Network requirements

DataSync network requirements depend on how you plan to transfer data (for example, over the public internet or using a more private connection).

Use the following tables to help you configure network access for DataSync agents that transfer data from your self-managed storage system and through virtual private cloud (VPC), public service, Federal Information Processing Standard (FIPS) endpoints.
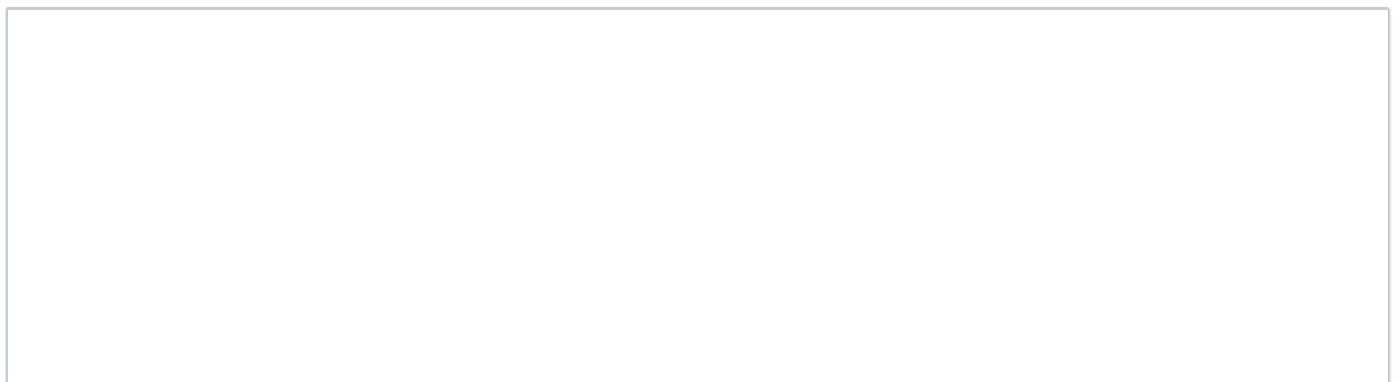
Topics

- Network requirements to connect to your self-managed storage (#on-premises-network-requirements)
- Network requirements when using VPC endpoints (#using-vpc-endpoint)
- Network requirements when using public service endpoints or FIPS endpoints (#using-public-endpoints)
- Required network interfaces for data transfers (#required-network-interfaces)

Network requirements to connect to your self-managed storage

To minimize network latency, deploy the DataSync agent close to your self-managed storage. This ensures that files travel over the network between the DataSync agent and the DataSync service using our purpose-built, accelerated protocol that significantly speeds up transfers.

The following ports are required for communication between the DataSync agent and your Network File System (NFS) server, Hadoop Distributed File System (HDFS) cluster, Server Message Block (SMB) server, or Amazon S3 API compatible storage.

From	To	Protocol	Port	How used by the DataSync agent
Agent	NFS server	TCP/UDP	2049 (NFS)	To mount a source NFS file system. Supports NFS v3.x, NFS v4.0, and NFS v4.1.
Agent	SMB server	TCP/UDP	139 (SMB) or 445 (SMB)	To mount a source SMB file share. Supports SMB 2.1 and SMB 3 versions.
Agent	Self-managed object storage	TCP	443 (HTTPS) or 80 (HTTP)	To access your self-managed object storage.
Agent	Hadoop cluster	TCP	NameNode port (default is 8020) In most clusters, you can find this port number in the `core-site.xml` file under the `fs.default` or `fs.default.name` property (depending on the Hadoop distribution).	To access the NameNodes in your Hadoop cluster. Specify the port used when creating an HDFS location.

Agent	Hadoop cluster	TCP	DataNode port (default is 50010) In most clusters, you can find this port number in the `hdfs-site.xml` file under the `dfs.datanode.address` property.	To access the DataNodes in your Hadoop cluster. The DataSync agent automatically determines the port to use.
Agent	Hadoop Key Management Server (KMS)	TCP	KMS port (default 9600)	To access the KMS for your Hadoop cluster.
Agent	Kerberos Key Distribution Center (KDC) server	TCP	KDC port (default 88)	When authenticating to the Kerberos realm. This port is used only with HDFS.

Network requirements when using VPC endpoints

If you use only private IP addresses, you can ensure that your VPC can't be reached over the internet, and you can prevent any packets from entering or exiting the network. By using private IP addresses, you can eliminate all internet access from your self-managed systems, and still use DataSync for data transfers to and from AWS.

DataSync requires the following ports for its operation when your agent is using private endpoints.

From	To	Protocol	Port	How used

From	To	Protocol	Port	How used
Your web browser	Your DataSync agent	TCP	80 (HTTP)	By your computer to obtain the agent activation key. After successful activation, DataSync closes the agent's port 80. 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. ⓘ **Note** Alternatively, you can obtain the activation key from the agent's local console. This method does not require connectivity between the browser and your agent. For more information about using the local console to get the activation key, see Obtaining an activation key using the local console (./local-console-vm.html#get-activation-key) .

From	To	Protocol	Port	How used
Agent	Your DataSync VPC endpoint To find the correct IP address, open the Amazon VPC console ☒ (https://console.aws.amazon.com/vpc/) , and choose **Endpoints** from the left navigation pane. Choose the DataSync endpoint, and check the **Subnets** list to find the private IP address that corresponds to the subnet that you chose for your VPC endpoint setup. For more information, see step 5 in Configuring DataSync to use private IP addresses for data transfer (./datasync-in-vpc.html#create-agent-steps-vpc) .	TCP	1024–10664	For control traffic between the DataSync agent and the AWS service.
Agent	Your task's elastic network interfaces To find the related IP addresses, open the Amazon EC2 console and choose **Network Interfaces** from the left navigation pane. To see the four network interfaces for the task, enter your task ID in the search filter. For more information, see step 9 in Configuring DataSync to use private IP addresses for data transfer (./datasync-in-vpc.html#create-agent-steps-vpc) .	TCP	443 (HTTPS)	For data transfer from the DataSync VM to the AWS service.

From	To	Protocol	Port	How used
Agent	Your DataSync VPC endpoint	TCP	22 (Support channel)	To allow AWS Support to access your DataSync to help you with troubleshooting DataSync issues. You don't need this port open for normal operation, but it's required for troubleshooting.

Following is an illustration of the ports required by DataSync when using private endpoints.
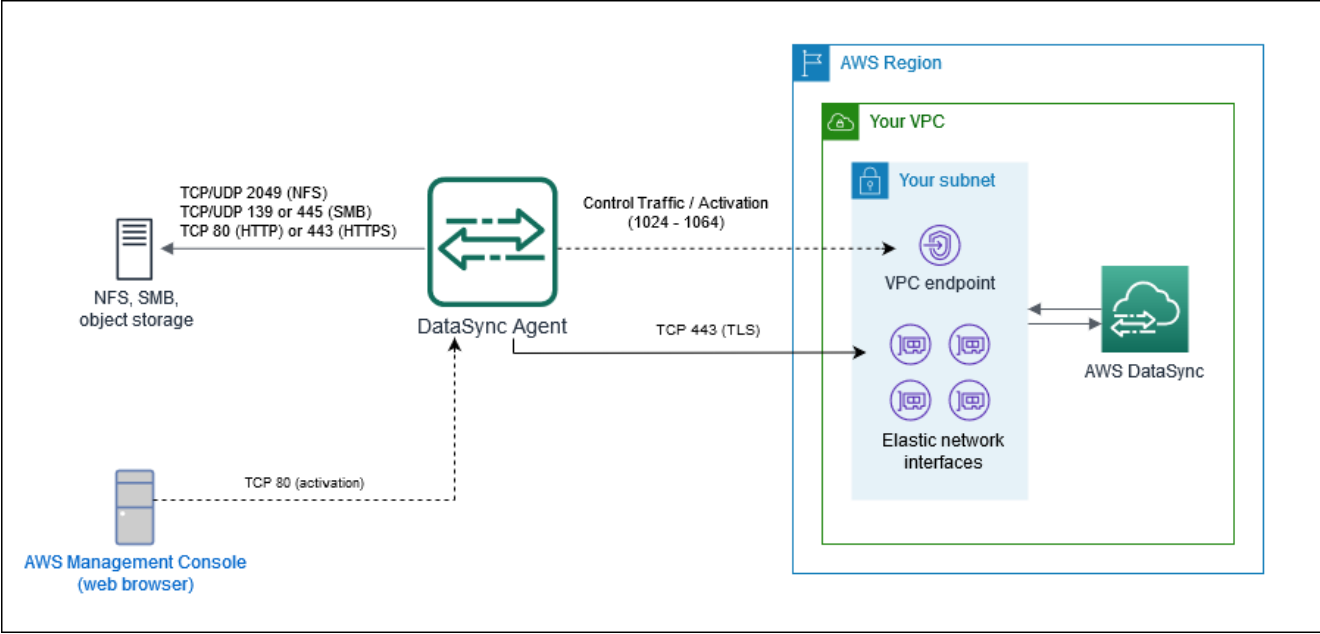
Network requirements when using public service endpoints or FIPS endpoints

Your agent VM requires access to the following endpoints to communicate with AWS when using public service endpoints, or when using FIPS endpoints. Enabling this access is not necessary when using DataSync with VPC endpoints.

If you use a firewall or router to filter or limit network traffic, configure your firewall or router to allow these service endpoints. They're required to enable outbound communication between your network and AWS.

From	To	Protocol	Port	How used	Endpoints accessed by the agent

From	To	Protocol	Port	How used	Endpoints accessed by the agent
Your web browser	DataSync agent	TCP	80 (HTTP)	Used by your computer to obtain the agent activation key. After successful activation, DataSync closes the agent's port 80. 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. ⓘ **Note** Alternatively, you can obtain the activation key from the agent's local console. This method does not require connectivity between the browser and your agent. For more information about using the local console to get the activation key, see Obtaining an activation key using the local console (./local-console-vm.html#get-activation-key) .	N/A

From	To	Protocol	Port	How used	Endpoints accessed by the agent

Agent	AWS	TCP	443 (HTTPS)	Used by the DataSync agent to activate with your AWS account. You can block the public endpoints after activation.	For public endpoint activation:
					
					`activation.datasync.`*`us-east-2`*`.amazonaws.com`
					
					For FIPS endpoint activation:
					
					`activation.datasync-fips.`*`us-east-2`*`.amazonaws.com`

From	To	Protocol	Port	How used	Endpoints accessed by the agent

Agent	AWS	TCP	443 (HTTPS)	For communication between the DataSync agent and the AWS service endpoint.
				
				For information about Regions and service endpoints, see Choose a service endpoint for AWS DataSync (./choose-service-endpoint.html) .
					API endpoints:
					
					`datasync.`*`us-east-2`*`.amazonaws.com`
					
					Data transfer endpoints:
					
					`yourTaskId``.datasync-dp.`*`us-east-2`*`.amazonaws.com`
					
					`cp.datasync.`*`us-east-2`*`.amazonaws.com`
					
					Data transfer endpoints for FIPS:
					
					`cp.datasync-fips.`*`us-east-2`*`.amazonaws.com`

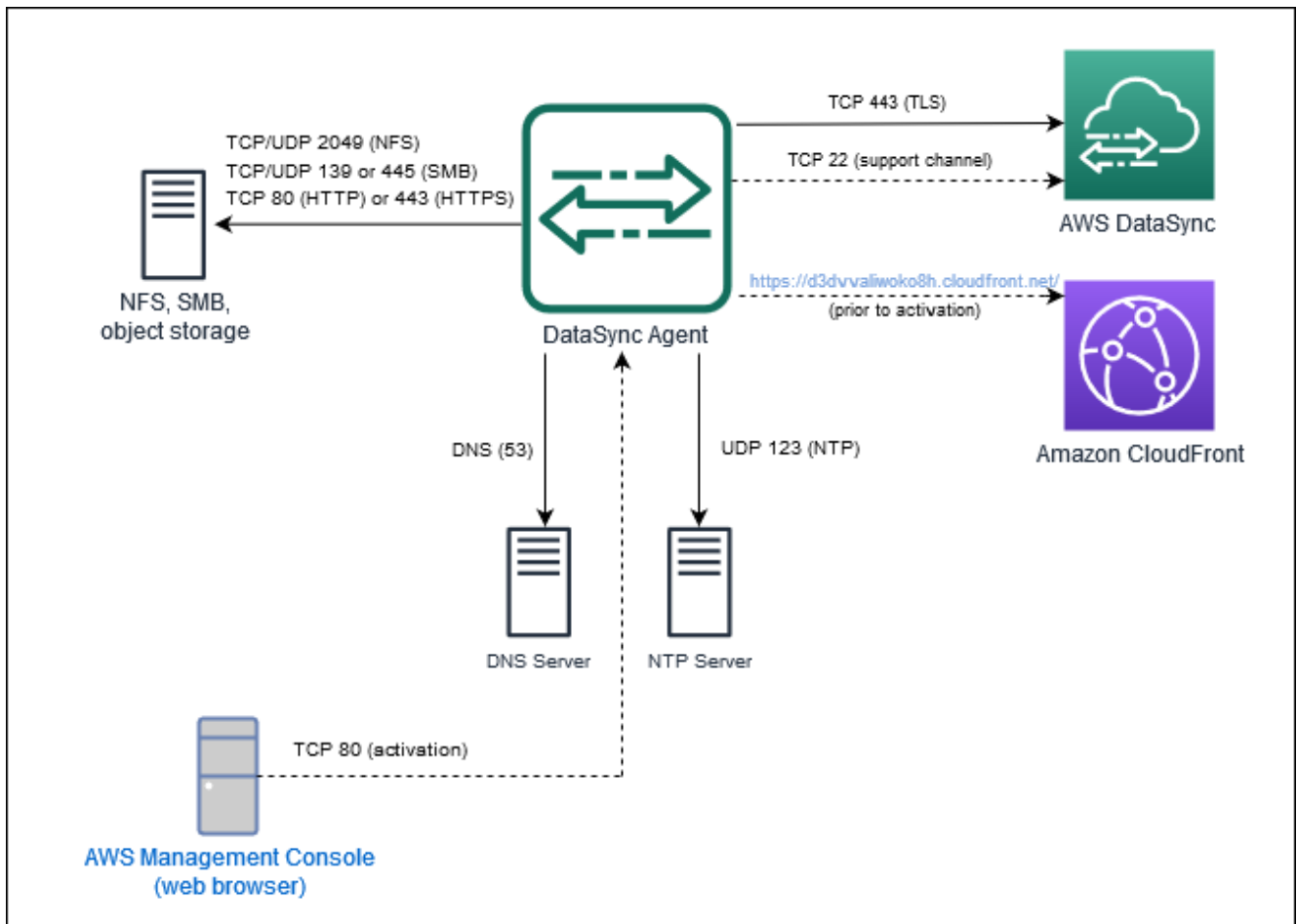| From	To	Protocol	Port	How used	Endpoints accessed by the agent

Agent	AWS	TCP	80 (HTTP)	Allows the DataSync agent to get updates from AWS.	The *activation_region* variable is the AWS Region you used to activate your DataSync agent.
					
					`repo.default.amazonaws.com`
					
					`packages.us-west-1.amazonaws.com`
					
					`packages.sa-east-1.amazonaws.com`
					
					`repo.$activation_region.amazonaws.com`
					
					`packages.$activation_region.amazonaws.com`

From	To	Protocol	Port	How used	Endpoints accessed by the agent

Agent	AWS	TCP	443 (HTTPS)	Allows the DataSync agent to get updates from AWS.	The *activation_region* variable is the AWS Region you used to activate your DataSync agent.
					
					`amazonlinux.default.amazonaws.com`
					
					`cdn.amazonlinux.com`
					
					`amazonlinux-2-repos-$activation_region.s3.dualstack.$activation_region.amazonaws.com`
					
					`amazonlinux-2-repos-$activation_region.s3.$activation_region.amazonaws.com`
					
					`*.s3.$activation_region.amazonaws.com`

From	To	Protocol	Port	How used	Endpoints accessed by the agent
Agent	Domain Name Service (DNS) server	TCP/UDP	53 (DNS)	For communication between DataSync agent and the DNS server.	N/A

From	To	Protocol	Port	How used	Endpoints accessed by the agent
Agent	AWS	TCP	22 (Support channel)	Allows AWS Support to access your DataSync to help you with troubleshooting DataSync issues. You don't need this port open for normal operation, but it's required for troubleshooting.	AWS support channel: `54.201.223.107`

From	To	Protocol	Port	How used	Endpoints accessed by the agent

Agent	Network Time Protocol (NTP) server	UDP	123 (NTP)	Used by local systems to synchronize the VM time to the host time.	NTP:
					
					`0.amazon.pool.ntp.org`
					
					`1.amazon.pool.ntp.org`
					
					`2.amazon.pool.ntp.org`
					
					`3.amazon.pool.ntp.org`
					
					ⓘ **Note**
					If you want to change the default NTP configuration of your VM agent to use a different NTP server using the local console, see Configuring a Network Time Protocol (NTP) server for VMware agents (./local-console-vm.html#time-management) .

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

TCP 443 (TLS)

TCP 22 (support channel)

AWS DataSync

TCP/UDP 2049 (NFS)
TCP/UDP 139 or 445 (SMB)
TCP 80 (HTTP) or 443 (HTTPS)

NFS, SMB,
object storage

DataSync Agent

https://d3dvvaliwoko8h.cloudfront.net/
(prior to activation)

Amazon CloudFront

DNS (53)

UDP 123 (NTP)

DNS Server

NTP Server

TCP 80 (activation)

AWS Management Console
(web browser)

Required network interfaces for data transfers

For every task you run, DataSync automatically creates and manages elastic network interfaces (https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/using-eni.html) for data transfer traffic. How many network interfaces DataSync creates and where they're created depends on the following details about your task:

- Whether your task requires a DataSync agent.
- Your source and destination locations (where you're copying data from and to).
- The type of endpoint used to activate your agent.

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.

Transfers with agents

You need a DataSync agent when copying data between a self-managed storage system and an AWS storage service.

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 (network interfaces aren't needed since DataSync communicates directly with the S3 bucket)	The subnet you specified 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 preferred file server for the file system.	
Amazon FSx for Lustre	4	The same subnet as the file system.	
Amazon FSx for OpenZFS	4	The same subnet as the file system.	

Transfers without agents

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

> ⓘ **Note**
>
> The total number of network interfaces depends on your DataSync task locations. For example, transferring from an Amazon EFS location to FSx for Lustre requires four network interfaces. Meanwhile, transferring from an FSx for Windows File Server to an Amazon S3 bucket requires two network interfaces.

Location	Network interfaces created by default	Where network interfaces are created

Location	Network interfaces created by default	Where network interfaces are created
Amazon S3	N/A (network interfaces aren't needed since DataSync communicates directly with the S3 bucket)	
Amazon EFS	2	The subnet you specify when creating the Amazon EFS location.
Amazon FSx for Windows File Server	2	The same subnet as the preferred file server for the file system.
Amazon FSx for Lustre	2	The same subnet as the file system.
Amazon FSx for OpenZFS	2	The same subnet as the file system.

To see the network interfaces allocated for your DataSync task, use the DescribeTask (https://docs.aws.amazon.com/datasync/latest/userguide/API_DescribeTask.html) operation.