required: Yes

### Response Syntax

```
{  
  "CreationTime": number,  
  "Domain": "string",  
  "LocationArn": "string",  
  "LocationUri": "string",  
  "SecurityGroupArns": [ "string" ],  
  "User": "string"  
}
```

## Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

### CreationTime

The time that the FSx for Windows File Server location was created.

Type: Timestamp

### Domain

The name of the Microsoft Active Directory domain that the FSx for Windows File Server file system belongs to.

Type: String

Length Constraints: Maximum length of 253.

Pattern: `^[A-Za-z0-9]( (\. | -+)? [A-Za-z0-9] ){0,252}$`

### LocationArn

The ARN of the FSx for Windows File Server location.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^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

The uniform resource identifier (URI) of the FSx for Windows File Server location.

Type: String

Length Constraints: Maximum length of 4360.

Pattern: `^(efs|nfs|s3|smb|hdfs|fsx[a-z0-9-]+)://[a-zA-Z0-9.:\-]+$`

## SecurityGroupArns

The ARNs of the security groups that are configured for the FSx for Windows File Server file system.

Type: Array of strings

Array Members: Minimum number of 1 item. Maximum number of 5 items.

Length Constraints: Maximum length of 128.

Pattern: `^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

The user with the permissions to mount and access the FSx for Windows File Server file system.

Type: String

Length Constraints: Maximum length of 104.

Pattern: `^[^\x5B\x5D\\\/:;|=, +*?]{1,104}$`

## Errors

For information about the errors that are common to all actions, see [Common Errors](#).

### InternalException

This exception is thrown when an error occurs in the AWS DataSync service.

HTTP Status Code: 500

### InvalidRequestException

This exception is thrown when the client submits a malformed request.

HTTP Status Code: 400

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)



```
{
  "Hostname": "string",
  "Port": number
},
"QopConfiguration": {
  "DataTransferProtection": "string",
  "RpcProtection": "string"
},
"ReplicationFactor": number,
"SimpleUser": "string"
}
```

## Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

### AgentArns

The ARNs of the DataSync agents that can connect with your HDFS cluster.

Type: Array of strings

Array Members: Minimum number of 1 item. Maximum number of 4 items.

Length Constraints: Maximum length of 128.

Pattern: `^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

The type of authentication used to determine the identity of the user.

Type: String

Valid Values: SIMPLE | KERBEROS

### BlockSize

The size of the data blocks to write into the HDFS cluster.

Type: Integer

Valid Range: Minimum value of 1048576. Maximum value of 1073741824.

### CreationTime

The time that the HDFS location was created.

Type: Timestamp

### KerberosPrincipal

The Kerberos principal with access to the files and folders on the HDFS cluster. This parameter is used if the `AuthenticationType` is defined as `KERBEROS`.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 256.

Pattern: `^\.+`\$

### KmsKeyProviderUri

The URI of the HDFS cluster's Key Management Server (KMS).

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `^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

The ARN of the HDFS location.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^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

The URI of the HDFS location.

Type: String

Length Constraints: Maximum length of 4360.

Pattern: `^(efs|nfs|s3|smb|hdfs|fsx[a-z0-9-]+)://[a-zA-Z0-9.\:/\-]+$`

### NameNodes

The NameNode that manages the HDFS namespace.

Type: Array of [HdfsNameNode](#) objects

Array Members: Minimum number of 1 item.

### QopConfiguration

The Quality of Protection (QOP) configuration, which specifies the Remote Procedure Call (RPC) and data transfer protection settings configured on the HDFS cluster.

Type: [QopConfiguration](#) object

### ReplicationFactor

The number of DataNodes to replicate the data to when writing to the HDFS cluster.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 512.

### SimpleUser

The user name to identify the client on the host operating system. This parameter is used if the `AuthenticationType` is defined as SIMPLE.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 256.

Pattern: `^[_ .A-Za-z0-9][-_ .A-Za-z0-9]*$`

## Errors

For information about the errors that are common to all actions, see [Common Errors](#).

### InternalException

This exception is thrown when an error occurs in the AWS DataSync service.



HTTP Status Code: 500

## **InvalidRequestException**

This exception is thrown when the client submits a malformed request.

HTTP Status Code: 400

## **See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

## DescribeLocationNfs

Provides details about how an AWS DataSync transfer location for a Network File System (NFS) file server is configured.

### Request Syntax

```
{  
  "LocationArn": "string"  
}
```

### Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

#### LocationArn

Specifies the Amazon Resource Name (ARN) of the NFS location that you want information about.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^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}$`

Required: Yes

### Response Syntax

```
{  
  "CreationTime": number,  
  "LocationArn": "string",  
  "LocationUri": "string",  
  "MountOptions": {  
    "Version": "string"  
  },  
  "OnPremConfig": {
```

```
    "AgentArns": [ "string" ]  
  }  
}
```

## Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

### CreationTime

The time when the NFS location was created.

Type: Timestamp

### LocationArn

The ARN of the NFS location.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^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}$`

### LocationUri

The URI of the NFS location.

Type: String

Length Constraints: Maximum length of 4360.

Pattern: `^(efs|nfs|s3|smb|hdfs|fsx[a-z0-9-]+)://[a-zA-Z0-9.\:/\-]+$`

### MountOptions

The mount options that DataSync uses to mount your NFS file server.

Type: [NfsMountOptions](#) object

### OnPremConfig

The AWS DataSync agents that are connecting to a Network File System (NFS) location.

Type: [OnPremConfig](#) object

## Errors

For information about the errors that are common to all actions, see [Common Errors](#).

### InternalException

This exception is thrown when an error occurs in the AWS DataSync service.

HTTP Status Code: 500

### InvalidRequestException

This exception is thrown when the client submits a malformed request.

HTTP Status Code: 400

## Examples

### Example

The following example returns information about the NFS location specified in the sample request.

### Sample Request

```
{
  "LocationArn": "arn:aws:datasync:us-east-2:111222333444:location/
loc-07db7abfc326c50aa"
}
```

### Example

This example illustrates one usage of DescribeLocationNfs.

### Sample Response

```
{
  "CreationTime": 1532660733.39,
  "LocationArn": "arn:aws:datasync:us-east-2:111222333444:location/
loc-07db7abfc326c50aa",
  "LocationUri": "hostname.amazon.com",
}
```

```
"OnPremConfig": {  
  "AgentArns": [ "arn:aws:datsync:us-east-2:111222333444:agent/  
agent-0b0addbeef44b3nfs" ]  
}
```

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

## DescribeLocationObjectStorage

Provides details about how an AWS DataSync transfer location for an object storage system is configured.

### Request Syntax

```
{  
  "LocationArn": "string"  
}
```

### Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

#### LocationArn

Specifies the Amazon Resource Name (ARN) of the object storage system location.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^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}$`

Required: Yes

### Response Syntax

```
{  
  "AccessKey": "string",  
  "AgentArns": [ "string" ],  
  "CreationTime": number,  
  "LocationArn": "string",  
  "LocationUri": "string",  
  "ServerCertificate": blob,  
  "ServerPort": number,
```

```
"ServerProtocol": "string"  
}
```

## Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

### AccessKey

The access key (for example, a user name) required to authenticate with the object storage system.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 200.

Pattern: ^.\*\$

### AgentArns

The ARNs of the DataSync agents that can connect with your object storage system.

Type: Array of strings

Array Members: Minimum number of 1 item. Maximum number of 4 items.

Length Constraints: Maximum length of 128.

Pattern: ^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}\$

### CreationTime

The time that the location was created.

Type: Timestamp

### LocationArn

The ARN of the object storage system location.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^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

The URI of the object storage system location.

Type: String

Length Constraints: Maximum length of 4360.

Pattern: `^(efs|nfs|s3|smb|hdfs|fsx[a-z0-9-]+)://[a-zA-Z0-9.:/\\-]+$`

### ServerCertificate

The certificate chain for DataSync to authenticate with your object storage system if the system uses a private or self-signed certificate authority (CA).

Type: Base64-encoded binary data object

Length Constraints: Maximum length of 32768.

### ServerPort

The port that your object storage server accepts inbound network traffic on (for example, port 443).

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 65536.

### ServerProtocol

The protocol that your object storage system uses to communicate.

Type: String

Valid Values: HTTPS | HTTP

## Errors

For information about the errors that are common to all actions, see [Common Errors](#).



## InternalException

This exception is thrown when an error occurs in the AWS DataSync service.

HTTP Status Code: 500

## InvalidRequestException

This exception is thrown when the client submits a malformed request.

HTTP Status Code: 400

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

## DescribeLocationS3

Provides details about how an AWS DataSync transfer location for an S3 bucket is configured.

### Request Syntax

```
{
  "LocationArn": "string"
}
```

### Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

#### LocationArn

Specifies the Amazon Resource Name (ARN) of the Amazon S3 location.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^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}$`

Required: Yes

### Response Syntax

```
{
  "AgentArns": [ "string" ],
  "CreationTime": number,
  "LocationArn": "string",
  "LocationUri": "string",
  "S3Config": {
    "BucketAccessRoleArn": "string"
  },
  "S3StorageClass": "string"
}
```

```
}
```

## Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

### AgentArns

The ARNs of the DataSync agents deployed on your Outpost when using working with Amazon S3 on Outposts.

For more information, see [Deploy your DataSync agent on AWS Outposts](#).

Type: Array of strings

Array Members: Minimum number of 1 item. Maximum number of 4 items.

Length Constraints: Maximum length of 128.

Pattern: `^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

The time that the Amazon S3 location was created.

Type: Timestamp

### LocationArn

The ARN of the Amazon S3 location.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^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

The URL of the Amazon S3 location that was described.

Type: String

Length Constraints: Maximum length of 4360.

Pattern: `^(efs|nfs|s3|smb|hdfs|fsx[a-z0-9-]+)://[a-zA-Z0-9.:/\-]+$`

### S3Config

Specifies the Amazon Resource Name (ARN) of the AWS Identity and Access Management (IAM) role that DataSync uses to access your S3 bucket.

For more information, see [Accessing S3 buckets](#).

Type: [S3Config](#) object

### S3StorageClass

When Amazon S3 is a destination location, this is the storage class that you chose for your objects.

Some storage classes have behaviors that can affect your Amazon S3 storage costs. For more information, see [Storage class considerations with Amazon S3 transfers](#).

Type: String

Valid Values: STANDARD | STANDARD\_IA | ONEZONE\_IA | INTELLIGENT\_TIERING | GLACIER | DEEP\_ARCHIVE | OUTPOSTS | GLACIER\_INSTANT\_RETRIEVAL

## Errors

For information about the errors that are common to all actions, see [Common Errors](#).

### InternalException

This exception is thrown when an error occurs in the AWS DataSync service.

HTTP Status Code: 500

### InvalidRequestException

This exception is thrown when the client submits a malformed request.

HTTP Status Code: 400

## Examples

### Example

The following example returns information about the Amazon S3 location specified in the sample request.

### Sample Request

```
{
  "LocationArn": "arn:aws:datasync:us-east-2:111222333444:location/loc-07db7abfc326c50s3"
}
```

### Example

This example illustrates one usage of DescribeLocationS3.

### Sample Response

```
{
  "CreationTime": 1532660733.39,
  "LocationArn": "arn:aws:datasync:us-east-2:111222333444:location/loc-07db7abfc326c50s3",
  "LocationUri": "MyBucket.",
  "S3Config": {
    "BucketAccessRoleArn": "arn:aws:iam::111222333444:role/MyBucketAccessRole",
  }
  "S3StorageClass": "STANDARD"
}
```

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)

- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

## DescribeLocationSmb

Provides details about how an AWS DataSync transfer location for a Server Message Block (SMB) file server is configured.

### Request Syntax

```
{  
  "LocationArn": "string"  
}
```

### Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

#### LocationArn

Specifies the Amazon Resource Name (ARN) of the SMB location that you want information about.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^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}$`

Required: Yes

### Response Syntax

```
{  
  "AgentArns": [ "string" ],  
  "CreationTime": number,  
  "Domain": "string",  
  "LocationArn": "string",  
  "LocationUri": "string",  
  "MountOptions": {  
    "Version": "string"  
  }
```

```
  },  
  "User": "string"  
}
```

## Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

### AgentArns

The ARNs of the DataSync agents that can connect with your SMB file server.

Type: Array of strings

Array Members: Minimum number of 1 item. Maximum number of 4 items.

Length Constraints: Maximum length of 128.

Pattern: `^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

The time that the SMB location was created.

Type: Timestamp

### Domain

The name of the Microsoft Active Directory domain that the SMB file server belongs to.

Type: String

Length Constraints: Maximum length of 253.

Pattern: `^[A-Za-z0-9](\\.|-+)?[A-Za-z0-9]{0,252}$`

### LocationArn

The ARN of the SMB location.

Type: String



Length Constraints: Maximum length of 128.

Pattern: `^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

The URI of the SMB location.

Type: String

Length Constraints: Maximum length of 4360.

Pattern: `^(efs|nfs|s3|smb|hdfs|fsx[a-z0-9-]+)://[a-zA-Z0-9.:/\\-]+$`

### MountOptions

The protocol that DataSync use to access your SMB file.

Type: [SmbMountOptions](#) object

### User

The user that can mount and access the files, folders, and file metadata in your SMB file server.

Type: String

Length Constraints: Maximum length of 104.

Pattern: `^[^\\x5B\\x5D\\|\\/;|=, +*?]{1,104}$`

## Errors

For information about the errors that are common to all actions, see [Common Errors](#).

### **InternalException**

This exception is thrown when an error occurs in the AWS DataSync service.

HTTP Status Code: 500

### **InvalidRequestException**

This exception is thrown when the client submits a malformed request.

HTTP Status Code: 400

## Examples

### Example

This example illustrates one usage of DescribeLocationSmb.

### Sample Request

```
{
  "arn:aws:datsync:us-east-1:111222333444:location/loc-0f01451b140b2af49"
}
```

### Example

This example illustrates one usage of DescribeLocationSmb.

### Sample Response

```
{
  "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"
}
```

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)

- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

## DescribeStorageSystem

Returns information about an on-premises storage system that you're using with DataSync Discovery.

### Request Syntax

```
{  
  "StorageSystemArn": "string"  
}
```

### Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

#### StorageSystemArn

Specifies the Amazon Resource Name (ARN) of an on-premises storage system that you're using with DataSync Discovery.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^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}$`

Required: Yes

### Response Syntax

```
{  
  "AgentArns": [ "string" ],  
  "CloudWatchLogGroupArn": "string",  
  "ConnectivityStatus": "string",  
  "CreationTime": number,  
  "ErrorMessage": "string",  
  "Name": "string",  
  "SecretsManagerArn": "string",
```

```
"ServerConfiguration": {
  "ServerHostname": "string",
  "ServerPort": number
},
"StorageSystemArn": "string",
"SystemType": "string"
}
```

## Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

### AgentArns

The ARN of the DataSync agent that connects to and reads from your on-premises storage system.

Type: Array of strings

Array Members: Fixed number of 1 item.

Length Constraints: Maximum length of 128.

Pattern: `^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

The ARN of the Amazon CloudWatch log group that's used to monitor and log discovery job events.

Type: String

Length Constraints: Maximum length of 562.

Pattern: `^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):logs:[a-z\\-0-9]+:[0-9]{12}:log-group:([^\:]*)(:\:)?$`

### ConnectivityStatus

Indicates whether your DataSync agent can connect to your on-premises storage system.

Type: String

Valid Values: PASS | FAIL | UNKNOWN

### CreationTime

The time when you added the on-premises storage system to DataSync Discovery.

Type: Timestamp

### ErrorMessage

Describes the connectivity error that the DataSync agent is encountering with your on-premises storage system.

Type: String

Length Constraints: Maximum length of 128.

Pattern: .\*

### Name

The name that you gave your on-premises storage system when adding it to DataSync Discovery.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 256.

Pattern: ^[\p{L}\p{M}\p{N}\s+=.\_:@\|-]+

### SecretsManagerArn

The ARN of the secret that stores your on-premises storage system's credentials. DataSync Discovery stores these credentials in [AWS Secrets Manager](#).

Type: String

Length Constraints: Maximum length of 2048.

Pattern: ^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):secretsmanager:[a-z\-\0-9]+:[0-9]{12}:secret:.\*

### ServerConfiguration

The server name and network port required to connect with your on-premises storage system's management interface.

Type: [DiscoveryServerConfiguration](#) object

### [StorageSystemArn](#)

The ARN of the on-premises storage system that the discovery job looked at.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^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}$`

### [SystemType](#)

The type of on-premises storage system.

#### Note

DataSync Discovery currently only supports NetApp Fabric-Attached Storage (FAS) and All Flash FAS (AFF) systems running ONTAP 9.7 or later.

Type: String

Valid Values: NetAppONTAP

## Errors

For information about the errors that are common to all actions, see [Common Errors](#).

### InternalException

This exception is thrown when an error occurs in the AWS DataSync service.

HTTP Status Code: 500

### InvalidRequestException

This exception is thrown when the client submits a malformed request.

HTTP Status Code: 400

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)



## DescribeStorageSystemResourceMetrics

Returns information, including performance data and capacity usage, which DataSync Discovery collects about a specific resource in your-premises storage system.

### Request Syntax

```
{
  "DiscoveryJobArn": "string",
  "EndTime": number,
  "MaxResults": number,
  "NextToken": "string",
  "ResourceId": "string",
  "ResourceType": "string",
  "StartTime": number
}
```

### Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

#### [DiscoveryJobArn](#)

Specifies the Amazon Resource Name (ARN) of the discovery job that collects information about your on-premises storage system.

Type: String

Length Constraints: Maximum length of 256.

Pattern: `^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}/job/discovery-job-[a-f0-9]{8}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{12}$`

Required: Yes

#### [EndTime](#)

Specifies a time within the total duration that the discovery job ran. To see information gathered during a certain time frame, use this parameter with `StartTime`.

Type: Timestamp

Required: No

### MaxResults

Specifies how many results that you want in the response.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 100.

Required: No

### NextToken

Specifies an opaque string that indicates the position to begin the next list of results in the response.

Type: String

Length Constraints: Maximum length of 65535.

Pattern: [a-zA-Z0-9=\_-]+

Required: No

### ResourceId

Specifies the universally unique identifier (UUID) of the storage system resource that you want information about.

Type: String

Pattern: [a-f0-9]{8}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{12}

Required: Yes

### ResourceType

Specifies the kind of storage system resource that you want information about.

Type: String

Valid Values: SVM | VOLUME | CLUSTER

Required: Yes

## StartTime

Specifies a time within the total duration that the discovery job ran. To see information gathered during a certain time frame, use this parameter with EndTime.

Type: Timestamp

Required: No

## Response Syntax

```
{
  "Metrics": [
    {
      "Capacity": {
        "ClusterCloudStorageUsed": number,
        "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"
}
```

## Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

### Metrics

The details that your discovery job collected about your storage system resource.

Type: Array of [ResourceMetrics](#) objects

### NextToken

The opaque string that indicates the position to begin the next list of results in the response.

Type: String

Length Constraints: Maximum length of 65535.

Pattern: [a-zA-Z0-9=\_- ]+

## Errors

For information about the errors that are common to all actions, see [Common Errors](#).

### **InternalException**

This exception is thrown when an error occurs in the AWS DataSync service.

HTTP Status Code: 500

### **InvalidRequestException**

This exception is thrown when the client submits a malformed request.

HTTP Status Code: 400

## Examples

### Sample Request

The following example requests information about a volume in an on-premises storage system.

```
{
  "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"
}
```

### Sample Response

The following example response includes performance and capacity information about the volume.

```
{
  "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"
  }
]
```

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

## DescribeStorageSystemResources

Returns information that DataSync Discovery collects about resources in your on-premises storage system.

### Request Syntax

```
{
  "DiscoveryJobArn": "string",
  "Filter": {
    "string" : [ "string" ]
  },
  "MaxResults": number,
  "NextToken": "string",
  "ResourceIds": [ "string" ],
  "ResourceType": "string"
}
```

### Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

#### [DiscoveryJobArn](#)

Specifies the Amazon Resource Name (ARN) of the discovery job that's collecting data from your on-premises storage system.

Type: String

Length Constraints: Maximum length of 256.

Pattern: `^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}/job/discovery-job-[a-f0-9]{8}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{12}$`

Required: Yes

## Filter

Filters the storage system resources that you want returned. For example, this might be volumes associated with a specific storage virtual machine (SVM).

Type: String to array of strings map

Valid Keys: SVM

Length Constraints: Maximum length of 1024.

Pattern: `^\.{0,1024}$`

Required: No

## MaxResults

Specifies the maximum number of storage system resources that you want to list in a response.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 100.

Required: No

## NextToken

Specifies an opaque string that indicates the position to begin the next list of results in the response.

Type: String

Length Constraints: Maximum length of 65535.

Pattern: `[a-zA-Z0-9=_- ]+`

Required: No

## ResourceIds

Specifies the universally unique identifiers (UUIDs) of the storage system resources that you want information about. You can't use this parameter in combination with the `Filter` parameter.

Type: Array of strings

Array Members: Minimum number of 1 item. Maximum number of 100 items.



Pattern: `[a-f0-9]{8}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{12}`

Required: No

### ResourceType

Specifies what kind of storage system resources that you want information about.

Type: String

Valid Values: SVM | VOLUME | CLUSTER

Required: Yes

## Response Syntax

```
{
  "NextToken": "string",
  "ResourceDetails": {
    "NetAppONTAPClusters": [
      {
        "CifsShareCount": number,
        "ClusterBlockStorageLogicalUsed": number,
        "ClusterBlockStorageSize": number,
        "ClusterBlockStorageUsed": number,
        "ClusterCloudStorageUsed": number,
        "ClusterName": "string",
        "LunCount": number,
        "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" ],
        "LunCount": number,
        "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,
        "LunCount": 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"
    }
]

```

```
}  
}
```

## Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

### NextToken

The opaque string that indicates the position to begin the next list of results in the response.

Type: String

Length Constraints: Maximum length of 65535.

Pattern: [a-zA-Z0-9=\_- ]+

### ResourceDetails

The information collected about your storage system's resources. A response can also include AWS storage service recommendations.

For more information, see [storage resource information](#) collected by and [recommendations](#) provided by DataSync Discovery.

Type: [ResourceDetails](#) object

## Errors

For information about the errors that are common to all actions, see [Common Errors](#).

### **InternalException**

This exception is thrown when an error occurs in the AWS DataSync service.

HTTP Status Code: 500

### **InvalidRequestException**

This exception is thrown when the client submits a malformed request.

HTTP Status Code: 400

## Examples

### Sample Request

The following example requests information about volumes in an on-premises storage system.

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

### Sample Response

The following example response includes information about volumes in the on-premises storage system, including recommendations about where to move the volume data to in AWS (such as Amazon FSx for NetApp ONTAP or 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-EXAMPLEaaaaa",
    "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"
}
]
```

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)



## DescribeTask

Provides information about a *task*, which defines where and how AWS DataSync transfers your data.

### Request Syntax

```
{  
  "TaskArn": "string"  
}
```

### Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

#### TaskArn

Specifies the Amazon Resource Name (ARN) of the transfer task that you want information about.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^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}$`

Required: Yes

### Response Syntax

```
{  
  "CloudWatchLogGroupArn": "string",  
  "CreationTime": number,  
  "CurrentTaskExecutionArn": "string",  
  "DestinationLocationArn": "string",  
  "DestinationNetworkInterfaceArns": [ "string" ],  
  "ErrorCode": "string",  
  "ErrorDetail": "string",  
  "Excludes": [  
    "string",  
    "string",  
    "string"  
  ]  
}
```

```

    {
      "FilterType": "string",
      "Value": "string"
    }
  ],
  "Includes": [
    {
      "FilterType": "string",
      "Value": "string"
    }
  ],
  "ManifestConfig": {
    "Action": "string",
    "Format": "string",
    "Source": {
      "S3": {
        "BucketAccessRoleArn": "string",
        "ManifestObjectPath": "string",
        "ManifestObjectVersionId": "string",
        "S3BucketArn": "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",
    "Status": "string"
  }
}

```

```

},
"ScheduleDetails": {
  "DisabledBy": "string",
  "DisabledReason": "string",
  "StatusUpdateTime": number
},
"SourceLocationArn": "string",
"SourceNetworkInterfaceArns": [ "string" ],
"Status": "string",
"TaskArn": "string",
"TaskReportConfig": {
  "Destination": {
    "S3": {
      "BucketAccessRoleArn": "string",
      "S3BucketArn": "string",
      "Subdirectory": "string"
    }
  },
  "ObjectVersionIds": "string",
  "OutputType": "string",
  "Overrides": {
    "Deleted": {
      "ReportLevel": "string"
    },
    "Skipped": {
      "ReportLevel": "string"
    },
    "Transferred": {
      "ReportLevel": "string"
    },
    "Verified": {
      "ReportLevel": "string"
    }
  },
  "ReportLevel": "string"
}
}

```

## Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

## CloudWatchLogGroupArn

The Amazon Resource Name (ARN) of an Amazon CloudWatch log group for monitoring your task.

For more information, see [Monitoring DataSync with Amazon CloudWatch](#).

Type: String

Length Constraints: Maximum length of 562.

Pattern: `^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):logs:[a-z\-\0-9]+:[0-9]{12}:log-group:([\^:\*]*)(:\^*)?&`

## CreationTime

The time that the task was created.

Type: Timestamp

## CurrentTaskExecutionArn

The ARN of the most recent task execution.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^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

The ARN of your transfer's destination location.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^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

The ARNs of the [network interfaces](#) that DataSync created for your destination location.

Type: Array of strings

Length Constraints: Maximum length of 128.

Pattern: `^arn:aws[\-a-z]{0,}:ec2:[a-z\-\0-9]*:[0-9]{12}:network-interface/eni-[0-9a-f]+$`

### ErrorCode

If there's an issue with your task, you can use the error code to help you troubleshoot the problem. For more information, see [Troubleshooting issues with DataSync transfers](#).

Type: String

### ErrorDetail

If there's an issue with your task, you can use the error details to help you troubleshoot the problem. For more information, see [Troubleshooting issues with DataSync transfers](#).

Type: String

### Excludes

The exclude filters that define the files, objects, and folders in your source location that you don't want DataSync to transfer. For more information and examples, see [Specifying what DataSync transfers by using filters](#).

Type: Array of [FilterRule](#) objects

Array Members: Minimum number of 0 items. Maximum number of 1 item.

### Includes

The include filters that define the files, objects, and folders in your source location that you want DataSync to transfer. For more information and examples, see [Specifying what DataSync transfers by using filters](#).

Type: Array of [FilterRule](#) objects

Array Members: Minimum number of 0 items. Maximum number of 1 item.

### ManifestConfig

The configuration of the manifest that lists the files or objects that you want DataSync to transfer. For more information, see [Specifying what DataSync transfers by using a manifest](#).

Type: [ManifestConfig](#) object

## Name

The name of your task.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 256.

Pattern: `^[a-zA-Z0-9\s+=. _:@/-]+$`

## Options

The task's settings. For example, what file metadata gets preserved, how data integrity gets verified at the end of your transfer, bandwidth limits, among other options.

Type: [Options](#) object

## Schedule

The schedule for when you want your task to run. For more information, see [Scheduling your task](#).

Type: [TaskSchedule](#) object

## ScheduleDetails

The details about your [task schedule](#).

Type: [TaskScheduleDetails](#) object

## SourceLocationArn

The ARN of your transfer's source location.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^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

The ARNs of the [network interfaces](#) that DataSync created for your source location.

Type: Array of strings

Length Constraints: Maximum length of 128.

Pattern: `^arn:aws[\-a-z]{0,}:ec2:[a-z\-\0-9]*:[0-9]{12}:network-interface/eni-[0-9a-f]+$`

## Status

The status of your task. For information about what each status means, see [Task statuses](#).

Type: String

Valid Values: AVAILABLE | CREATING | QUEUED | RUNNING | UNAVAILABLE

## TaskArn

The ARN of your task.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^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}$`

## TaskReportConfig

The configuration of your task report, which provides detailed information about your DataSync transfer. For more information, see [Monitoring your DataSync transfers with task reports](#).

Type: [TaskReportConfig](#) object

## Errors

For information about the errors that are common to all actions, see [Common Errors](#).

### InternalException

This exception is thrown when an error occurs in the AWS DataSync service.

HTTP Status Code: 500

### InvalidRequestException

This exception is thrown when the client submits a malformed request.

HTTP Status Code: 400

## Examples

### Example

The following example returns information about the task specified in the sample request.

### Sample Request

```
{
  "TaskArn": "arn:aws:datsync:us-east-2:111222333444:task/task-08de6e6697796f026"
}
```

### Example

This example illustrates one usage of DescribeTask.

### Sample Response

```
{
  "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"
}
```



```
}
```

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

## DescribeTaskExecution

Provides information about an execution of your AWS DataSync task. You can use this operation to help monitor the progress of an ongoing transfer or check the results of the transfer.

### Request Syntax

```
{
  "TaskExecutionArn": "string"
}
```

### Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

#### TaskExecutionArn

Specifies the Amazon Resource Name (ARN) of the task execution that you want information about.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^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}$`

Required: Yes

### Response Syntax

```
{
  "BytesCompressed": number,
  "BytesTransferred": number,
  "BytesWritten": number,
  "EstimatedBytesToTransfer": number,
  "EstimatedFilesToDelete": number,
  "EstimatedFilesToTransfer": number,
  "Excludes": [
    {
```

```

        "FilterType": "string",
        "Value": "string"
    }
],
"FilesDeleted": number,
"FilesSkipped": number,
"FilesTransferred": number,
"FilesVerified": number,
"Includes": [
    {
        "FilterType": "string",
        "Value": "string"
    }
],
"ManifestConfig": {
    "Action": "string",
    "Format": "string",
    "Source": {
        "S3": {
            "BucketAccessRoleArn": "string",
            "ManifestObjectPath": "string",
            "ManifestObjectVersionId": "string",
            "S3BucketArn": "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"
},
"ReportResult": {

```

```

    "ErrorCode": "string",
    "ErrorDetail": "string",
    "Status": "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",
  "TaskReportConfig": {
    "Destination": {
      "S3": {
        "BucketAccessRoleArn": "string",
        "S3BucketArn": "string",
        "Subdirectory": "string"
      }
    },
    "ObjectVersionIds": "string",
    "OutputType": "string",
    "Overrides": {
      "Deleted": {
        "ReportLevel": "string"
      },
      "Skipped": {
        "ReportLevel": "string"
      },
      "Transferred": {
        "ReportLevel": "string"
      },
      "Verified": {
        "ReportLevel": "string"
      }
    }
  },
  "ReportLevel": "string"
}

```

```
}
```

## Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

### BytesCompressed

The physical number of bytes transferred over the network after compression was applied. In most cases, this number is less than `BytesTransferred` unless the data isn't compressible.

Type: Long

### BytesTransferred

The total number of bytes that are involved in the transfer. For the number of bytes sent over the network, see `BytesCompressed`.

Type: Long

### BytesWritten

The number of logical bytes written to the destination location.

Type: Long

### EstimatedBytesToTransfer

The estimated physical number of bytes that will transfer over the network.

Type: Long

### EstimatedFilesToDelete

The expected number of files, objects, and directories that DataSync will delete in your destination location. If you don't [configure your task](#) to delete data in the destination that isn't in the source, the value is always 0.

Type: Long

### EstimatedFilesToTransfer

The expected number of files, objects, and directories that DataSync will transfer over the network. This value is calculated during the task execution's PREPARING phase before the

TRANSFERRING phase. The calculation is based on comparing the content of the source and destination locations and finding the difference that needs to be transferred.

Type: Long

### Excludes

A list of filter rules that exclude specific data during your transfer. For more information and examples, see [Filtering data transferred by DataSync](#).

Type: Array of [FilterRule](#) objects

Array Members: Minimum number of 0 items. Maximum number of 1 item.

### FilesDeleted

The number of files, objects, and directories that DataSync deleted in your destination location. If you don't [configure your task](#) to delete data in the destination that isn't in the source, the value is always 0.

Type: Long

### FilesSkipped

The number of files, objects, and directories that DataSync skipped during your transfer.

Type: Long

### FilesTransferred

The actual number of files, objects, and directories that DataSync transferred over the network. This value is updated periodically during the task execution's TRANSFERRING phase when something is read from the source and sent over the network.

If DataSync fails to transfer something, this value can be less than `EstimatedFilesToTransfer`. In some cases, this value can also be greater than `EstimatedFilesToTransfer`. This element is implementation-specific for some location types, so don't use it as an exact indication of what transferred or to monitor your task execution.

Type: Long

### FilesVerified

The number of files, objects, and directories that DataSync verified during your transfer.

**Note**

When you configure your task to [verify only the data that's transferred](#), DataSync doesn't verify directories in some situations or files that fail to transfer.

Type: Long

**Includes**

A list of filter rules that include specific data during your transfer. For more information and examples, see [Filtering data transferred by DataSync](#).

Type: Array of [FilterRule](#) objects

Array Members: Minimum number of 0 items. Maximum number of 1 item.

**ManifestConfig**

The configuration of the manifest that lists the files or objects to transfer. For more information, see [Specifying what DataSync transfers by using a manifest](#).

Type: [ManifestConfig](#) object

**Options**

Indicates how your transfer task is configured. These options include how DataSync handles files, objects, and their associated metadata during your transfer. You also can specify how to verify data integrity, set bandwidth limits for your task, among other options.

Each option has a default value. Unless you need to, you don't have to configure any option before calling [StartTaskExecution](#).

You also can override your task options for each task execution. For example, you might want to adjust the `LogLevel` for an individual execution.

Type: [Options](#) object

**ReportResult**

Indicates whether DataSync generated a complete [task report](#) for your transfer.

Type: [ReportResult](#) object

## Result

The result of the task execution.

Type: [TaskExecutionResultDetail](#) object

## StartTime

The time when the task execution started.

Type: Timestamp

## Status

The status of the task execution.

For detailed information about task execution statuses, see [Task execution statuses](#).

Type: String

Valid Values: QUEUED | CANCELLING | LAUNCHING | PREPARING | TRANSFERRING | VERIFYING | SUCCESS | ERROR

## TaskExecutionArn

The ARN of the task execution that you wanted information about. TaskExecutionArn is hierarchical and includes TaskArn for the task that was executed.

For example, a TaskExecution value with the ARN `arn:aws:datsync:us-east-1:111222333444:task/task-0208075f79cedf4a2/execution/exec-08ef1e88ec491019b` executed the task with the ARN `arn:aws:datsync:us-east-1:111222333444:task/task-0208075f79cedf4a2`.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^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}$`

## TaskReportConfig

The configuration of your task report, which provides detailed information about for your DataSync transfer. For more information, see [Creating a task report](#).

Type: [TaskReportConfig](#) object



## Errors

For information about the errors that are common to all actions, see [Common Errors](#).

### InternalException

This exception is thrown when an error occurs in the AWS DataSync service.

HTTP Status Code: 500

### InvalidRequestException

This exception is thrown when the client submits a malformed request.

HTTP Status Code: 400

## Examples

### Sample Request

This example illustrates a DescribeTaskExecution request.

```
{
  "TaskExecutionArn": "arn:aws:datsync:us-east-1:111222333444:task/task-
aaaabbbbccccdddf/execution/exec-1234abcd1234abcd1"
}
```

### Sample Response

This example illustrates a DescribeTaskExecution response.

```
{
  "BytesCompressed": 3500,
  "BytesTransferred": 5000,
  "BytesWritten": 5000,
  "EstimatedBytesToTransfer": 5000,
  "EstimatedFilesToDelete": 10,
  "EstimatedFilesToTransfer": 100,
  "FilesDeleted": 10,
  "FilesSkipped": 0,
  "FilesTransferred": 100,
  "FilesVerified": 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-1:111222333444:task/task-
aaaabbbbccccdddf/execution/exec-1234abcd1234abcd1",
  "TaskReportConfig": {
    "Destination": {
      "S3": {
        "BucketAccessRoleArn": "arn:aws:iam::111222333444:role/my-datsync-
role",
        "S3BucketArn": "arn:aws:s3:::my-task-reports-bucket/*",
        "Subdirectory": "reports"
      }
    }
  },
  "ObjectVersionIds": "INCLUDE",
  "OutputType": "STANDARD",
  "Overrides": {
    "Deleted": {
      "ReportLevel": "ERRORS_ONLY"
    },
    "Skipped": {
      "ReportLevel": "SUCCESSSES_AND_ERRORS"
    },
    "Transferred": {
      "ReportLevel": "ERRORS_ONLY"
    }
  },

```

```
        "Verified": {
            "ReportLevel": "ERRORS_ONLY"
        },
        "ReportLevel": "ERRORS_ONLY"
    }
}
```

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

## GenerateRecommendations

Creates recommendations about where to migrate your data to in AWS. Recommendations are generated based on information that DataSync Discovery collects about your on-premises storage system's resources. For more information, see [Recommendations provided by DataSync Discovery](#).

Once generated, you can view your recommendations by using the [DescribeStorageSystemResources](#) operation.

### Request Syntax

```
{  
  "DiscoveryJobArn": "string",  
  "ResourceIds": [ "string" ],  
  "ResourceType": "string"  
}
```

### Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

#### [DiscoveryJobArn](#)

Specifies the Amazon Resource Name (ARN) of the discovery job that collects information about your on-premises storage system.

Type: String

Length Constraints: Maximum length of 256.

Pattern: `^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}/job/discovery-job-[a-f0-9]{8}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{12}$`

Required: Yes

#### [ResourceIds](#)

Specifies the universally unique identifiers (UUIDs) of the resources in your storage system that you want recommendations on.

Type: Array of strings

Array Members: Minimum number of 1 item. Maximum number of 100 items.

Pattern: `[a-f0-9]{8}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{12}`

Required: Yes

### ResourceType

Specifies the type of resource in your storage system that you want recommendations on.

Type: String

Valid Values: SVM | VOLUME | CLUSTER

Required: Yes

## Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

## Errors

For information about the errors that are common to all actions, see [Common Errors](#).

### InternalException

This exception is thrown when an error occurs in the AWS DataSync service.

HTTP Status Code: 500

### InvalidRequestException

This exception is thrown when the client submits a malformed request.

HTTP Status Code: 400

## Examples

### Sample Request

The following example generates AWS storage recommendations for a volume in an on-premises storage system.

```
{
  "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"
}
```

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

## ListAgents

Returns a list of AWS DataSync agents that belong to an AWS account in the AWS Region specified in the request.

With pagination, you can reduce the number of agents returned in a response. If you get a truncated list of agents in a response, the response contains a marker that you can specify in your next request to fetch the next page of agents.

ListAgents is eventually consistent. This means the result of running the operation might not reflect that you just created or deleted an agent. For example, if you create an agent with [CreateAgent](#) and then immediately run ListAgents, that agent might not show up in the list right away. In situations like this, you can always confirm whether an agent has been created (or deleted) by using [DescribeAgent](#).

### Request Syntax

```
{
  "MaxResults": number,
  "NextToken": "string"
}
```

### Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

#### [MaxResults](#)

Specifies the maximum number of DataSync agents to list in a response. By default, a response shows a maximum of 100 agents.

Type: Integer

Valid Range: Minimum value of 0. Maximum value of 100.

Required: No

#### [NextToken](#)

Specifies an opaque string that indicates the position to begin the next list of results in the response.

Type: String

Length Constraints: Maximum length of 65535.

Pattern: [a-zA-Z0-9=\_- ]+

Required: No

## Response Syntax

```
{
  "Agents": [
    {
      "AgentArn": "string",
      "Name": "string",
      "Platform": {
        "Version": "string"
      },
      "Status": "string"
    }
  ],
  "NextToken": "string"
}
```

## Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

### Agents

A list of DataSync agents in your AWS account in the AWS Region specified in the request. The list is ordered by the agents' Amazon Resource Names (ARNs).

Type: Array of [AgentListEntry](#) objects

### NextToken

The opaque string that indicates the position to begin the next list of results in the response.

Type: String



Length Constraints: Maximum length of 65535.

Pattern: [a-zA-Z0-9=\_- ]+

## Errors

For information about the errors that are common to all actions, see [Common Errors](#).

### InternalException

This exception is thrown when an error occurs in the AWS DataSync service.

HTTP Status Code: 500

### InvalidRequestException

This exception is thrown when the client submits a malformed request.

HTTP Status Code: 400

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

## ListDiscoveryJobs

Provides a list of the existing discovery jobs in the AWS Region and AWS account where you're using DataSync Discovery.

### Request Syntax

```
{
  "MaxResults": number,
  "NextToken": "string",
  "StorageSystemArn": "string"
}
```

### Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

#### [MaxResults](#)

Specifies how many results you want in the response.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 100.

Required: No

#### [NextToken](#)

Specifies an opaque string that indicates the position to begin the next list of results in the response.

Type: String

Length Constraints: Maximum length of 65535.

Pattern: [a-zA-Z0-9=\_-]+

Required: No

## StorageSystemArn

Specifies the Amazon Resource Name (ARN) of an on-premises storage system. Use this parameter if you only want to list the discovery jobs that are associated with a specific storage system.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^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}$`

Required: No

## Response Syntax

```
{
  "DiscoveryJobs": [
    {
      "DiscoveryJobArn": "string",
      "Status": "string"
    }
  ],
  "NextToken": "string"
}
```

## Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

### DiscoveryJobs

The discovery jobs that you've run.

Type: Array of [DiscoveryJobListEntry](#) objects

### NextToken

The opaque string that indicates the position to begin the next list of results in the response.

Type: String

Length Constraints: Maximum length of 65535.

Pattern: [a-zA-Z0-9=\_- ]+

## Errors

For information about the errors that are common to all actions, see [Common Errors](#).

### InternalException

This exception is thrown when an error occurs in the AWS DataSync service.

HTTP Status Code: 500

### InvalidRequestException

This exception is thrown when the client submits a malformed request.

HTTP Status Code: 400

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

## ListLocations

Returns a list of source and destination locations.

If you have more locations than are returned in a response (that is, the response returns only a truncated list of your agents), the response contains a token that you can specify in your next request to fetch the next page of locations.

### Request Syntax

```
{
  "Filters": [
    {
      "Name": "string",
      "Operator": "string",
      "Values": [ "string" ]
    }
  ],
  "MaxResults": number,
  "NextToken": "string"
}
```

### Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

#### [Filters](#)

You can use API filters to narrow down the list of resources returned by `ListLocations`. For example, to retrieve all tasks on a specific source location, you can use `ListLocations` with filter name `LocationType S3` and `Operator Equals`.

Type: Array of [LocationFilter](#) objects

Required: No

#### [MaxResults](#)

The maximum number of locations to return.

Type: Integer

Valid Range: Minimum value of 0. Maximum value of 100.

Required: No

### NextToken

An opaque string that indicates the position at which to begin the next list of locations.

Type: String

Length Constraints: Maximum length of 65535.

Pattern: [a-zA-Z0-9=\_-]+

Required: No

## Response Syntax

```
{
  "Locations": [
    {
      "LocationArn": "string",
      "LocationUri": "string"
    }
  ],
  "NextToken": "string"
}
```

## Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

### Locations

An array that contains a list of locations.

Type: Array of [LocationListEntry](#) objects

### NextToken

An opaque string that indicates the position at which to begin returning the next list of locations.

Type: String

Length Constraints: Maximum length of 65535.

Pattern: [a-zA-Z0-9=\_- ]+

## Errors

For information about the errors that are common to all actions, see [Common Errors](#).

### InternalException

This exception is thrown when an error occurs in the AWS DataSync service.

HTTP Status Code: 500

### InvalidRequestException

This exception is thrown when the client submits a malformed request.

HTTP Status Code: 400

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

## ListStorageSystems

Lists the on-premises storage systems that you're using with DataSync Discovery.

### Request Syntax

```
{  
  "MaxResults": number,  
  "NextToken": "string"  
}
```

### Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

#### MaxResults

Specifies how many results you want in the response.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 100.

Required: No

#### NextToken

Specifies an opaque string that indicates the position to begin the next list of results in the response.

Type: String

Length Constraints: Maximum length of 65535.

Pattern: [a-zA-Z0-9=\_- ]+

Required: No

### Response Syntax

```
{
```



```
"NextToken": "string",
"StorageSystems": [
  {
    "Name": "string",
    "StorageSystemArn": "string"
  }
]
```

## Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

### NextToken

The opaque string that indicates the position to begin the next list of results in the response.

Type: String

Length Constraints: Maximum length of 65535.

Pattern: [a-zA-Z0-9=\_-]+

### StorageSystems

The Amazon Resource Names (ARNs) of the on-premises storage systems that you're using with DataSync Discovery.

Type: Array of [StorageSystemListEntry](#) objects

## Errors

For information about the errors that are common to all actions, see [Common Errors](#).

### **InternalException**

This exception is thrown when an error occurs in the AWS DataSync service.

HTTP Status Code: 500

### **InvalidRequestException**

This exception is thrown when the client submits a malformed request.

HTTP Status Code: 400

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

## ListTagsForResource

Returns all the tags associated with an AWS resource.

### Request Syntax

```
{
  "MaxResults": number,
  "NextToken": "string",
  "ResourceArn": "string"
}
```

### Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

#### MaxResults

Specifies how many results that you want in the response.

Type: Integer

Valid Range: Minimum value of 0. Maximum value of 100.

Required: No

#### NextToken

Specifies an opaque string that indicates the position to begin the next list of results in the response.

Type: String

Length Constraints: Maximum length of 65535.

Pattern: [a-zA-Z0-9=\_-]+

Required: No

#### ResourceArn

Specifies the Amazon Resource Name (ARN) of the resource that you want tag information on.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^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})?&`

Required: Yes

## Response Syntax

```
{
  "NextToken": "string",
  "Tags": [
    {
      "Key": "string",
      "Value": "string"
    }
  ]
}
```

## Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

### NextToken

The opaque string that indicates the position to begin the next list of results in the response.

Type: String

Length Constraints: Maximum length of 65535.

Pattern: `[a-zA-Z0-9=_-]+`

### Tags

An array of tags applied to the specified resource.

Type: Array of [TagListEntry](#) objects

Array Members: Minimum number of 0 items. Maximum number of 55 items.

## Errors

For information about the errors that are common to all actions, see [Common Errors](#).

### InternalException

This exception is thrown when an error occurs in the AWS DataSync service.

HTTP Status Code: 500

### InvalidRequestException

This exception is thrown when the client submits a malformed request.

HTTP Status Code: 400

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

## ListTaskExecutions

Returns a list of executions for an AWS DataSync transfer task.

### Request Syntax

```
{
  "MaxResults": number,
  "NextToken": "string",
  "TaskArn": "string"
}
```

### Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

#### MaxResults

Specifies how many results you want in the response.

Type: Integer

Valid Range: Minimum value of 0. Maximum value of 100.

Required: No

#### NextToken

Specifies an opaque string that indicates the position at which to begin the next list of results in the response.

Type: String

Length Constraints: Maximum length of 65535.

Pattern: [a-zA-Z0-9=\_-]+

Required: No

#### TaskArn

Specifies the Amazon Resource Name (ARN) of the task that you want execution information about.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^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}$`

Required: No

## Response Syntax

```
{
  "NextToken": "string",
  "TaskExecutions": [
    {
      "Status": "string",
      "TaskExecutionArn": "string"
    }
  ]
}
```

## Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

### NextToken

The opaque string that indicates the position to begin the next list of results in the response.

Type: String

Length Constraints: Maximum length of 65535.

Pattern: `[a-zA-Z0-9=_-]+`

### TaskExecutions

A list of the task's executions.

Type: Array of [TaskExecutionListEntry](#) objects

## Errors

For information about the errors that are common to all actions, see [Common Errors](#).

### InternalException

This exception is thrown when an error occurs in the AWS DataSync service.

HTTP Status Code: 500

### InvalidRequestException

This exception is thrown when the client submits a malformed request.

HTTP Status Code: 400

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)



## ListTasks

Returns a list of the AWS DataSync tasks you created.

### Request Syntax

```
{
  "Filters": [
    {
      "Name": "string",
      "Operator": "string",
      "Values": [ "string" ]
    }
  ],
  "MaxResults": number,
  "NextToken": "string"
}
```

### Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

#### Filters

You can use API filters to narrow down the list of resources returned by `ListTasks`. For example, to retrieve all tasks on a specific source location, you can use `ListTasks` with filter name `LocationId` and `Operator` `Equals` with the ARN for the location.

Type: Array of [TaskFilter](#) objects

Required: No

#### MaxResults

The maximum number of tasks to return.

Type: Integer

Valid Range: Minimum value of 0. Maximum value of 100.

Required: No

## NextToken

An opaque string that indicates the position at which to begin the next list of tasks.

Type: String

Length Constraints: Maximum length of 65535.

Pattern: [a-zA-Z0-9=\_-]+

Required: No

## Response Syntax

```
{
  "NextToken": "string",
  "Tasks": [
    {
      "Name": "string",
      "Status": "string",
      "TaskArn": "string"
    }
  ]
}
```

## Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

### NextToken

An opaque string that indicates the position at which to begin returning the next list of tasks.

Type: String

Length Constraints: Maximum length of 65535.

Pattern: [a-zA-Z0-9=\_-]+

### Tasks

A list of all the tasks that are returned.

Type: Array of [TaskListEntry](#) objects

## Errors

For information about the errors that are common to all actions, see [Common Errors](#).

### InternalException

This exception is thrown when an error occurs in the AWS DataSync service.

HTTP Status Code: 500

### InvalidRequestException

This exception is thrown when the client submits a malformed request.

HTTP Status Code: 400

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

## RemoveStorageSystem

Permanently removes a storage system resource from DataSync Discovery, including the associated discovery jobs, collected data, and recommendations.

### Request Syntax

```
{
  "StorageSystemArn": "string"
}
```

### Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

#### StorageSystemArn

Specifies the Amazon Resource Name (ARN) of the storage system that you want to permanently remove from DataSync Discovery.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^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}$`

Required: Yes

### Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

### Errors

For information about the errors that are common to all actions, see [Common Errors](#).

## InternalException

This exception is thrown when an error occurs in the AWS DataSync service.

HTTP Status Code: 500

## InvalidRequestException

This exception is thrown when the client submits a malformed request.

HTTP Status Code: 400

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

## StartDiscoveryJob

Runs a DataSync discovery job on your on-premises storage system. If you haven't added the storage system to DataSync Discovery yet, do this first by using the [AddStorageSystem](#) operation.

### Request Syntax

```
{
  "ClientToken": "string",
  "CollectionDurationMinutes": number,
  "StorageSystemArn": "string",
  "Tags": [
    {
      "Key": "string",
      "Value": "string"
    }
  ]
}
```

### Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

#### [ClientToken](#)

Specifies a client token to make sure requests with this API operation are idempotent. If you don't specify a client token, DataSync generates one for you automatically.

Type: String

Pattern: `[a-f0-9]{8}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{12}`

Required: Yes

#### [CollectionDurationMinutes](#)

Specifies in minutes how long you want the discovery job to run.

**Note**

For more accurate recommendations, we recommend a duration of at least 14 days. Longer durations allow time to collect a sufficient number of data points and provide a realistic representation of storage performance and utilization.

Type: Integer

Valid Range: Minimum value of 60. Maximum value of 44640.

Required: Yes

**StorageSystemArn**

Specifies the Amazon Resource Name (ARN) of the on-premises storage system that you want to run the discovery job on.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^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}$`

Required: Yes

**Tags**

Specifies labels that help you categorize, filter, and search for your AWS resources.

Type: Array of [TagListEntry](#) objects

Array Members: Minimum number of 0 items. Maximum number of 50 items.

Required: No

**Response Syntax**

```
{  
  "DiscoveryJobArn": "string"
```

```
}
```

## Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

### DiscoveryJobArn

The ARN of the discovery job that you started.

Type: String

Length Constraints: Maximum length of 256.

Pattern: `^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}$`

## Errors

For information about the errors that are common to all actions, see [Common Errors](#).

### **InternalException**

This exception is thrown when an error occurs in the AWS DataSync service.

HTTP Status Code: 500

### **InvalidRequestException**

This exception is thrown when the client submits a malformed request.

HTTP Status Code: 400

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:



- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

## StartTaskExecution

Starts an AWS DataSync transfer task. For each task, you can only run one task execution at a time.

There are several phases to a task execution. For more information, see [Task execution statuses](#).

### Important

If you're planning to transfer data to or from an Amazon S3 location, review [how DataSync can affect your S3 request charges](#) and the [DataSync pricing page](#) before you begin.

## Request Syntax

```
{
  "Excludes": [
    {
      "FilterType": "string",
      "Value": "string"
    }
  ],
  "Includes": [
    {
      "FilterType": "string",
      "Value": "string"
    }
  ],
  "ManifestConfig": {
    "Action": "string",
    "Format": "string",
    "Source": {
      "S3": {
        "BucketAccessRoleArn": "string",
        "ManifestObjectPath": "string",
        "ManifestObjectVersionId": "string",
        "S3BucketArn": "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",
  "TaskReportConfig": {
    "Destination": {
      "S3": {
        "BucketAccessRoleArn": "string",
        "S3BucketArn": "string",
        "Subdirectory": "string"
      }
    }
  },
  "ObjectVersionIds": "string",
  "OutputType": "string",
  "Overrides": {
    "Deleted": {
      "ReportLevel": "string"
    },
    "Skipped": {
      "ReportLevel": "string"
    },
    "Transferred": {
      "ReportLevel": "string"
    },
    "Verified": {
      "ReportLevel": "string"
    }
  }
}

```

```
    },  
    "ReportLevel": "string"  
  }  
}
```

## Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

### Excludes

Specifies a list of filter rules that determines which files to exclude from a task. The list contains a single filter string that consists of the patterns to exclude. The patterns are delimited by "|" (that is, a pipe), for example, "/folder1|/folder2".

Type: Array of [FilterRule](#) objects

Array Members: Minimum number of 0 items. Maximum number of 1 item.

Required: No

### Includes

Specifies a list of filter rules that determines which files to include when running a task. The pattern should contain a single filter string that consists of the patterns to include. The patterns are delimited by "|" (that is, a pipe), for example, "/folder1|/folder2".

Type: Array of [FilterRule](#) objects

Array Members: Minimum number of 0 items. Maximum number of 1 item.

Required: No

### ManifestConfig

Configures a manifest, which is a list of files or objects that you want DataSync to transfer. For more information and configuration examples, see [Specifying what DataSync transfers by using a manifest](#).

When using this parameter, your caller identity (the role that you're using DataSync with) must have the `iam:PassRole` permission. The [AWSDataSyncFullAccess](#) policy includes this permission.

To remove a manifest configuration, specify this parameter with an empty value.

Type: [ManifestConfig](#) object

Required: No

## [OverrideOptions](#)

Indicates how your transfer task is configured. These options include how DataSync handles files, objects, and their associated metadata during your transfer. You also can specify how to verify data integrity, set bandwidth limits for your task, among other options.

Each option has a default value. Unless you need to, you don't have to configure any option before calling [StartTaskExecution](#).

You also can override your task options for each task execution. For example, you might want to adjust the `LogLevel` for an individual execution.

Type: [Options](#) object

Required: No

## [Tags](#)

Specifies the tags that you want to apply to the Amazon Resource Name (ARN) representing the task execution.

*Tags* are key-value pairs that help you manage, filter, and search for your DataSync resources.

Type: Array of [TagListEntry](#) objects

Array Members: Minimum number of 0 items. Maximum number of 50 items.

Required: No

## [TaskArn](#)

Specifies the Amazon Resource Name (ARN) of the task that you want to start.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^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}$`

Required: Yes

### **TaskReportConfig**

Specifies how you want to configure a task report, which provides detailed information about your DataSync transfer. For more information, see [Monitoring your DataSync transfers with task reports](#).

When using this parameter, your caller identity (the role that you're using DataSync with) must have the `iam:PassRole` permission. The [AWSDataSyncFullAccess](#) policy includes this permission.

To remove a task report configuration, specify this parameter as empty.

Type: [TaskReportConfig](#) object

Required: No

## **Response Syntax**

```
{
  "TaskExecutionArn": "string"
}
```

## **Response Elements**

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

### **TaskExecutionArn**

The ARN of the running task execution.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^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}$`

## Errors

For information about the errors that are common to all actions, see [Common Errors](#).

### InternalException

This exception is thrown when an error occurs in the AWS DataSync service.

HTTP Status Code: 500

### InvalidRequestException

This exception is thrown when the client submits a malformed request.

HTTP Status Code: 400

## Examples

### Sample Request

The following example starts a task execution using the default options for the specified task.

```
{
  "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:datsync:us-east-2:111222333444:task/task-08de6e6697796f026"
}
```

### Sample Response

This example illustrates one usage of StartTaskExecution.

```
{
```

```
"TaskExecutionArn": "arn:aws:datsync:us-east-2:111222333444:task/
task-08de6e6697796f026/execution/exec-04ce9d516d69bd52f"
}
```

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)



## StopDiscoveryJob

Stops a running DataSync discovery job.

You can stop a discovery job anytime. A job that's stopped before it's scheduled to end likely will provide you some information about your on-premises storage system resources. To get recommendations for a stopped job, you must use the [GenerateRecommendations](#) operation.

### Request Syntax

```
{
  "DiscoveryJobArn": "string"
}
```

### Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

#### [DiscoveryJobArn](#)

Specifies the Amazon Resource Name (ARN) of the discovery job that you want to stop.

Type: String

Length Constraints: Maximum length of 256.

Pattern: `^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}/job/discovery-job-[a-f0-9]{8}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{12}$`

Required: Yes

### Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

### Errors

For information about the errors that are common to all actions, see [Common Errors](#).

## InternalException

This exception is thrown when an error occurs in the AWS DataSync service.

HTTP Status Code: 500

## InvalidRequestException

This exception is thrown when the client submits a malformed request.

HTTP Status Code: 400

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

## TagResource

Applies a *tag* to an AWS resource. Tags are key-value pairs that can help you manage, filter, and search for your resources.

These include AWS DataSync resources, such as locations, tasks, and task executions.

### Request Syntax

```
{
  "ResourceArn": "string",
  "Tags": [
    {
      "Key": "string",
      "Value": "string"
    }
  ]
}
```

### Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

#### ResourceArn

Specifies the Amazon Resource Name (ARN) of the resource to apply the tag to.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^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})?&`

Required: Yes

#### Tags

Specifies the tags that you want to apply to the resource.

Type: Array of [TagListEntry](#) objects

Array Members: Minimum number of 0 items. Maximum number of 50 items.

Required: Yes

## Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

## Errors

For information about the errors that are common to all actions, see [Common Errors](#).

### InternalException

This exception is thrown when an error occurs in the AWS DataSync service.

HTTP Status Code: 500

### InvalidRequestException

This exception is thrown when the client submits a malformed request.

HTTP Status Code: 400

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)

- [AWS SDK for Ruby V3](#)

# UntagResource

Removes tags from an AWS resource.

## Request Syntax

```
{  
  "Keys": [ "string" ],  
  "ResourceArn": "string"  
}
```

## Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

### Keys

Specifies the keys in the tags that you want to remove.

Type: Array of strings

Array Members: Minimum number of 1 item. Maximum number of 50 items.

Length Constraints: Minimum length of 1. Maximum length of 256.

Pattern: `^[a-zA-Z0-9\s+=. _:/-]+$`

Required: Yes

### ResourceArn

Specifies the Amazon Resource Name (ARN) of the resource to remove the tags from.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^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})?&`

Required: Yes

## Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

## Errors

For information about the errors that are common to all actions, see [Common Errors](#).

### InternalException

This exception is thrown when an error occurs in the AWS DataSync service.

HTTP Status Code: 500

### InvalidRequestException

This exception is thrown when the client submits a malformed request.

HTTP Status Code: 400

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

# UpdateAgent

Updates the name of an AWS DataSync agent.

## Request Syntax

```
{  
  "AgentArn": "string",  
  "Name": "string"  
}
```

## Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

### AgentArn

The Amazon Resource Name (ARN) of the agent to update.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^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}$`

Required: Yes

### Name

The name that you want to use to configure the agent.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 256.

Pattern: `^[a-zA-Z0-9\s+=. _:@/-]+$`

Required: No



## Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

## Errors

For information about the errors that are common to all actions, see [Common Errors](#).

### InternalException

This exception is thrown when an error occurs in the AWS DataSync service.

HTTP Status Code: 500

### InvalidRequestException

This exception is thrown when the client submits a malformed request.

HTTP Status Code: 400

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

# UpdateDiscoveryJob

Edits a DataSync discovery job configuration.

## Request Syntax

```
{
  "CollectionDurationMinutes": number,
  "DiscoveryJobArn": "string"
}
```

## Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

### CollectionDurationMinutes

Specifies in minutes how long that you want the discovery job to run. (You can't set this parameter to less than the number of minutes that the job has already run for.)

Type: Integer

Valid Range: Minimum value of 60. Maximum value of 44640.

Required: Yes

### DiscoveryJobArn

Specifies the Amazon Resource Name (ARN) of the discovery job that you want to update.

Type: String

Length Constraints: Maximum length of 256.

Pattern: `^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}/job/discovery-job-[a-f0-9]{8}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{12}$`

Required: Yes

## Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

## Errors

For information about the errors that are common to all actions, see [Common Errors](#).

### InternalException

This exception is thrown when an error occurs in the AWS DataSync service.

HTTP Status Code: 500

### InvalidRequestException

This exception is thrown when the client submits a malformed request.

HTTP Status Code: 400

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

## UpdateLocationAzureBlob

Modifies some configurations of the Microsoft Azure Blob Storage transfer location that you're using with AWS DataSync.

### Request Syntax

```
{
  "AccessTier": "string",
  "AgentArns": [ "string" ],
  "AuthenticationType": "string",
  "BlobType": "string",
  "LocationArn": "string",
  "SasConfiguration": {
    "Token": "string"
  },
  "Subdirectory": "string"
}
```

### Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

#### [AccessTier](#)

Specifies the access tier that you want your objects or files transferred into. This only applies when using the location as a transfer destination. For more information, see [Access tiers](#).

Type: String

Valid Values: HOT | COOL | ARCHIVE

Required: No

#### [AgentArns](#)

Specifies the Amazon Resource Name (ARN) of the DataSync agent that can connect with your Azure Blob Storage container.

You can specify more than one agent. For more information, see [Using multiple agents for your transfer](#).

Type: Array of strings

Array Members: Minimum number of 1 item. Maximum number of 4 items.

Length Constraints: Maximum length of 128.

Pattern: `^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}$`

Required: No

### AuthenticationType

Specifies the authentication method DataSync uses to access your Azure Blob Storage. DataSync can access blob storage using a shared access signature (SAS).

Type: String

Valid Values: SAS

Required: No

### BlobType

Specifies the type of blob that you want your objects or files to be when transferring them into Azure Blob Storage. Currently, DataSync only supports moving data into Azure Blob Storage as block blobs. For more information on blob types, see the [Azure Blob Storage documentation](#).

Type: String

Valid Values: BLOCK

Required: No

### LocationArn

Specifies the ARN of the Azure Blob Storage transfer location that you're updating.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^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}$`

Required: Yes

### SasConfiguration

Specifies the SAS configuration that allows DataSync to access your Azure Blob Storage.

Type: [AzureBlobSasConfiguration](#) object

Required: No

### Subdirectory

Specifies path segments if you want to limit your transfer to a virtual directory in your container (for example, /my/images).

Type: String

Length Constraints: Maximum length of 1024.

Pattern: `^[\\p{L}\\p{M}\\p{Z}\\p{S}\\p{N}\\p{P}\\p{C}]*$`

Required: No

## Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

## Errors

For information about the errors that are common to all actions, see [Common Errors](#).

### InternalException

This exception is thrown when an error occurs in the AWS DataSync service.

HTTP Status Code: 500

### InvalidRequestException

This exception is thrown when the client submits a malformed request.

HTTP Status Code: 400

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

## UpdateLocationHdfs

Updates some parameters of a previously created location for a Hadoop Distributed File System cluster.

### Request Syntax

```
{
  "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"
}
```

### Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

#### AgentArns

The ARNs of the agents that are used to connect to the HDFS cluster.

Type: Array of strings



Array Members: Minimum number of 1 item. Maximum number of 4 items.

Length Constraints: Maximum length of 128.

Pattern: `^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}$`

Required: No

### AuthenticationType

The type of authentication used to determine the identity of the user.

Type: String

Valid Values: SIMPLE | KERBEROS

Required: No

### BlockSize

The size of the data blocks to write into the HDFS cluster.

Type: Integer

Valid Range: Minimum value of 1048576. Maximum value of 1073741824.

Required: No

### KerberosKeytab

The Kerberos key table (keytab) that contains mappings between the defined Kerberos principal and the encrypted keys. You can load the keytab from a file by providing the file's address. If you use the AWS CLI, it performs base64 encoding for you. Otherwise, provide the base64-encoded text.

Type: Base64-encoded binary data object

Length Constraints: Maximum length of 65536.

Required: No

### KerberosKrb5Conf

The `krb5.conf` file that contains the Kerberos configuration information. You can load the `krb5.conf` file by providing the file's address. If you're using the AWS CLI, it performs the base64 encoding for you. Otherwise, provide the base64-encoded text.

Type: Base64-encoded binary data object

Length Constraints: Maximum length of 131072.

Required: No

### KerberosPrincipal

The Kerberos principal with access to the files and folders on the HDFS cluster.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 256.

Pattern: ^.+

Required: No

### KmsKeyProviderUri

The URI of the HDFS cluster's Key Management Server (KMS).

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: ^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\$

Required: No

### LocationArn

The Amazon Resource Name (ARN) of the source HDFS cluster location.

Type: String

Length Constraints: Maximum length of 128.

Pattern: ^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}\$

Required: Yes

## NameNodes

The NameNode that manages the HDFS namespace. The NameNode performs operations such as opening, closing, and renaming files and directories. The NameNode contains the information to map blocks of data to the DataNodes. You can use only one NameNode.

Type: Array of [HdfsNameNode](#) objects

Array Members: Minimum number of 1 item.

Required: No

## QopConfiguration

The Quality of Protection (QOP) configuration specifies the Remote Procedure Call (RPC) and data transfer privacy settings configured on the Hadoop Distributed File System (HDFS) cluster.

Type: [QopConfiguration](#) object

Required: No

## ReplicationFactor

The number of DataNodes to replicate the data to when writing to the HDFS cluster.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 512.

Required: No

## SimpleUser

The user name used to identify the client on the host operating system.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 256.

Pattern: `^[_ .A-Za-z0-9][-_ .A-Za-z0-9]*$`

Required: No

## Subdirectory

A subdirectory in the HDFS cluster. This subdirectory is used to read data from or write data to the HDFS cluster.

Type: String

Length Constraints: Maximum length of 4096.

Pattern: `^[a-zA-Z0-9_\- \+\.\/(\)\$\p{Zs}]+$`

Required: No

## Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

## Errors

For information about the errors that are common to all actions, see [Common Errors](#).

### InternalException

This exception is thrown when an error occurs in the AWS DataSync service.

HTTP Status Code: 500

### InvalidRequestException

This exception is thrown when the client submits a malformed request.

HTTP Status Code: 400

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)

- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

## UpdateLocationNfs

Modifies some configurations of the Network File System (NFS) transfer location that you're using with AWS DataSync.

For more information, see [Configuring transfers to or from an NFS file server](#).

### Request Syntax

```
{
  "LocationArn": "string",
  "MountOptions": {
    "Version": "string"
  },
  "OnPremConfig": {
    "AgentArns": [ "string" ]
  },
  "Subdirectory": "string"
}
```

### Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

#### [LocationArn](#)

Specifies the Amazon Resource Name (ARN) of the NFS transfer location that you want to update.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^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}$`

Required: Yes

#### [MountOptions](#)

Specifies how DataSync can access a location using the NFS protocol.

Type: [NfsMountOptions](#) object

Required: No

### [OnPremConfig](#)

The AWS DataSync agents that are connecting to a Network File System (NFS) location.

Type: [OnPremConfig](#) object

Required: No

### [Subdirectory](#)

Specifies the export path in your NFS file server that you want DataSync to mount.

This path (or a subdirectory of the path) is where DataSync transfers data to or from. For information on configuring an export for DataSync, see [Accessing NFS file servers](#).

Type: String

Length Constraints: Maximum length of 4096.

Pattern: `^[a-zA-Z0-9_\-\.\/\(\)\p{Zs}]+$`

Required: No

## Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

## Errors

For information about the errors that are common to all actions, see [Common Errors](#).

### **InternalException**

This exception is thrown when an error occurs in the AWS DataSync service.

HTTP Status Code: 500

### **InvalidRequestException**

This exception is thrown when the client submits a malformed request.

HTTP Status Code: 400

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)



# UpdateLocationObjectStorage

Updates some parameters of an existing AWS DataSync location for an object storage system.

## Request Syntax

```
{
  "AccessKey": "string",
  "AgentArns": [ "string" ],
  "LocationArn": "string",
  "SecretKey": "string",
  "ServerCertificate": blob,
  "ServerPort": number,
  "ServerProtocol": "string",
  "Subdirectory": "string"
}
```

## Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

### [AccessKey](#)

Specifies the access key (for example, a user name) if credentials are required to authenticate with the object storage server.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 200.

Pattern: `^.*$`

Required: No

### [AgentArns](#)

Specifies the Amazon Resource Names (ARNs) of the DataSync agents that can securely connect with your location.

Type: Array of strings

Array Members: Minimum number of 1 item. Maximum number of 4 items.

Length Constraints: Maximum length of 128.

Pattern: `^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}$`

Required: No

### LocationArn

Specifies the ARN of the object storage system location that you're updating.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^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}$`

Required: Yes

### SecretKey

Specifies the secret key (for example, a password) if credentials are required to authenticate with the object storage server.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 200.

Pattern: `^.*$`

Required: No

### ServerCertificate

Specifies a certificate chain for DataSync to authenticate with your object storage system if the system uses a private or self-signed certificate authority (CA). You must specify a single .pem file with a full certificate chain (for example, `file:///home/user/.ssh/object_storage_certificates.pem`).

The certificate chain might include:

- The object storage system's certificate
- All intermediate certificates (if there are any)

- The root certificate of the signing CA

You can concatenate your certificates into a `.pem` file (which can be up to 32768 bytes before base64 encoding). The following example `cat` command creates an `object_storage_certificates.pem` file that includes three certificates:

```
cat object_server_certificate.pem intermediate_certificate.pem  
ca_root_certificate.pem > object_storage_certificates.pem
```

To use this parameter, configure `ServerProtocol` to `HTTPS`.

Updating this parameter doesn't interfere with tasks that you have in progress.

Type: Base64-encoded binary data object

Length Constraints: Maximum length of 32768.

Required: No

### ServerPort

Specifies the port that your object storage server accepts inbound network traffic on (for example, port 443).

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 65536.

Required: No

### ServerProtocol

Specifies the protocol that your object storage server uses to communicate.

Type: String

Valid Values: `HTTPS` | `HTTP`

Required: No

### Subdirectory

Specifies the object prefix for your object storage server. If this is a source location, DataSync only copies objects with this prefix. If this is a destination location, DataSync writes all objects with this prefix.

Type: String

Length Constraints: Maximum length of 4096.

Pattern: `^[a-zA-Z0-9_\-\.\/(\)\p{Zs}]*$`

Required: No

## Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

## Errors

For information about the errors that are common to all actions, see [Common Errors](#).

### InternalException

This exception is thrown when an error occurs in the AWS DataSync service.

HTTP Status Code: 500

### InvalidRequestException

This exception is thrown when the client submits a malformed request.

HTTP Status Code: 400

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)

- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

## UpdateLocationSmb

Updates some of the parameters of a Server Message Block (SMB) file server location that you can use for AWS DataSync transfers.

### Request Syntax

```
{
  "AgentArns": [ "string" ],
  "Domain": "string",
  "LocationArn": "string",
  "MountOptions": {
    "Version": "string"
  },
  "Password": "string",
  "Subdirectory": "string",
  "User": "string"
}
```

### Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

#### AgentArns

Specifies the DataSync agent (or agents) which you want to connect to your SMB file server. You specify an agent by using its Amazon Resource Name (ARN).

Type: Array of strings

Array Members: Minimum number of 1 item. Maximum number of 4 items.

Length Constraints: Maximum length of 128.

Pattern: `^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}$`

Required: No

#### Domain

Specifies the Windows domain name that your SMB file server belongs to.

If you have multiple domains in your environment, configuring this parameter makes sure that DataSync connects to the right file server.

For more information, see [required permissions](#) for SMB locations.

Type: String

Length Constraints: Maximum length of 253.

Pattern: `^[A-Za-z0-9](\.|-+)?[A-Za-z0-9]{0,252}$`

Required: No

### LocationArn

Specifies the ARN of the SMB location that you want to update.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^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}$`

Required: Yes

### MountOptions

Specifies the version of the Server Message Block (SMB) protocol that AWS DataSync uses to access an SMB file server.

Type: [SmbMountOptions](#) object

Required: No

### Password

Specifies the password of the user who can mount your SMB file server and has permission to access the files and folders involved in your transfer.

For more information, see [required permissions](#) for SMB locations.

Type: String

Length Constraints: Maximum length of 104.

Pattern: `^\.{0,104}$`

Required: No

### Subdirectory

Specifies the name of the share exported by your SMB file server where DataSync will read or write data. You can include a subdirectory in the share path (for example, `/path/to/subdirectory`). Make sure that other SMB clients in your network can also mount this path.

To copy all data in the specified subdirectory, DataSync must be able to mount the SMB share and access all of its data. For more information, see [required permissions](#) for SMB locations.

Type: String

Length Constraints: Maximum length of 4096.

Pattern: `^[a-zA-Z0-9_\-\+\.\(\)\$\p{Zs}]+$`

Required: No

### User

Specifies the user name that can mount your SMB file server and has permission to access the files and folders involved in your transfer.

For information about choosing a user with the right level of access for your transfer, see [required permissions](#) for SMB locations.

Type: String

Length Constraints: Maximum length of 104.

Pattern: `^[^\\x5B\\x5D\\\/:;|=, +*?]{1,104}$`

Required: No

## Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

## Errors

For information about the errors that are common to all actions, see [Common Errors](#).



## InternalException

This exception is thrown when an error occurs in the AWS DataSync service.

HTTP Status Code: 500

## InvalidRequestException

This exception is thrown when the client submits a malformed request.

HTTP Status Code: 400

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

# UpdateStorageSystem

Modifies some configurations of an on-premises storage system resource that you're using with DataSync Discovery.

## Request Syntax

```
{
  "AgentArns": [ "string" ],
  "CloudWatchLogGroupArn": "string",
  "Credentials": {
    "Password": "string",
    "Username": "string"
  },
  "Name": "string",
  "ServerConfiguration": {
    "ServerHostname": "string",
    "ServerPort": number
  },
  "StorageSystemArn": "string"
}
```

## Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

### AgentArns

Specifies the Amazon Resource Name (ARN) of the DataSync agent that connects to and reads your on-premises storage system. You can only specify one ARN.

Type: Array of strings

Array Members: Fixed number of 1 item.

Length Constraints: Maximum length of 128.

Pattern: `^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}$`

Required: No

### CloudWatchLogGroupArn

Specifies the ARN of the Amazon CloudWatch log group for monitoring and logging discovery job events.

Type: String

Length Constraints: Maximum length of 562.

Pattern: `^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):logs:[a-z\-\0-9]+:[0-9]{12}:log-group:([^\:]*)(:\:)*?$`

Required: No

### Credentials

Specifies the user name and password for accessing your on-premises storage system's management interface.

Type: [Credentials](#) object

Required: No

### Name

Specifies a familiar name for your on-premises storage system.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 256.

Pattern: `^[a-zA-Z0-9\-\_:\@\/-]+$`

Required: No

### ServerConfiguration

Specifies the server name and network port required to connect with your on-premises storage system's management interface.

Type: [DiscoveryServerConfiguration](#) object

Required: No

## StorageSystemArn

Specifies the ARN of the on-premises storage system that you want reconfigure.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^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}$`

Required: Yes

## Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

## Errors

For information about the errors that are common to all actions, see [Common Errors](#).

### InternalException

This exception is thrown when an error occurs in the AWS DataSync service.

HTTP Status Code: 500

### InvalidRequestException

This exception is thrown when the client submits a malformed request.

HTTP Status Code: 400

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)

- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

# UpdateTask

Updates the configuration of a *task*, which defines where and how AWS DataSync transfers your data.

## Request Syntax

```
{
  "CloudWatchLogGroupArn": "string",
  "Excludes": [
    {
      "FilterType": "string",
      "Value": "string"
    }
  ],
  "Includes": [
    {
      "FilterType": "string",
      "Value": "string"
    }
  ],
  "ManifestConfig": {
    "Action": "string",
    "Format": "string",
    "Source": {
      "S3": {
        "BucketAccessRoleArn": "string",
        "ManifestObjectPath": "string",
        "ManifestObjectVersionId": "string",
        "S3BucketArn": "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",
    "Status": "string"
  },
  "TaskArn": "string",
  "TaskReportConfig": {
    "Destination": {
      "S3": {
        "BucketAccessRoleArn": "string",
        "S3BucketArn": "string",
        "Subdirectory": "string"
      }
    }
  },
  "ObjectVersionIds": "string",
  "OutputType": "string",
  "Overrides": {
    "Deleted": {
      "ReportLevel": "string"
    },
    "Skipped": {
      "ReportLevel": "string"
    },
    "Transferred": {
      "ReportLevel": "string"
    },
    "Verified": {
      "ReportLevel": "string"
    }
  },
  "ReportLevel": "string"
}
}

```

## Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

### [CloudWatchLogGroupArn](#)

Specifies the Amazon Resource Name (ARN) of an Amazon CloudWatch log group for monitoring your task.

For more information, see [Monitoring DataSync with Amazon CloudWatch](#).

Type: String

Length Constraints: Maximum length of 562.

Pattern: `^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):logs:[a-z\-\0-9]+:[0-9]{12}:log-group:([^\:]*)(:\:)*?$`

Required: No

### [Excludes](#)

Specifies exclude filters that define the files, objects, and folders in your source location that you don't want DataSync to transfer. For more information and examples, see [Specifying what DataSync transfers by using filters](#).

Type: Array of [FilterRule](#) objects

Array Members: Minimum number of 0 items. Maximum number of 1 item.

Required: No

### [Includes](#)

Specifies include filters define the files, objects, and folders in your source location that you want DataSync to transfer. For more information and examples, see [Specifying what DataSync transfers by using filters](#).

Type: Array of [FilterRule](#) objects

Array Members: Minimum number of 0 items. Maximum number of 1 item.



Required: No

## **ManifestConfig**

Configures a manifest, which is a list of files or objects that you want DataSync to transfer. For more information and configuration examples, see [Specifying what DataSync transfers by using a manifest](#).

When using this parameter, your caller identity (the IAM role that you're using DataSync with) must have the `iam:PassRole` permission. The [AWSDataSyncFullAccess](#) policy includes this permission.

To remove a manifest configuration, specify this parameter as empty.

Type: [ManifestConfig](#) object

Required: No

## **Name**

Specifies the name of your task.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 256.

Pattern: `^[a-zA-Z0-9\s+=. _:@/-]+$`

Required: No

## **Options**

Indicates how your transfer task is configured. These options include how DataSync handles files, objects, and their associated metadata during your transfer. You also can specify how to verify data integrity, set bandwidth limits for your task, among other options.

Each option has a default value. Unless you need to, you don't have to configure any option before calling [StartTaskExecution](#).

You also can override your task options for each task execution. For example, you might want to adjust the `LogLevel` for an individual execution.

Type: [Options](#) object

Required: No

### Schedule

Specifies a schedule for when you want your task to run. For more information, see [Scheduling your task](#).

Type: [TaskSchedule](#) object

Required: No

### TaskArn

Specifies the ARN of the task that you want to update.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^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}$`

Required: Yes

### TaskReportConfig

Specifies how you want to configure a task report, which provides detailed information about your DataSync transfer. For more information, see [Monitoring your DataSync transfers with task reports](#).

When using this parameter, your caller identity (the IAM role that you're using DataSync with) must have the `iam:PassRole` permission. The [AWSDataSyncFullAccess](#) policy includes this permission.

To remove a task report configuration, specify this parameter as empty.

Type: [TaskReportConfig](#) object

Required: No

## Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

## Errors

For information about the errors that are common to all actions, see [Common Errors](#).

### InternalException

This exception is thrown when an error occurs in the AWS DataSync service.

HTTP Status Code: 500

### InvalidRequestException

This exception is thrown when the client submits a malformed request.

HTTP Status Code: 400

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

# UpdateTaskExecution

Updates the configuration of a running AWS DataSync task execution.

## Note

Currently, the only Option that you can modify with UpdateTaskExecution is [BytesPerSecond](#) , which throttles bandwidth for a running or queued task execution.

## Request Syntax

```
{
  "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"
}
```

## Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

## Options

Indicates how your transfer task is configured. These options include how DataSync handles files, objects, and their associated metadata during your transfer. You also can specify how to verify data integrity, set bandwidth limits for your task, among other options.

Each option has a default value. Unless you need to, you don't have to configure any option before calling [StartTaskExecution](#).

You also can override your task options for each task execution. For example, you might want to adjust the `LogLevel` for an individual execution.

Type: [Options](#) object

Required: Yes

## TaskExecutionArn

Specifies the Amazon Resource Name (ARN) of the task execution that you're updating.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^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}$`

Required: Yes

## Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

## Errors

For information about the errors that are common to all actions, see [Common Errors](#).

### InternalException

This exception is thrown when an error occurs in the AWS DataSync service.

HTTP Status Code: 500

## InvalidRequestException

This exception is thrown when the client submits a malformed request.

HTTP Status Code: 400

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

## Data Types

The following data types are supported:

- [AgentListEntry](#)
- [AzureBlobSasConfiguration](#)
- [Capacity](#)
- [Credentials](#)
- [DiscoveryJobListEntry](#)
- [DiscoveryServerConfiguration](#)
- [Ec2Config](#)
- [FilterRule](#)
- [FsxProtocol](#)

- [FsxProtocolNfs](#)
- [FsxProtocolSmb](#)
- [HdfsNameNode](#)
- [IOPS](#)
- [Latency](#)
- [LocationFilter](#)
- [LocationListEntry](#)
- [ManifestConfig](#)
- [MaxP95Performance](#)
- [NetAppONTAPCluster](#)
- [NetAppONTAPSVM](#)
- [NetAppONTAPVolume](#)
- [NfsMountOptions](#)
- [OnPremConfig](#)
- [Options](#)
- [P95Metrics](#)
- [Platform](#)
- [PrivateLinkConfig](#)
- [QopConfiguration](#)
- [Recommendation](#)
- [ReportDestination](#)
- [ReportDestinationS3](#)
- [ReportOverride](#)
- [ReportOverrides](#)
- [ReportResult](#)
- [ResourceDetails](#)
- [ResourceMetrics](#)
- [S3Config](#)
- [S3ManifestConfig](#)
- [SmbMountOptions](#)

- [SourceManifestConfig](#)
- [StorageSystemListEntry](#)
- [TagListEntry](#)
- [TaskExecutionListEntry](#)
- [TaskExecutionResultDetail](#)
- [TaskFilter](#)
- [TaskListEntry](#)
- [TaskReportConfig](#)
- [TaskSchedule](#)
- [TaskScheduleDetails](#)
- [Throughput](#)



## AgentListEntry

Represents a single entry in a list (or array) of AWS DataSync agents when you call the [ListAgents](#) operation.

### Contents

#### AgentArn

The Amazon Resource Name (ARN) of a DataSync agent.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^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}$`

Required: No

#### Name

The name of an agent.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 256.

Pattern: `^[a-zA-Z0-9\s+=. _:@/-]+$`

Required: No

#### Platform

The platform-related details about the agent, such as the version number.

Type: [Platform](#) object

Required: No

#### Status

The status of an agent.

- If the status is ONLINE, the agent is configured properly and ready to use.

- If the status is OFFLINE, the agent has been out of contact with DataSync for five minutes or longer. This can happen for a few reasons. For more information, see [What do I do if my agent is offline?](#)

Type: String

Valid Values: ONLINE | OFFLINE

Required: No

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

# AzureBlobSasConfiguration

The shared access signature (SAS) configuration that allows AWS DataSync to access your Microsoft Azure Blob Storage.

For more information, see [SAS tokens](#) for accessing your Azure Blob Storage.

## Contents

### Token

Specifies a SAS token that provides permissions to access your Azure Blob Storage.

The token is part of the SAS URI string that comes after the storage resource URI and a question mark. A token looks something like this:

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

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `^.+`

Required: Yes

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

## Capacity

The storage capacity of an on-premises storage system resource (for example, a volume).

### Contents

#### ClusterCloudStorageUsed

The amount of space in the cluster that's in cloud storage (for example, if you're using data tiering).

Type: Long

Valid Range: Minimum value of 0.

Required: No

#### LogicalUsed

The amount of space that's being used in a storage system resource without accounting for compression or deduplication.

Type: Long

Valid Range: Minimum value of 0.

Required: No

#### Provisioned

The total amount of space available in a storage system resource.

Type: Long

Valid Range: Minimum value of 0.

Required: No

#### Used

The amount of space that's being used in a storage system resource.

Type: Long

Valid Range: Minimum value of 0.

Required: No

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

# Credentials

The credentials that provide DataSync Discovery read access to your on-premises storage system's management interface.

DataSync Discovery stores these credentials in [AWS Secrets Manager](#). For more information, see [Accessing your on-premises storage system](#).

## Contents

### Password

Specifies the password for your storage system's management interface.

Type: String

Length Constraints: Maximum length of 1024.

Pattern: `^(?!.*[:\"])[^:"]*$).+$`

Required: Yes

### Username

Specifies the user name for your storage system's management interface.

Type: String

Length Constraints: Maximum length of 1024.

Pattern: `^(?!.*[:\"])[^:"]*$).+$`

Required: Yes

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)



# DiscoveryJobListEntry

The details about a specific DataSync discovery job.

## Contents

### DiscoveryJobArn

The Amazon Resource Name (ARN) of a discovery job.

Type: String

Length Constraints: Maximum length of 256.

Pattern: `^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}$`

Required: No

### Status

The status of a discovery job. For more information, see [Discovery job statuses](#).

Type: String

Valid Values: RUNNING | WARNING | TERMINATED | FAILED | STOPPED | COMPLETED | COMPLETED\_WITH\_ISSUES

Required: No

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)



# DiscoveryServerConfiguration

The network settings that DataSync Discovery uses to connect with your on-premises storage system's management interface.

## Contents

### ServerHostname

The domain name or IP address of your storage system's management interface.

Type: String

Length Constraints: Maximum length of 255.

Pattern: `^(([a-zA-Z0-9\-\ ]*[a-zA-Z0-9])\.)*([A-Za-z0-9\-\ ]*[A-Za-z0-9])$`

Required: Yes

### ServerPort

The network port for accessing the storage system's management interface.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 65535.

Required: No

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

## Ec2Config

The subnet and security groups that AWS DataSync uses to access your Amazon EFS file system.

### Contents

#### SecurityGroupArns

Specifies the Amazon Resource Names (ARNs) of the security groups associated with an Amazon EFS file system's mount target.

Type: Array of strings

Array Members: Minimum number of 1 item. Maximum number of 5 items.

Length Constraints: Maximum length of 128.

Pattern: `^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]+$`

Required: Yes

#### SubnetArn

Specifies the ARN of a subnet where DataSync creates the [network interfaces](#) for managing traffic during your transfer.

The subnet must be located:

- In the same virtual private cloud (VPC) as the Amazon EFS file system.
- In the same Availability Zone as at least one mount target for the Amazon EFS file system.

#### Note

You don't need to specify a subnet that includes a file system mount target.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):ec2:[a-z\-\0-9]*:[0-9]{12}:subnet/.*$`

Required: Yes

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

## FilterRule

Specifies which files, folders, and objects to include or exclude when transferring files from source to destination.

### Contents

#### FilterType

The type of filter rule to apply. AWS DataSync only supports the SIMPLE\_PATTERN rule type.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^[A-Z0-9_]+$`

Valid Values: SIMPLE\_PATTERN

Required: No

#### Value

A single filter string that consists of the patterns to include or exclude. The patterns are delimited by "|" (that is, a pipe), for example: `/folder1|/folder2`

Type: String

Length Constraints: Maximum length of 102400.

Pattern: `^[^\x00]+$`

Required: No

### See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)

- [AWS SDK for Ruby V3](#)

# FsxProtocol

Specifies the data transfer protocol that AWS DataSync uses to access your Amazon FSx file system.

## Contents

### NFS

Specifies the Network File System (NFS) protocol configuration that DataSync uses to access your FSx for OpenZFS file system or FSx for ONTAP file system's storage virtual machine (SVM).

Type: [FsxProtocolNfs](#) object

Required: No

### SMB

Specifies the Server Message Block (SMB) protocol configuration that DataSync uses to access your FSx for ONTAP file system's SVM.

Type: [FsxProtocolSmb](#) object

Required: No

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

## FsxProtocolNfs

Specifies the Network File System (NFS) protocol configuration that AWS DataSync uses to access your Amazon FSx for OpenZFS or Amazon FSx for NetApp ONTAP file system.

### Contents

#### MountOptions

Specifies how DataSync can access a location using the NFS protocol.

Type: [NfsMountOptions](#) object

Required: No

### See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

## FsxProtocolSmb

Specifies the Server Message Block (SMB) protocol configuration that AWS DataSync uses to access your Amazon FSx for NetApp ONTAP file system. For more information, see [Accessing FSx for ONTAP file systems](#).

### Contents

#### Password

Specifies the password of a user who has permission to access your SVM.

Type: String

Length Constraints: Maximum length of 104.

Pattern: `^.{0,104}$`

Required: Yes

#### User

Specifies a user that can mount and access the files, folders, and metadata in your SVM.

For information about choosing a user with the right level of access for your transfer, see [Using the SMB protocol](#).

Type: String

Length Constraints: Maximum length of 104.

Pattern: `^[^\x5B\x5D\\\/:;|=,+*?]{1,104}$`

Required: Yes

#### Domain

Specifies the fully qualified domain name (FQDN) of the Microsoft Active Directory that your storage virtual machine (SVM) belongs to.

If you have multiple domains in your environment, configuring this setting makes sure that DataSync connects to the right SVM.

Type: String



Length Constraints: Maximum length of 253.

Pattern: `^[A-Za-z0-9](\.|-+)?[A-Za-z0-9]{0,252}$`

Required: No

## MountOptions

Specifies the version of the Server Message Block (SMB) protocol that AWS DataSync uses to access an SMB file server.

Type: [SmbMountOptions](#) object

Required: No

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

# HdfsNameNode

The NameNode of the Hadoop Distributed File System (HDFS). The NameNode manages the file system's namespace. The NameNode performs operations such as opening, closing, and renaming files and directories. The NameNode contains the information to map blocks of data to the DataNodes.

## Contents

### Hostname

The hostname of the NameNode in the HDFS cluster. This value is the IP address or Domain Name Service (DNS) name of the NameNode. An agent that's installed on-premises uses this hostname to communicate with the NameNode in the network.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `^(([a-zA-Z0-9\-.]*[a-zA-Z0-9])\.)*([A-Za-z0-9\-.]*[A-Za-z0-9])$`

Required: Yes

### Port

The port that the NameNode uses to listen to client requests.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 65536.

Required: Yes

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)



# IOPS

The IOPS peaks for an on-premises storage system resource. Each data point represents the 95th percentile peak value during a 1-hour interval.

## Contents

### Other

Peak IOPS unrelated to read and write operations.

Type: Double

Valid Range: Minimum value of 0.

Required: No

### Read

Peak IOPS related to read operations.

Type: Double

Valid Range: Minimum value of 0.

Required: No

### Total

Peak total IOPS on your on-premises storage system resource.

Type: Double

Valid Range: Minimum value of 0.

Required: No

### Write

Peak IOPS related to write operations.

Type: Double

Valid Range: Minimum value of 0.

Required: No

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

# Latency

The latency peaks for an on-premises storage system resource. Each data point represents the 95th percentile peak value during a 1-hour interval.

## Contents

### Other

Peak latency for operations unrelated to read and write operations.

Type: Double

Valid Range: Minimum value of 0.

Required: No

### Read

Peak latency for read operations.

Type: Double

Valid Range: Minimum value of 0.

Required: No

### Write

Peak latency for write operations.

Type: Double

Valid Range: Minimum value of 0.

Required: No

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)

- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

## LocationFilter

Narrow down the list of resources returned by `ListLocations`. For example, to see all your Amazon S3 locations, create a filter using `"Name": "LocationType", "Operator": "Equals",` and `"Values": "S3"`.

For more information, see [filtering resources](#).

### Contents

#### Name

The name of the filter being used. Each API call supports a list of filters that are available for it (for example, `LocationType` for `ListLocations`).

Type: String

Valid Values: `LocationUri` | `LocationType` | `CreationTime`

Required: Yes

#### Operator

The operator that is used to compare filter values (for example, `Equals` or `Contains`).

Type: String

Valid Values: `Equals` | `NotEquals` | `In` | `LessThanOrEqualTo` | `LessThan` | `GreaterThanOrEqual` | `GreaterThan` | `Contains` | `NotContains` | `BeginsWith`

Required: Yes

#### Values

The values that you want to filter for. For example, you might want to display only Amazon S3 locations.

Type: Array of strings

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `^[0-9a-zA-Z_\ \-\:\*\.\ \\/\?\-]*$`

Required: Yes



## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

## LocationListEntry

Represents a single entry in a list of locations. `LocationListEntry` returns an array that contains a list of locations when the [ListLocations](#) operation is called.

### Contents

#### LocationArn

The Amazon Resource Name (ARN) of the location. For Network File System (NFS) or Amazon EFS, the location is the export path. For Amazon S3, the location is the prefix path that you want to mount and use as the root of the location.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^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}$`

Required: No

#### LocationUri

Represents a list of URIs of a location. `LocationUri` returns an array that contains a list of locations when the [ListLocations](#) operation is called.

Format: `TYPE://GLOBAL_ID/SUBDIR`.

TYPE designates the type of location (for example, `nfs` or `s3`).

GLOBAL\_ID is the globally unique identifier of the resource that backs the location. An example for EFS is `us-east-2.fs-abcd1234`. An example for Amazon S3 is the bucket name, such as `myBucket`. An example for NFS is a valid IPv4 address or a hostname that is compliant with Domain Name Service (DNS).

SUBDIR is a valid file system path, delimited by forward slashes as is the \*nix convention. For NFS and Amazon EFS, it's the export path to mount the location. For Amazon S3, it's the prefix path that you mount to and treat as the root of the location.

Type: String

Length Constraints: Maximum length of 4360.

Pattern: `^(efs|nfs|s3|smb|hdfs|fsx[a-z0-9-]+)://[a-zA-Z0-9.\:/\-\-]+$`

Required: No

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

# ManifestConfig

Configures a manifest, which is a list of files or objects that you want AWS DataSync to transfer. For more information and configuration examples, see [Specifying what DataSync transfers by using a manifest](#).

## Contents

### Action

Specifies what DataSync uses the manifest for.

Type: String

Valid Values: TRANSFER

Required: No

### Format

Specifies the file format of your manifest. For more information, see [Creating a manifest](#).

Type: String

Valid Values: CSV

Required: No

### Source

Specifies the manifest that you want DataSync to use and where it's hosted.

#### Note

You must specify this parameter if you're configuring a new manifest on or after February 7, 2024.

If you don't, you'll get a 400 status code and `ValidationException` error stating that you're missing the IAM role for DataSync to access the S3 bucket where you're hosting your manifest. For more information, see [Providing DataSync access to your manifest](#).

Type: [SourceManifestConfig](#) object

Required: No

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

## MaxP95Performance

The performance data that DataSync Discovery collects about an on-premises storage system resource.

### Contents

#### **lopsOther**

Peak IOPS unrelated to read and write operations.

Type: Double

Valid Range: Minimum value of 0.

Required: No

#### **lopsRead**

Peak IOPS related to read operations.

Type: Double

Valid Range: Minimum value of 0.

Required: No

#### **lopsTotal**

Peak total IOPS on your on-premises storage system resource.

Type: Double

Valid Range: Minimum value of 0.

Required: No

#### **lopsWrite**

Peak IOPS related to write operations.

Type: Double

Valid Range: Minimum value of 0.

Required: No

### **LatencyOther**

Peak latency for operations unrelated to read and write operations.

Type: Double

Valid Range: Minimum value of 0.

Required: No

### **LatencyRead**

Peak latency for read operations.

Type: Double

Valid Range: Minimum value of 0.

Required: No

### **LatencyWrite**

Peak latency for write operations.

Type: Double

Valid Range: Minimum value of 0.

Required: No

### **ThroughputOther**

Peak throughput unrelated to read and write operations.

Type: Double

Valid Range: Minimum value of 0.

Required: No

### **ThroughputRead**

Peak throughput related to read operations.

Type: Double

Valid Range: Minimum value of 0.

Required: No

### **ThroughputTotal**

Peak total throughput on your on-premises storage system resource.

Type: Double

Valid Range: Minimum value of 0.

Required: No

### **ThroughputWrite**

Peak throughput related to write operations.

Type: Double

Valid Range: Minimum value of 0.

Required: No

## **See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)



# NetAppONTAPCluster

The information that DataSync Discovery collects about an on-premises storage system cluster.

## Contents

### CifsShareCount

The number of CIFS shares in the cluster.

Type: Long

Valid Range: Minimum value of 0.

Required: No

### ClusterBlockStorageLogicalUsed

The storage space that's being used in the cluster without accounting for compression or deduplication.

Type: Long

Valid Range: Minimum value of 0.

Required: No

### ClusterBlockStorageSize

The total storage space that's available in the cluster.

Type: Long

Valid Range: Minimum value of 0.

Required: No

### ClusterBlockStorageUsed

The storage space that's being used in a cluster.

Type: Long

Valid Range: Minimum value of 0.

Required: No

### **ClusterCloudStorageUsed**

The amount of space in the cluster that's in cloud storage (for example, if you're using data tiering).

Type: Long

Valid Range: Minimum value of 0.

Required: No

### **ClusterName**

The name of the cluster.

Type: String

Length Constraints: Maximum length of 1024.

Pattern: `^\.{0,1024}$`

Required: No

### **LunCount**

The number of LUNs (logical unit numbers) in the cluster.

Type: Long

Valid Range: Minimum value of 0.

Required: No

### **MaxP95Performance**

The performance data that DataSync Discovery collects about the cluster.

Type: [MaxP95Performance](#) object

Required: No

### **NfsExportedVolumes**

The number of NFS volumes in the cluster.

Type: Long

Valid Range: Minimum value of 0.

Required: No

## Recommendations

The AWS storage services that DataSync Discovery recommends for the cluster. For more information, see [Recommendations provided by DataSync Discovery](#).

Type: Array of [Recommendation](#) objects

Required: No

## RecommendationStatus

Indicates whether DataSync Discovery recommendations for the cluster are ready to view, incomplete, or can't be determined.

For more information, see [Recommendation statuses](#).

Type: String

Valid Values: NONE | IN\_PROGRESS | COMPLETED | FAILED

Required: No

## ResourceId

The universally unique identifier (UUID) of the cluster.

Type: String

Pattern: `[a-f0-9]{8}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{12}`

Required: No

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)

- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

# NetAppONTAPSVM

The information that DataSync Discovery collects about a storage virtual machine (SVM) in your on-premises storage system.

## Contents

### CifsShareCount

The number of CIFS shares in the SVM.

Type: Long

Valid Range: Minimum value of 0.

Required: No

### ClusterUuid

The universally unique identifier (UUID) of the cluster associated with the SVM.

Type: String

Pattern: `[a-f0-9]{8}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{12}`

Required: No

### EnabledProtocols

The data transfer protocols (such as NFS) configured for the SVM.

Type: Array of strings

Length Constraints: Maximum length of 1024.

Pattern: `^.{0,1024}$`

Required: No

### LunCount

The number of LUNs (logical unit numbers) in the SVM.

Type: Long

Valid Range: Minimum value of 0.

Required: No

### **MaxP95Performance**

The performance data that DataSync Discovery collects about the SVM.

Type: [MaxP95Performance](#) object

Required: No

### **NfsExportedVolumes**

The number of NFS volumes in the SVM.

Type: Long

Valid Range: Minimum value of 0.

Required: No

### **Recommendations**

The AWS storage services that DataSync Discovery recommends for the SVM. For more information, see [Recommendations provided by DataSync Discovery](#).

Type: Array of [Recommendation](#) objects

Required: No

### **RecommendationStatus**

Indicates whether DataSync Discovery recommendations for the SVM are ready to view, incomplete, or can't be determined.

For more information, see [Recommendation statuses](#).

Type: String

Valid Values: NONE | IN\_PROGRESS | COMPLETED | FAILED

Required: No

### **ResourceId**

The UUID of the SVM.

Type: String

Pattern: [a-f0-9]{8}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{12}

Required: No

### **SvmName**

The name of the SVM

Type: String

Length Constraints: Maximum length of 1024.

Pattern: ^.{0,1024}\$

Required: No

### **TotalCapacityProvisioned**

The total storage space that's available in the SVM.

Type: Long

Valid Range: Minimum value of 0.

Required: No

### **TotalCapacityUsed**

The storage space that's being used in the SVM.

Type: Long

Valid Range: Minimum value of 0.

Required: No

### **TotalLogicalCapacityUsed**

The storage space that's being used in the SVM without accounting for compression or deduplication.

Type: Long

Valid Range: Minimum value of 0.

Required: No

### **TotalSnapshotCapacityUsed**

The amount of storage in the SVM that's being used for snapshots.

Type: Long

Valid Range: Minimum value of 0.

Required: No

### **See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)



# NetAppONTAPVolume

The information that DataSync Discovery collects about a volume in your on-premises storage system.

## Contents

### CapacityProvisioned

The total storage space that's available in the volume.

Type: Long

Valid Range: Minimum value of 0.

Required: No

### CapacityUsed

The storage space that's being used in the volume.

Type: Long

Valid Range: Minimum value of 0.

Required: No

### CifsShareCount

The number of CIFS shares in the volume.

Type: Long

Valid Range: Minimum value of 0.

Required: No

### LogicalCapacityUsed

The storage space that's being used in the volume without accounting for compression or deduplication.

Type: Long

Valid Range: Minimum value of 0.

Required: No

### **LunCount**

The number of LUNs (logical unit numbers) in the volume.

Type: Long

Valid Range: Minimum value of 0.

Required: No

### **MaxP95Performance**

The performance data that DataSync Discovery collects about the volume.

Type: [MaxP95Performance](#) object

Required: No

### **NfsExported**

The number of NFS volumes in the volume.

Type: Boolean

Required: No

### **Recommendations**

The AWS storage services that DataSync Discovery recommends for the volume. For more information, see [Recommendations provided by DataSync Discovery](#).

Type: Array of [Recommendation](#) objects

Required: No

### **RecommendationStatus**

Indicates whether DataSync Discovery recommendations for the volume are ready to view, incomplete, or can't be determined.

For more information, see [Recommendation statuses](#).

Type: String

Valid Values: NONE | IN\_PROGRESS | COMPLETED | FAILED

Required: No

### **ResourceId**

The universally unique identifier (UUID) of the volume.

Type: String

Pattern: `[a-f0-9]{8}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{12}`

Required: No

### **SecurityStyle**

The volume's security style (such as Unix or NTFS).

Type: String

Length Constraints: Maximum length of 1024.

Pattern: `^.{0,1024}$`

Required: No

### **SnapshotCapacityUsed**

The amount of storage in the volume that's being used for snapshots.

Type: Long

Valid Range: Minimum value of 0.

Required: No

### **SvmName**

The name of the SVM associated with the volume.

Type: String

Length Constraints: Maximum length of 1024.

Pattern: `^.{0,1024}$`

Required: No

## SvmUuid

The UUID of the storage virtual machine (SVM) associated with the volume.

Type: String

Pattern: `[a-f0-9]{8}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{12}`

Required: No

## VolumeName

The name of the volume.

Type: String

Length Constraints: Maximum length of 1024.

Pattern: `^.{0,1024}$`

Required: No

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

# NfsMountOptions

Specifies how DataSync can access a location using the NFS protocol.

## Contents

### Version

Specifies the NFS version that you want DataSync to use when mounting your NFS share. If the server refuses to use the version specified, the task fails.

You can specify the following options:

- **AUTOMATIC** (default): DataSync chooses NFS version 4.1.
- **NFS3**: Stateless protocol version that allows for asynchronous writes on the server.
- **NFSv4\_0**: Stateful, firewall-friendly protocol version that supports delegations and pseudo file systems.
- **NFSv4\_1**: Stateful protocol version that supports sessions, directory delegations, and parallel data processing. NFS version 4.1 also includes all features available in version 4.0.

#### Note

DataSync currently only supports NFS version 3 with Amazon FSx for NetApp ONTAP locations.

Type: String

Valid Values: AUTOMATIC | NFS3 | NFS4\_0 | NFS4\_1

Required: No

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)

- [AWS SDK for Ruby V3](#)

# OnPremConfig

The AWS DataSync agents that are connecting to a Network File System (NFS) location.

## Contents

### AgentArns

The Amazon Resource Names (ARNs) of the agents connecting to a transfer location.

Type: Array of strings

Array Members: Minimum number of 1 item. Maximum number of 4 items.

Length Constraints: Maximum length of 128.

Pattern: `^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}$`

Required: Yes

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

## Options

Indicates how your transfer task is configured. These options include how DataSync handles files, objects, and their associated metadata during your transfer. You also can specify how to verify data integrity, set bandwidth limits for your task, among other options.

Each option has a default value. Unless you need to, you don't have to configure any option before calling [StartTaskExecution](#).

You also can override your task options for each task execution. For example, you might want to adjust the `LogLevel` for an individual execution.

## Contents

### Atime

Specifies whether to preserve metadata indicating the last time a file was read or written to.

#### Note

The behavior of `Atime` isn't fully standard across platforms, so DataSync can only do this on a best-effort basis.

- `BEST Effort` (default) - DataSync attempts to preserve the original `Atime` attribute on all source files (that is, the version before the `PREPARING` phase of the task execution). This option is recommended.
- `NONE` - Ignores `Atime`.

#### Note

If `Atime` is set to `BEST Effort`, `Mtime` must be set to `PRESERVE`.  
If `Atime` is set to `NONE`, `Mtime` must also be `NONE`.

Type: String

Valid Values: `NONE` | `BEST Effort`

Required: No



## BytesPerSecond

Limits the bandwidth used by a DataSync task. For example, if you want DataSync to use a maximum of 1 MB, set this value to 1048576 (=1024\*1024).

Type: Long

Valid Range: Minimum value of -1.

Required: No

## Gid

Specifies the POSIX group ID (GID) of the file's owners.

- INT\_VALUE (default) - Preserves the integer value of user ID (UID) and GID, which is recommended.
- NONE - Ignores UID and GID.

For more information, see [Metadata copied by DataSync](#).

Type: String

Valid Values: NONE | INT\_VALUE | NAME | BOTH

Required: No

## LogLevel

Specifies the type of logs that DataSync publishes to a Amazon CloudWatch Logs log group. To specify the log group, see [CloudWatchLogGroupArn](#).

- BASIC - Publishes logs with only basic information (such as transfer errors).
- TRANSFER - Publishes logs for all files or objects that your DataSync task transfers and performs data-integrity checks on.
- OFF - No logs are published.

Type: String

Valid Values: OFF | BASIC | TRANSFER

Required: No

## Mtime

Specifies whether to preserve metadata indicating the last time that a file was written to before the PREPARING phase of your task execution. This option is required when you need to run the a task more than once.

- PRESERVE (default) - Preserves original Mtime, which is recommended.
- NONE - Ignores Mtime.

### Note

If Mtime is set to PRESERVE, Atime must be set to BEST\_EFFORT.  
If Mtime is set to NONE, Atime must also be set to NONE.

Type: String

Valid Values: NONE | PRESERVE

Required: No

## ObjectTags

Specifies whether you want DataSync to PRESERVE object tags (default behavior) when transferring between object storage systems. If you want your DataSync task to ignore object tags, specify the NONE value.

Type: String

Valid Values: PRESERVE | NONE

Required: No

## OverwriteMode

Specifies whether DataSync should modify or preserve data at the destination location.

- ALWAYS (default) - DataSync modifies data in the destination location when source data (including metadata) has changed.

If DataSync overwrites objects, you might incur additional charges for certain Amazon S3 storage classes (for example, for retrieval or early deletion). For more information, see [Storage class considerations with Amazon S3 transfers](#).

- NEVER - DataSync doesn't overwrite data in the destination location even if the source data has changed. You can use this option to protect against overwriting changes made to files or objects in the destination.

Type: String

Valid Values: ALWAYS | NEVER

Required: No

### PosixPermissions

Specifies which users or groups can access a file for a specific purpose such as reading, writing, or execution of the file.

For more information, see [Metadata copied by DataSync](#).

- PRESERVE (default) - Preserves POSIX-style permissions, which is recommended.
- NONE - Ignores POSIX-style permissions.

#### Note

DataSync can preserve extant permissions of a source location.

Type: String

Valid Values: NONE | PRESERVE

Required: No

### PreserveDeletedFiles

Specifies whether files in the destination location that don't exist in the source should be preserved. This option can affect your Amazon S3 storage cost. If your task deletes objects, you might incur minimum storage duration charges for certain storage classes. For detailed information, see [Considerations when working with Amazon S3 storage classes in DataSync](#).

- PRESERVE (default) - Ignores such destination files, which is recommended.
- REMOVE - Deletes destination files that aren't present in the source.

**Note**

If you set this parameter to REMOVE, you can't set TransferMode to ALL. When you transfer all data, DataSync doesn't scan your destination location and doesn't know what to delete.

Type: String

Valid Values: PRESERVE | REMOVE

Required: No

**PreserveDevices**

Specifies whether DataSync should preserve the metadata of block and character devices in the source location and recreate the files with that device name and metadata on the destination. DataSync copies only the name and metadata of such devices.

**Note**

DataSync can't copy the actual contents of these devices because they're nonterminal and don't return an end-of-file (EOF) marker.

- NONE (default) - Ignores special devices (recommended).
- PRESERVE - Preserves character and block device metadata. This option currently isn't supported for Amazon EFS.

Type: String

Valid Values: NONE | PRESERVE

Required: No

**SecurityDescriptorCopyFlags**

Specifies which components of the SMB security descriptor are copied from source to destination objects.

This value is only used for transfers between SMB and Amazon FSx for Windows File Server locations or between two FSx for Windows File Server locations. For more information, see [how DataSync handles metadata](#).

- OWNER\_DACL (default) - For each copied object, DataSync copies the following metadata:
  - The object owner.
  - NTFS discretionary access control lists (DACLS), which determine whether to grant access to an object.

DataSync won't copy NTFS system access control lists (SACLs) with this option.

- OWNER\_DACL\_SACL - For each copied object, DataSync copies the following metadata:
  - The object owner.
  - NTFS discretionary access control lists (DACLS), which determine whether to grant access to an object.
  - SACLs, which are used by administrators to log attempts to access a secured object.

Copying SACLs requires granting additional permissions to the Windows user that DataSync uses to access your SMB location. For information about choosing a user with the right permissions, see required permissions for [SMB](#), [FSx for Windows File Server](#), or [FSx for ONTAP](#) (depending on the type of location in your transfer).

- NONE - None of the SMB security descriptor components are copied. Destination objects are owned by the user that was provided for accessing the destination location. DACLS and SACLs are set based on the destination server's configuration.

Type: String

Valid Values: NONE | OWNER\_DACL | OWNER\_DACL\_SACL

Required: No

## TaskQueueing

Specifies whether your transfer tasks should be put into a queue during certain scenarios when [running multiple tasks](#). This is ENABLED by default.

Type: String

Valid Values: ENABLED | DISABLED

Required: No

## TransferMode

Determines whether DataSync transfers only the data and metadata that differ between the source and the destination location or transfers all the content from the source (without comparing what's in the destination).

- **CHANGED** (default) - DataSync copies only data or metadata that is new or different content from the source location to the destination location.
- **ALL** - DataSync copies everything in the source to the destination without comparing differences between the locations.

Type: String

Valid Values: CHANGED | ALL

Required: No

## Uid

Specifies the POSIX user ID (UID) of the file's owner.

- **INT\_VALUE** (default) - Preserves the integer value of UID and group ID (GID), which is recommended.
- **NONE** - Ignores UID and GID.

For more information, see [Metadata copied by DataSync](#).

Type: String

Valid Values: NONE | INT\_VALUE | NAME | BOTH

Required: No

## VerifyMode

Specifies how and when DataSync checks the integrity of your data during a transfer.

- **ONLY\_FILES\_TRANSFERRED** (recommended) - DataSync calculates the checksum of transferred files and metadata at the source location. At the end of the transfer, DataSync then compares this checksum to the checksum calculated on those files at the destination.

We recommend this option when transferring to S3 Glacier Flexible Retrieval or S3 Glacier Deep Archive storage classes. For more information, see [Storage class considerations with Amazon S3 locations](#).

- `POINT_IN_TIME_CONSISTENT` (default) - At the end of the transfer, DataSync scans the entire source and destination to verify that both locations are fully synchronized.

If you use a [manifest](#), DataSync only scans and verifies what's listed in the manifest.

You can't use this option when transferring to S3 Glacier Flexible Retrieval or S3 Glacier Deep Archive storage classes. For more information, see [Storage class considerations with Amazon S3 locations](#).

- `NONE` - DataSync doesn't run additional verification at the end of the transfer. All data transmissions are still integrity-checked with checksum verification during the transfer.

Type: String

Valid Values: `POINT_IN_TIME_CONSISTENT` | `ONLY_FILES_TRANSFERRED` | `NONE`

Required: No

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

## P95Metrics

The types of performance data that DataSync Discovery collects about an on-premises storage system resource.

### Contents

#### IOPS

The IOPS peaks for an on-premises storage system resource. Each data point represents the 95th percentile peak value during a 1-hour interval.

Type: [IOPS](#) object

Required: No

#### Latency

The latency peaks for an on-premises storage system resource. Each data point represents the 95th percentile peak value during a 1-hour interval.

Type: [Latency](#) object

Required: No

#### Throughput

The throughput peaks for an on-premises storage system resource. Each data point represents the 95th percentile peak value during a 1-hour interval.

Type: [Throughput](#) object

Required: No

### See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)





## Platform

The platform-related details about the AWS DataSync agent, such as the version number.

### Contents

#### Version

The version of the DataSync agent.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 256.

Pattern: `^[a-zA-Z0-9\s+=. _:@/-]+$`

Required: No

### See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

## PrivateLinkConfig

Specifies how your AWS DataSync agent connects to AWS using a [virtual private cloud \(VPC\) service endpoint](#). An agent that uses a VPC endpoint isn't accessible over the public internet.

### Contents

#### PrivateLinkEndpoint

Specifies the VPC endpoint provided by [AWS PrivateLink](#) that your agent connects to.

Type: String

Length Constraints: Minimum length of 7. Maximum length of 15.

Pattern: `\A(25[0-5]|2[0-4]\d|[0-1]?\d?\d)(\.(25[0-5]|2[0-4]\d|[0-1]?\d?\d)){3}\z`

Required: No

#### SecurityGroupArns

Specifies the Amazon Resource Names (ARN) of the security group that provides DataSync access to your VPC endpoint. You can only specify one ARN.

Type: Array of strings

Array Members: Fixed number of 1 item.

Length Constraints: Maximum length of 128.

Pattern: `^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]+$`

Required: No

#### SubnetArns

Specifies the ARN of the subnet where your VPC endpoint is located. You can only specify one ARN.

Type: Array of strings

Array Members: Fixed number of 1 item.

Length Constraints: Maximum length of 128.

Pattern: `^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):ec2:[a-z\-\0-9]*:[0-9]{12}:subnet/.*$`

Required: No

### **VpcEndpointId**

Specifies the ID of the VPC endpoint that your agent connects to.

Type: String

Pattern: `^vpce-[0-9a-f]{17}$`

Required: No

### **See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

# QopConfiguration

The Quality of Protection (QOP) configuration specifies the Remote Procedure Call (RPC) and data transfer privacy settings configured on the Hadoop Distributed File System (HDFS) cluster.

## Contents

### DataTransferProtection

The data transfer protection setting configured on the HDFS cluster. This setting corresponds to your `dfs.data.transfer.protection` setting in the `hdfs-site.xml` file on your Hadoop cluster.

Type: String

Valid Values: DISABLED | AUTHENTICATION | INTEGRITY | PRIVACY

Required: No

### RpcProtection

The RPC protection setting configured on the HDFS cluster. This setting corresponds to your `hadoop.rpc.protection` setting in your `core-site.xml` file on your Hadoop cluster.

Type: String

Valid Values: DISABLED | AUTHENTICATION | INTEGRITY | PRIVACY

Required: No

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

## Recommendation

The details about an AWS storage service that DataSync Discovery recommends for a resource in your on-premises storage system.

For more information, see [Recommendations provided by DataSync Discovery](#).

### Contents

#### EstimatedMonthlyStorageCost

The estimated monthly cost of the recommended AWS storage service.

Type: String

Length Constraints: Maximum length of 1024.

Pattern:  $^{\{0,1024\}}\$$

Required: No

#### StorageConfiguration

Information about how you can set up a recommended AWS storage service.

Type: String to string map

Key Length Constraints: Maximum length of 1024.

Key Pattern:  $^{\{0,1024\}}\$$

Value Length Constraints: Maximum length of 1024.

Value Pattern:  $^{\{0,1024\}}\$$

Required: No

#### StorageType

A recommended AWS storage service that you can migrate data to based on information that DataSync Discovery collects about your on-premises storage system.

Type: String

Length Constraints: Maximum length of 1024.

Pattern: `^\.{0,1024}$`

Required: No

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

# ReportDestination

Specifies where DataSync uploads your [task report](#).

## Contents

### S3

Specifies the Amazon S3 bucket where DataSync uploads your task report.

Type: [ReportDestinationS3](#) object

Required: No

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)



# ReportDestinationS3

Specifies the Amazon S3 bucket where DataSync uploads your [task report](#).

## Contents

### BucketAccessRoleArn

Specifies the Amazon Resource Name (ARN) of the IAM policy that allows DataSync to upload a task report to your S3 bucket. For more information, see [Allowing DataSync to upload a task report to an Amazon S3 bucket](#).

Type: String

Length Constraints: Maximum length of 2048.

Pattern: `^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):iam::[0-9]{12}:role/.*$`

Required: Yes

### S3BucketArn

Specifies the ARN of the S3 bucket where DataSync uploads your report.

Type: String

Length Constraints: Maximum length of 156.

Pattern: `^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):(s3|s3-outposts):[a-z\-\0-9]*:[0-9]*:.*$`

Required: Yes

### Subdirectory

Specifies a bucket prefix for your report.

Type: String

Length Constraints: Maximum length of 4096.

Pattern: `^[a-zA-Z0-9_\-\+\.\^(\)\p{Zs}]*$`

Required: No

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

# ReportOverride

Specifies the level of detail for a particular aspect of your DataSync [task report](#).

## Contents

### ReportLevel

Specifies whether your task report includes errors only or successes and errors.

For example, your report might mostly include only what didn't go well in your transfer (ERRORS\_ONLY). At the same time, you want to verify that your [task filter](#) is working correctly. In this situation, you can get a list of what files DataSync successfully skipped and if something transferred that you didn't to transfer (SUCCESSSES\_AND\_ERRORS).

Type: String

Valid Values: ERRORS\_ONLY | SUCCESSSES\_AND\_ERRORS

Required: No

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

# ReportOverrides

The level of detail included in each aspect of your DataSync [task report](#).

## Contents

### Deleted

Specifies the level of reporting for the files, objects, and directories that DataSync attempted to delete in your destination location. This only applies if you [configure your task](#) to delete data in the destination that isn't in the source.

Type: [ReportOverride](#) object

Required: No

### Skipped

Specifies the level of reporting for the files, objects, and directories that DataSync attempted to skip during your transfer.

Type: [ReportOverride](#) object

Required: No

### Transferred

Specifies the level of reporting for the files, objects, and directories that DataSync attempted to transfer.

Type: [ReportOverride](#) object

Required: No

### Verified

Specifies the level of reporting for the files, objects, and directories that DataSync attempted to verify at the end of your transfer.

Type: [ReportOverride](#) object

Required: No

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

# ReportResult

Indicates whether DataSync created a complete [task report](#) for your transfer.

## Contents

### ErrorCode

Indicates the code associated with the error if DataSync can't create a complete report.

Type: String

Required: No

### ErrorDetail

Provides details about issues creating a report.

Type: String

Required: No

### Status

Indicates whether DataSync is still working on your report, created a report, or can't create a complete report.

Type: String

Valid Values: PENDING | SUCCESS | ERROR

Required: No

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

## ResourceDetails

Information provided by DataSync Discovery about the resources in your on-premises storage system.

### Contents

#### NetAppONTAPClusters

The information that DataSync Discovery collects about the cluster in your on-premises storage system.

Type: Array of [NetAppONTAPCluster](#) objects

Required: No

#### NetAppONTAPSVMS

The information that DataSync Discovery collects about storage virtual machines (SVMs) in your on-premises storage system.

Type: Array of [NetAppONTAP SVM](#) objects

Required: No

#### NetAppONTAPVolumes

The information that DataSync Discovery collects about volumes in your on-premises storage system.

Type: Array of [NetAppONTAPVolume](#) objects

Required: No

### See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)





## ResourceMetrics

Information, including performance data and capacity usage, provided by DataSync Discovery about a resource in your on-premises storage system.

### Contents

#### Capacity

The storage capacity of the on-premises storage system resource.

Type: [Capacity](#) object

Required: No

#### P95Metrics

The types of performance data that DataSync Discovery collects about the on-premises storage system resource.

Type: [P95Metrics](#) object

Required: No

#### ResourceId

The universally unique identifier (UUID) of the on-premises storage system resource.

Type: String

Pattern: `[a-f0-9]{8}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{12}`

Required: No

#### ResourceType

The type of on-premises storage system resource.

Type: String

Valid Values: SVM | VOLUME | CLUSTER

Required: No

## Timestamp

The time when DataSync Discovery collected this information from the resource.

Type: Timestamp

Required: No

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

## S3Config

Specifies the Amazon Resource Name (ARN) of the AWS Identity and Access Management (IAM) role that DataSync uses to access your S3 bucket.

For more information, see [Accessing S3 buckets](#).

### Contents

#### BucketAccessRoleArn

Specifies the ARN of the IAM role that DataSync uses to access your S3 bucket.

Type: String

Length Constraints: Maximum length of 2048.

Pattern: `^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):iam::[0-9]{12}:role/.*$`

Required: Yes

### See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

## S3ManifestConfig

Specifies the S3 bucket where you're hosting the manifest that you want AWS DataSync to use. For more information and configuration examples, see [Specifying what DataSync transfers by using a manifest](#).

### Contents

#### BucketAccessRoleArn

Specifies the AWS Identity and Access Management (IAM) role that allows DataSync to access your manifest. For more information, see [Providing DataSync access to your manifest](#).

Type: String

Length Constraints: Maximum length of 2048.

Pattern: `^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):iam::[0-9]{12}:role/.*$`

Required: Yes

#### ManifestObjectPath

Specifies the Amazon S3 object key of your manifest. This can include a prefix (for example, `prefix/my-manifest.csv`).

Type: String

Length Constraints: Maximum length of 4096.

Pattern: `^[a-zA-Z0-9_-\+\.\^(\)\p{Zs}]*$`

Required: Yes

#### S3BucketArn

Specifies the Amazon Resource Name (ARN) of the S3 bucket where you're hosting your manifest.

Type: String

Length Constraints: Maximum length of 156.

Pattern: `^arn:(aws|aws-cn|aws-us-gov|aws-iso|aws-iso-b):(s3|s3-outposts):[a-z\-\0-9]*:[0-9]*:.*$`

Required: Yes

### **ManifestObjectVersionId**

Specifies the object version ID of the manifest that you want DataSync to use. If you don't set this, DataSync uses the latest version of the object.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 100.

Pattern: `^\.+ $`

Required: No

### **See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

# SmbMountOptions

Specifies the version of the Server Message Block (SMB) protocol that AWS DataSync uses to access an SMB file server.

## Contents

### Version

By default, DataSync automatically chooses an SMB protocol version based on negotiation with your SMB file server. You also can configure DataSync to use a specific SMB version, but we recommend doing this only if DataSync has trouble negotiating with the SMB file server automatically.

These are the following options for configuring the SMB version:

- **AUTOMATIC** (default): DataSync and the SMB file server negotiate the highest version of SMB that they mutually support between 2.1 and 3.1.1.

This is the recommended option. If you instead choose a specific version that your file server doesn't support, you may get an `Operation Not Supported` error.

- **SMB3**: Restricts the protocol negotiation to only SMB version 3.0.2.
- **SMB2**: Restricts the protocol negotiation to only SMB version 2.1.
- **SMB2\_0**: Restricts the protocol negotiation to only SMB version 2.0.
- **SMB1**: Restricts the protocol negotiation to only SMB version 1.0.

#### Note

The SMB1 option isn't available when [creating an Amazon FSx for NetApp ONTAP location](#).

Type: String

Valid Values: AUTOMATIC | SMB2 | SMB3 | SMB1 | SMB2\_0

Required: No

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

# SourceManifestConfig

Specifies the manifest that you want AWS DataSync to use and where it's hosted. For more information and configuration examples, see [Specifying what DataSync transfers by using a manifest](#).

## Contents

### S3

Specifies the S3 bucket where you're hosting your manifest.

Type: [S3ManifestConfig](#) object

Required: Yes

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)



## StorageSystemListEntry

Information that identifies an on-premises storage system that you're using with DataSync Discovery.

### Contents

#### Name

The name of an on-premises storage system that you added to DataSync Discovery.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 256.

Pattern: `^\p{L}\p{M}\p{N}\s+=. _:@\|-]+$`

Required: No

#### StorageSystemArn

The Amazon Resource Names (ARN) of an on-premises storage system that you added to DataSync Discovery.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^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}$`

Required: No

### See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)

- [AWS SDK for Ruby V3](#)

## TagListEntry

A key-value pair representing a single tag that's been applied to an AWS resource.

### Contents

#### Key

The key for an AWS resource tag.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 256.

Pattern: `^[a-zA-Z0-9\s+=. _:/-]+$`

Required: Yes

#### Value

The value for an AWS resource tag.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 256.

Pattern: `^[a-zA-Z0-9\s+=. _:@/-]+$`

Required: No

### See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

## TaskExecutionListEntry

Represents a single entry in a list of AWS DataSync task executions that's returned with the [ListTaskExecutions](#) operation.

### Contents

#### Status

The status of a task execution. For more information, see [Task execution statuses](#).

Type: String

Valid Values: QUEUED | CANCELLING | LAUNCHING | PREPARING | TRANSFERRING | VERIFYING | SUCCESS | ERROR

Required: No

#### TaskExecutionArn

The Amazon Resource Name (ARN) of a task execution.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^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}$`

Required: No

### See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

## TaskExecutionResultDetail

Describes the detailed result of a `TaskExecution` operation. This result includes the time in milliseconds spent in each phase, the status of the task execution, and the errors encountered.

### Contents

#### ErrorCode

Errors that AWS DataSync encountered during execution of the task. You can use this error code to help troubleshoot issues.

Type: String

Required: No

#### ErrorDetail

Detailed description of an error that was encountered during the task execution. You can use this information to help troubleshoot issues.

Type: String

Required: No

#### PrepareDuration

The total time in milliseconds that AWS DataSync spent in the `PREPARING` phase.

Type: Long

Valid Range: Minimum value of 0.

Required: No

#### PrepareStatus

The status of the `PREPARING` phase.

Type: String

Valid Values: `PENDING` | `SUCCESS` | `ERROR`

Required: No

**TotalDuration**

The total time in milliseconds that AWS DataSync took to transfer the file from the source to the destination location.

Type: Long

Valid Range: Minimum value of 0.

Required: No

**TransferDuration**

The total time in milliseconds that AWS DataSync spent in the TRANSFERRING phase.

Type: Long

Valid Range: Minimum value of 0.

Required: No

**TransferStatus**

The status of the TRANSFERRING phase.

Type: String

Valid Values: PENDING | SUCCESS | ERROR

Required: No

**VerifyDuration**

The total time in milliseconds that AWS DataSync spent in the VERIFYING phase.

Type: Long

Valid Range: Minimum value of 0.

Required: No

**VerifyStatus**

The status of the VERIFYING phase.

Type: String

Valid Values: PENDING | SUCCESS | ERROR

Required: No

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

## TaskFilter

You can use API filters to narrow down the list of resources returned by `ListTasks`. For example, to retrieve all tasks on a source location, you can use `ListTasks` with filter name `LocationId` and `Operator` `Equals` with the ARN for the location.

For more information, see [filtering DataSync resources](#).

### Contents

#### Name

The name of the filter being used. Each API call supports a list of filters that are available for it. For example, `LocationId` for `ListTasks`.

Type: String

Valid Values: `LocationId` | `CreationTime`

Required: Yes

#### Operator

The operator that is used to compare filter values (for example, `Equals` or `Contains`).

Type: String

Valid Values: `Equals` | `NotEquals` | `In` | `LessThanOrEqualTo` | `LessThan` | `GreaterThanOrEqual` | `GreaterThan` | `Contains` | `NotContains` | `BeginsWith`

Required: Yes

#### Values

The values that you want to filter for. For example, you might want to display only tasks for a specific destination location.

Type: Array of strings

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `^[0-9a-zA-Z_\ \-\:\*\.\ \\/\?\-]*$`

Required: Yes



## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

# TaskListEntry

Represents a single entry in a list of tasks. `TaskListEntry` returns an array that contains a list of tasks when the [ListTasks](#) operation is called. A task includes the source and destination file systems to sync and the options to use for the tasks.

## Contents

### Name

The name of the task.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 256.

Pattern: `^[a-zA-Z0-9\s+=. _:@/-]+$`

Required: No

### Status

The status of the task.

Type: String

Valid Values: AVAILABLE | CREATING | QUEUED | RUNNING | UNAVAILABLE

Required: No

### TaskArn

The Amazon Resource Name (ARN) of the task.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `^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}$`

Required: No

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

## TaskReportConfig

Specifies how you want to configure a task report, which provides detailed information about for your AWS DataSync transfer.

For more information, see [Task reports](#).

### Contents

#### Destination

Specifies the Amazon S3 bucket where DataSync uploads your task report. For more information, see [Task reports](#).

Type: [ReportDestination](#) object

Required: No

#### ObjectVersionIds

Specifies whether your task report includes the new version of each object transferred into an S3 bucket. This only applies if you [enable versioning on your bucket](#). Keep in mind that setting this to INCLUDE can increase the duration of your task execution.

Type: String

Valid Values: INCLUDE | NONE

Required: No

#### OutputType

Specifies the type of task report that you want:

- **SUMMARY\_ONLY**: Provides necessary details about your task, including the number of files, objects, and directories transferred and transfer duration.
- **STANDARD**: Provides complete details about your task, including a full list of files, objects, and directories that were transferred, skipped, verified, and more.

Type: String

Valid Values: SUMMARY\_ONLY | STANDARD

Required: No

## Overrides

Customizes the reporting level for aspects of your task report. For example, your report might generally only include errors, but you could specify that you want a list of successes and errors just for the files that DataSync attempted to delete in your destination location.

Type: [ReportOverrides](#) object

Required: No

## ReportLevel

Specifies whether you want your task report to include only what went wrong with your transfer or a list of what succeeded and didn't.

- **ERRORS\_ONLY**: A report shows what DataSync was unable to transfer, skip, verify, and delete.
- **SUCCESSES\_AND\_ERRORS**: A report shows what DataSync was able and unable to transfer, skip, verify, and delete.

Type: String

Valid Values: **ERRORS\_ONLY** | **SUCCESSES\_AND\_ERRORS**

Required: No

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

# TaskSchedule

Configures your AWS DataSync task to run on a [schedule](#) (at a minimum interval of 1 hour).

## Contents

### ScheduleExpression

Specifies your task schedule by using a cron expression in UTC time. For information about cron expression syntax, see the [Amazon EventBridge User Guide](#).

Type: String

Length Constraints: Maximum length of 256.

Pattern: `^[a-zA-Z0-9\ \_ \* \? \, \ | \ ^ \ - \ \# \ $ \ ( \) \ + ] * $`

Required: Yes

### Status

Specifies whether to enable or disable your task schedule. Your schedule is enabled by default, but there can be situations where you need to disable it. For example, you might need to pause a recurring transfer to fix an issue with your task or perform maintenance on your storage system.

DataSync might disable your schedule automatically if your task fails repeatedly with the same error. For more information, see [TaskScheduleDetails](#).

Type: String

Valid Values: ENABLED | DISABLED

Required: No

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)

- [AWS SDK for Ruby V3](#)

## TaskScheduleDetails

Provides information about your AWS DataSync [task schedule](#).

### Contents

#### DisabledBy

Indicates how your task schedule was disabled.

- **USER** - Your schedule was manually disabled by using the [UpdateTask](#) operation or DataSync console.
- **SERVICE** - Your schedule was automatically disabled by DataSync because the task failed repeatedly with the same error.

Type: String

Valid Values: USER | SERVICE

Required: No

#### DisabledReason

Provides a reason if the task schedule is disabled.

If your schedule is disabled by **USER**, you see a `Manually disabled by user.` message.

If your schedule is disabled by **SERVICE**, you see an error message to help you understand why the task keeps failing. For information on resolving DataSync errors, see [Troubleshooting issues with DataSync transfers](#).

Type: String

Length Constraints: Maximum length of 8192.

Pattern: `^[\\w\\s.,'?!:;\\/=|<>()-]*$`

Required: No

#### StatusUpdateTime

Indicates the last time the status of your task schedule changed. For example, if DataSync automatically disables your schedule because of a repeated error, you can see when the schedule was disabled.



Type: Timestamp

Required: No

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

# Throughput

The throughput peaks for an on-premises storage system volume. Each data point represents the 95th percentile peak value during a 1-hour interval.

## Contents

### Other

Peak throughput unrelated to read and write operations.

Type: Double

Valid Range: Minimum value of 0.

Required: No

### Read

Peak throughput related to read operations.

Type: Double

Valid Range: Minimum value of 0.

Required: No

### Total

Peak total throughput on your on-premises storage system resource.

Type: Double

Valid Range: Minimum value of 0.

Required: No

### Write

Peak throughput related to write operations.

Type: Double

Valid Range: Minimum value of 0.

Required: No

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

## Common Errors

This section lists the errors common to the API actions of all AWS services. For errors specific to an API action for this service, see the topic for that API action.

### AccessDeniedException

You do not have sufficient access to perform this action.

HTTP Status Code: 400

### IncompleteSignature

The request signature does not conform to AWS standards.

HTTP Status Code: 400

### InternalFailure

The request processing has failed because of an unknown error, exception or failure.

HTTP Status Code: 500

### InvalidAction

The action or operation requested is invalid. Verify that the action is typed correctly.

HTTP Status Code: 400

### InvalidClientTokenId

The X.509 certificate or AWS access key ID provided does not exist in our records.

HTTP Status Code: 403

### **NotAuthorized**

You do not have permission to perform this action.

HTTP Status Code: 400

### **OptInRequired**

The AWS access key ID needs a subscription for the service.

HTTP Status Code: 403

### **RequestExpired**

The request reached the service more than 15 minutes after the date stamp on the request or more than 15 minutes after the request expiration date (such as for pre-signed URLs), or the date stamp on the request is more than 15 minutes in the future.

HTTP Status Code: 400

### **ServiceUnavailable**

The request has failed due to a temporary failure of the server.

HTTP Status Code: 503

### **ThrottlingException**

The request was denied due to request throttling.

HTTP Status Code: 400

### **ValidationError**

The input fails to satisfy the constraints specified by an AWS service.

HTTP Status Code: 400

## **Common Parameters**

The following list contains the parameters that all actions use for signing Signature Version 4 requests with a query string. Any action-specific parameters are listed in the topic for that action.

For more information about Signature Version 4, see [Signing AWS API requests](#) in the *IAM User Guide*.

### Action

The action to be performed.

Type: string

Required: Yes

### Version

The API version that the request is written for, expressed in the format YYYY-MM-DD.

Type: string

Required: Yes

### X-Amz-Algorithm

The hash algorithm that you used to create the request signature.

Condition: Specify this parameter when you include authentication information in a query string instead of in the HTTP authorization header.

Type: string

Valid Values: AWS4-HMAC-SHA256

Required: Conditional

### X-Amz-Credential

The credential scope value, which is a string that includes your access key, the date, the region you are targeting, the service you are requesting, and a termination string ("aws4\_request"). The value is expressed in the following format: *access\_key/YYYYMMDD/region/service/aws4\_request*.

For more information, see [Create a signed AWS API request](#) in the *IAM User Guide*.

Condition: Specify this parameter when you include authentication information in a query string instead of in the HTTP authorization header.

Type: string

Required: Conditional

### **X-Amz-Date**

The date that is used to create the signature. The format must be ISO 8601 basic format (YYYYMMDD'T'HHMMSS'Z'). For example, the following date time is a valid X-Amz-Date value: 20120325T120000Z.

Condition: X-Amz-Date is optional for all requests; it can be used to override the date used for signing requests. If the Date header is specified in the ISO 8601 basic format, X-Amz-Date is not required. When X-Amz-Date is used, it always overrides the value of the Date header. For more information, see [Elements of an AWS API request signature](#) in the *IAM User Guide*.

Type: string

Required: Conditional

### **X-Amz-Security-Token**

The temporary security token that was obtained through a call to AWS Security Token Service (AWS STS). For a list of services that support temporary security credentials from AWS STS, see [AWS services that work with IAM](#) in the *IAM User Guide*.

Condition: If you're using temporary security credentials from AWS STS, you must include the security token.

Type: string

Required: Conditional

### **X-Amz-Signature**

Specifies the hex-encoded signature that was calculated from the string to sign and the derived signing key.

Condition: Specify this parameter when you include authentication information in a query string instead of in the HTTP authorization header.

Type: string

Required: Conditional

## X-Amz-SignedHeaders

Specifies all the HTTP headers that were included as part of the canonical request. For more information about specifying signed headers, see [Create a signed AWS API request](#) in the *IAM User Guide*.

Condition: Specify this parameter when you include authentication information in a query string instead of in the HTTP authorization header.

Type: string

Required: Conditional

# Document history

The following table describes important additions to the AWS DataSync documentation. We also update the documentation frequently to address feedback that you send us.

To get notified about updates to this documentation, subscribe to the RSS feed.

Change	Description	Date
<a href="#">Updated S3 cross-account tutorial</a>	Removed some source account user permissions that are no longer required for this transfer.	June 10, 2024
<a href="#">New task execution status</a>	The CANCELLING status indicates when a task execution is being cancelled.	May 15, 2024
<a href="#">New option for pausing task schedules</a>	You can disable your AWS DataSync task schedule when you need to troubleshoot issues or perform storage system maintenance.	April 24, 2024
<a href="#">Updated TLS cipher for FIPS endpoints</a>	AWS DataSync uses the TLS_AES_128_GCM_SHA256 (secp256r1) cipher for Federal Information Processing Standard (FIPS) service endpoints.	April 22, 2024
<a href="#">AWS managed policy updates - Update to an existing policy</a>	The AWSDataSyncFullAccess policy has a new permission for services that work with DataSync.	February 16, 2024



---

<a href="#">Transfer specific files or objects with a manifest</a>	AWS DataSync can transfer a list of files or objects by using a manifest.	February 7, 2024
<a href="#">New AWS Region</a>	AWS DataSync is now available for data transfers in the Canada West (Calgary) Region.	December 20, 2023
<a href="#">Support for transfers with additional cloud providers</a>	AWS DataSync can now transfer data between AWS storage services and IBM Cloud Object Storage or Seagate Lyve Cloud.	November 7, 2023
<a href="#">Support for transfers with Alibaba Cloud Object Storage Service</a>	AWS DataSync can now transfer data between AWS storage services and Alibaba Cloud Object Storage Service.	September 25, 2023
<a href="#">Support for task reports</a>	Monitor your AWS DataSync transfers with task reports.	August 30, 2023
<a href="#">New AWS Region</a>	AWS DataSync is now available for data transfers in the Israel (Tel Aviv) Region.	August 23, 2023
<a href="#">Support for transfers with additional cloud providers</a>	AWS DataSync can now transfer data between AWS storage services and several other cloud providers (such as Wasabi Cloud Storage, DigitalOcean Spaces, and Oracle Cloud Infrastructure Object Storage).	August 8, 2023

---

<a href="#">General availability of Microsoft Azure Blob Storage support</a>	AWS DataSync can now transfer objects to and from Microsoft Azure Blob Storage.	July 25, 2023
<a href="#">TLS 1.3 support</a>	When transferring between storage locations, AWS DataSync now encrypts all network traffic with Transport Layer Security (TLS) 1.3.	June 28, 2023
<a href="#">New DataSync Discovery metrics</a>	AWS DataSync Discovery can now tell you how many LUNs (logical unit numbers) are in a storage resource cluster, storage virtual machine (SVM), or volume.	June 28, 2023
<a href="#">New AWS Region</a>	AWS DataSync is now available for data transfers in the Asia Pacific (Melbourne) Region.	May 24, 2023
<a href="#">Support for with S3 compatible storage on Snowball Edge</a>	You can use AWS DataSync to transfer data between Amazon S3 compatible storage on AWS Snowball Edge and AWS storage services.	May 18, 2023
<a href="#">AWS managed policy updates - Update to an existing policy</a>	The <code>AWSDataSyncFullAccess</code> policy has new permissions for services that work with DataSync.	May 2, 2023
<a href="#">General availability of AWS DataSync Discovery</a>	Use DataSync Discovery to help accelerate your migration to AWS.	April 25, 2023

---

<a href="#">Public preview release of Microsoft Azure Blob Storage support</a>	AWS DataSync can now transfer objects from Microsoft Azure Blob Storage.	March 29, 2023
<a href="#">New IAM policy</a>	To support the DataSync Discovery feature, DataSync uses the service-linked role named <code>AWSServiceRoleForDataSyncDiscovery</code> .	March 21, 2023
<a href="#">New AWS Regions</a>	AWS DataSync is now available in the following AWS Regions: Asia Pacific (Hyderabad), Europe (Spain), and Europe (Zurich).	February 6, 2023
<a href="#">Using tags in task executions</a>	You can now tag your AWS DataSync task executions.	December 16, 2022
<a href="#">Support for S3 Glacier Instant Retrieval</a>	You can now transfer objects directly into the S3 Glacier Instant Retrieval storage class.	December 16, 2022
<a href="#">Copying object system metadata</a>	AWS DataSync can now copy system metadata when transferring between an object storage system and Amazon S3.	December 16, 2022
<a href="#">New AWS Regions</a>	AWS DataSync is now available in the China (Beijing) and China (Ningxia) Regions.	December 14, 2022
<a href="#">New AWS Region</a>	AWS DataSync is now available in the Middle East (UAE) Region.	November 16, 2022

---

<a href="#">Support for self-signed certificates with object storage locations</a>	AWS DataSync can connect to object storage locations that use self-signed or private certificates.	October 25, 2022
<a href="#">Get data compression information</a>	AWS DataSync can provide the physical number of bytes transferred over the network after compression was applied.	October 25, 2022
<a href="#">Public preview release of AWS DataSync Discovery</a>	Use DataSync Discovery to help accelerate your migration to AWS.	September 21, 2022
<a href="#">New option for migrating data to or from Google Cloud Storage</a>	You can transfer data to or from Google Cloud Storage by deploying an AWS DataSync agent in Google Cloud.	July 21, 2022
<a href="#">Support for Amazon FSx for NetApp ONTAP file systems</a>	AWS DataSync can now transfer files and folders to and from FSx for ONTAP file systems.	June 28, 2022
<a href="#">New security options for Amazon EFS locations</a>	AWS DataSync can access Amazon EFS file systems using TLS, access points, and IAM roles.	May 31, 2022

---

<a href="#">Migrating data to or from Google Cloud Storage and Azure Files</a>	With AWS DataSync, you can transfer data to or from Google Cloud Storage and Azure Files. For more information, see <a href="#">Creating a location for object storage</a> and <a href="#">Creating a location for SMB</a> .	May 24, 2022
<a href="#">New AWS DataSync task setting</a>	With the <b>Copy object tags</b> option, you can specify whether to maintain object tags when transferring between object storage systems.	May 5, 2022
<a href="#">New AWS Region</a>	AWS DataSync is now available in the Asia Pacific (Jakarta) Region.	April 19, 2022
<a href="#">Support for Amazon FSx for OpenZFS file systems</a>	AWS DataSync can now transfer files and folders to and from FSx for OpenZFS file systems.	April 5, 2022
<a href="#">Support for Amazon FSx for Lustre file systems</a>	AWS DataSync can now transfer files and folders to and from FSx for Lustre file systems.	December 10, 2021
<a href="#">Support for Hadoop Distributed File Systems (HDFS)</a>	AWS DataSync now supports transferring files and folders to and from HDFS clusters.	November 3, 2021
<a href="#">New AWS Region</a>	AWS DataSync is now available in the Asia Pacific (Osaka) Region.	July 28, 2021

[Fully automated transfers between AWS storage services](#)

AWS DataSync can now transfer files or objects between Amazon S3, Amazon EFS, or FSx for Windows File Server with just a few clicks in the DataSync console.

November 9, 2020

[Adjusting the network bandwidth used by a running task](#)

AWS DataSync now enables customers to adjust the network bandwidth used by a running DataSync task. This helps to minimize impact on other users or applications when a task spans multiple days.

November 9, 2020

[Enhanced support for on-premises DataSync virtual machine \(VM\) functions](#)

The AWS DataSync agent VM host console now supports enhanced functions, including activating an agent from the local console.

October 19, 2020

[AWS DataSync can now transfer data to and from AWS Outposts](#)

DataSync now supports transferring objects to and from Amazon S3 on AWS Outposts.

September 30, 2020

[Support for API filtering](#)

AWS DataSync now supports filtering for the `ListTasks` and `ListLocations` API calls, enabling you to easily retrieve configuration of data transfer tasks by using filters such as the source or destination for the data transfer.

August 18, 2020

[Support for copying data from your self-managed object storage](#)

AWS DataSync now supports data transfer between self-managed object storage and Amazon S3, Amazon Elastic File System, or FSx for Windows File Server.

July 27, 2020

[Support for Linux Kernel-based Virtual Machine \(KVM\) and Microsoft Hyper-V hypervisors](#)

AWS DataSync now provides the ability to deploy on-premises agents on the KVM and Microsoft Hyper-V virtualization platforms , in addition to the existing VMware and Amazon EC2 options.

July 1, 2020

[AWS DataSync can now automatically configure your Amazon CloudWatch Logs configuration](#)

When using DataSync, you now have the option of automatically generating the CloudWatch log group and resource policy required to publish logs for your data transfer, simplifying task creation and monitoring setup.

July 1, 2020

---

<a href="#">AWS DataSync can now transfer data to and from AWS Snowcone</a>	DataSync now supports transferring files to and from AWS Snowcone, the smallest member of the AWS Snow Family of edge computing and data transfer devices. Snowcone is portable, ruggedized, and secure—small and light enough to fit in a backpack and able to withstand harsh environments.	June 17, 2020
<a href="#">New AWS Region</a>	AWS DataSync is now available in the Africa (Cape Town) Region and the Europe (Milan) Region.	June 16, 2020
<a href="#">Enhanced monitoring capabilities with file-level logging</a>	You can now enable detailed logging for files and objects copied between your NFS servers, SMB servers, Amazon S3 buckets, Amazon EFS file systems, and FSx for Windows File Server file systems.	April 24, 2020
<a href="#">Support for copying data between your SMB share and Amazon FSx for Windows File Server</a>	You can now copy data between your SMB share and FSx for Windows File Server.	January 24, 2020
<a href="#">Support for scheduling tasks</a>	You can now run tasks manually or schedule them to run based on a specified schedule.	November 20, 2019



---

<a href="#">New AWS Region</a>	AWS DataSync is now available in the Asia Pacific (Hong Kong) Region, Asia Pacific (Mumbai) Region, Europe (Stockholm) Region, South America (São Paulo) Region, and AWS GovCloud (US-East) Region.	November 20, 2019
<a href="#">New AWS Region</a>	AWS DataSync is now available in the Canada (Central) Region, Europe (London) Region, and Europe (Paris) Region.	October 2, 2019
<a href="#">Support for Amazon S3 storage classes</a>	You can now transfer objects directly into Amazon S3 storage classes.	September 24, 2019
<a href="#">New AWS Region</a>	AWS DataSync is now available in the Middle East (Bahrain) Region.	August 28, 2019
<a href="#">Support for copying data between your Server Message Block (SMB) share and Amazon S3 or Amazon EFS</a>	You can now copy data between your SMB file share and Amazon S3 or Amazon EFS.	August 22, 2019
<a href="#">Support for using virtual private cloud (VPC) endpoints</a>	You can now create a private connection between your agent and AWS and run tasks in a private network. Doing this increases the security of your data as it's copied over the network.	August 5, 2019

---

<a href="#">Support for Federal Information Processing Standard (FIPS) endpoints</a>	You can now use FIPS endpoints to create agents and run tasks.	August 5, 2019
<a href="#">New AWS Region</a>	AWS DataSync is now available in the AWS GovCloud (US-West) Region.	June 11, 2019
<a href="#">Support for filtering</a>	You can now apply filters to transfer only a subset of the files in your source location when you transfer data from your source to your destination location.	May 22, 2019
<a href="#">First release of AWS DataSync</a>	General release of the AWS DataSync service.	November 26, 2018

# AWS Glossary

For the latest AWS terminology, see the [AWS glossary](#) in the *AWS Glossary Reference*.