SAN concepts

ONTAP 9

NetApp

December 01, 2022

Table of Contents

SAN concepts

About SAN host provisioning

SAN provisioning with iSCSI

In SAN environments, storage systems are targets that have storage target devices. For iSCSI and FC, the storage target devices are referred to as LUNs (logical units). For Non-Volatile Memory Express (NVMe) over Fibre Channel, the storage target devices are referred to as namespaces.

You configure storage by creating LUNs for iSCSI and FC or by creating namespaces for NVMe. The LUNs or namespaces are then accessed by hosts using Internet Small Computer Systems Interface (iSCSI) or Fibre Channel (FC) protocol networks.

To connect to iSCSI networks, hosts can use standard Ethernet network adapters (NICs), TCP offload engine (TOE) cards with software initiators, converged network adapters (CNAs), or dedicated iSCSI host bus adapters (HBAs).

To connect to FC networks, hosts require FC HBAs or CNAs.

Supported FC protocols include:

- FC
- FCoE
- NVMe

iSCSI target node network connections and names

iSCSI target nodes can connect to the network in several ways:

- Over Ethernet interfaces using software that is integrated into ONTAP.
- Over multiple system interfaces, with an interface used for iSCSI that can also transmit traffic for other protocols, such as SMB and NFS.
- Using a unified target adapter (UTA) or a converged network adapter (CNA).

Every iSCSI node must have a node name.

The two formats, or type designators, for iSCSI node names are *iqn* and *eui*. The SVM iSCSI target always uses the iqn-type designator. The initiator can use either the iqn-type or eui-type designator.

Storage system node name

Each SVM running iSCSI has a default node name based on a reverse domain name and a unique encoding number.

The node name is displayed in the following format:

iqn.1992-08.com.netapp:sn.*`unique-encoding-number`*

The following example shows the default node name for a storage system with a unique encoding number:

iqn.1992-08.com.netapp:sn.812921059e6c11e097b3123478563412:vs.6

TCP port for iSCSI

The iSCSI protocol is configured in ONTAP to use TCP port number 3260.

ONTAP does not support changing the port number for iSCSI. Port number 3260 is registered as part of the iSCSI specification and cannot be used by any other application or service.

Related information

NetApp Documentation: ONTAP SAN Host Configuration

iSCSI service management

iSCSI service management

You can manage the availability of the iSCSI service on the iSCSI logical interfaces of the storage virtual machine (SVM) by using the `vserver iscsi interface enable` or `vserver iscsi interface disable` commands.

By default, the iSCSI service is enabled on all iSCSI logical interfaces.

How iSCSI is implemented on the host

iSCSI can be implemented on the host using hardware or software.

You can implement iSCSI in one of the following ways:

- Using Initiator software that uses the host's standard Ethernet interfaces.
- Through an iSCSI host bus adapter (HBA): An iSCSI HBA appears to the host operating system as a SCSI disk adapter with local disks.
- Using a TCP Offload Engine (TOE) adapter that offloads TCP/IP processing.

  The iSCSI protocol processing is still performed by host software.

How iSCSI authentication works

During the initial stage of an iSCSI session, the initiator sends a login request to the storage system to begin an iSCSI session. The storage system then either permits or denies the login request, or determine that a login is not required.

iSCSI authentication methods are:

- Challenge Handshake Authentication Protocol (CHAP)--The initiator logs in using a CHAP user name and password.

  You can specify a CHAP password or generate a hexadecimal secret password. There are two types of CHAP user names and passwords:

  - Inbound—The storage system authenticates the initiator.

Inbound settings are required if you are using CHAP authentication.

  - Outbound—This is an optional setting to enable the initiator to authenticate the storage system.

    You can use outbound settings only if you define an inbound user name and password on the storage system.

- deny—The initiator is denied access to the storage system.
- none—The storage system does not require authentication for the initiator.

You can define the list of initiators and their authentication methods. You can also define a default authentication method that applies to initiators that are not on this list.

Related information

Windows Multipathing Options with Data ONTAP: Fibre Channel and iSCSI

iSCSI initiator security management

ONTAP provides a number of features for managing security for iSCSI initiators. You can define a list of iSCSI initiators and the authentication method for each, display the initiators and their associated authentication methods in the authentication list, add and remove initiators from the authentication list, and define the default iSCSI initiator authentication method for initiators not in the list.

iSCSI endpoint isolation

Beginning with ONTAP 9.1 existing iSCSI security commands were enhanced to accept an IP address range, or multiple IP addresses.

All iSCSI initiators must provide origination IP addresses when establishing a session or connection with a target. This new functionality prevents an initiator from logging into the cluster if the origination IP address is unsupported or unknown, providing a unique identification scheme. Any initiator originating from an unsupported or unknown IP address will have their login rejected at the iSCSI session layer, preventing the initiator from accessing any LUN or volume within the cluster.

Implement this new functionality with two new commands to help manage pre-existing entries.

Add initiator address range

Improve iSCSI initiator security management by adding an IP address range, or multiple IP addresses with the `vserver iscsi security add-initiator-address-range` command.

```
cluster1::> vserver iscsi security add-initiator-address-range
```

Remove initiator address range

Remove an IP address range, or multiple IP addresses, with the `vserver iscsi security remove-initiator-address-range` command.

```
cluster1::> vserver iscsi security remove-initiator-address-range
```

What CHAP authentication is

The Challenge Handshake Authentication Protocol (CHAP) enables authenticated communication between iSCSI initiators and targets. When you use CHAP authentication, you define CHAP user names and passwords on both the initiator and the storage system.

During the initial stage of an iSCSI session, the initiator sends a login request to the storage system to begin the session. The login request includes the initiator's CHAP user name and CHAP algorithm. The storage system responds with a CHAP challenge. The initiator provides a CHAP response. The storage system verifies the response and authenticates the initiator. The CHAP password is used to compute the response.

Guidelines for using CHAP authentication

You should follow certain guidelines when using CHAP authentication.

- If you define an inbound user name and password on the storage system, you must use the same user name and password for outbound CHAP settings on the initiator. If you also define an outbound user name and password on the storage system to enable bidirectional authentication, you must use the same user name and password for inbound CHAP settings on the initiator.
- You cannot use the same user name and password for inbound and outbound settings on the storage system.
- CHAP user names can be 1 to 128 bytes.

  A null user name is not allowed.

- CHAP passwords (secrets) can be 1 to 512 bytes.

  Passwords can be hexadecimal values or strings. For hexadecimal values, you should enter the value with a prefix of "0x" or "0X". A null password is not allowed.

- For additional restrictions, you should see the initiator's documentation.

  For example, the Microsoft iSCSI software initiator requires both the initiator and target CHAP passwords to be at least 12 bytes if IPsec encryption is not being used. The maximum password length is 16 bytes regardless of whether IPsec is used.

How using iSCSI interface access lists to limit initiator interfaces can increase performance and security

ISCSI interface access lists can be used to limit the number of LIFs in an SVM that an initiator can access, thereby increasing performance and security.

When an initiator begins a discovery session using an iSCSI `SendTargets` command, it receives the IP addresses associated with the LIF (network interface) that is in the access list. By default, all initiators have access to all iSCSI LIFs in the SVM. You can use the access list to restrict the number of LIFs in an SVM that an initiator has access to.

iSNS server registration requirement

What iSNS is

The Internet Storage Name Service (iSNS) is a protocol that enables automated discovery and management of iSCSI devices on a TCP/IP storage network. An iSNS server maintains information about active iSCSI devices on the network, including their IP addresses, iSCSI node names IQN’s, and portal groups.

You can obtain an iSNS server from a third-party vendor. If you have an iSNS server on your network configured and enabled for use by the initiator and target, you can use the management LIF for a storage virtual machine (SVM) to register all the iSCSI LIFs for that SVM on the iSNS server. After the registration is complete, the iSCSI initiator can query the iSNS server to discover all the LIFs for that particular SVM.

If you decide to use an iSNS service, you must ensure that your storage virtual machines (SVMs) are properly registered with an Internet Storage Name Service (iSNS) server.

If you do not have an iSNS server on your network, you must manually configure each target to be visible to the host.

What an iSNS server does

An iSNS server uses the Internet Storage Name Service (iSNS) protocol to maintain information about active iSCSI devices on the network, including their IP addresses, iSCSI node names (IQNs), and portal groups.

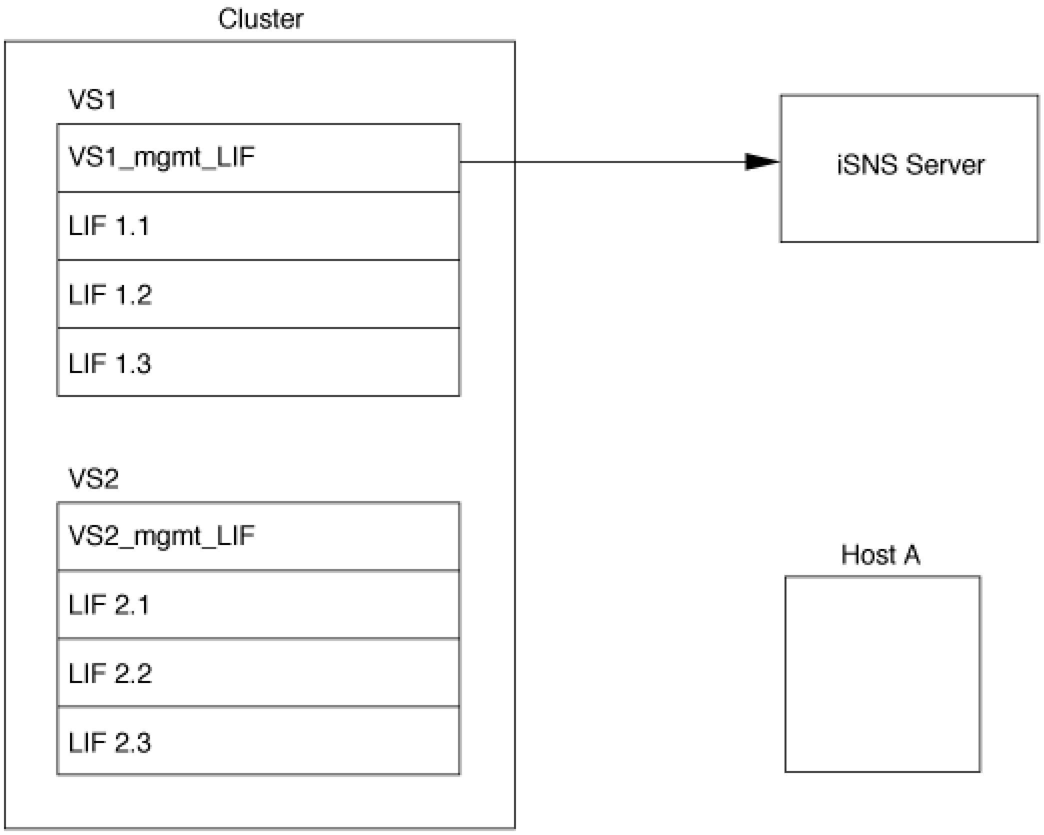
The iSNS protocol enables automated discovery and management of iSCSI devices on an IP storage network. An iSCSI initiator can query the iSNS server to discover iSCSI target devices.

NetApp does not supply or resell iSNS servers. You can obtain these servers from a vendor supported by NetApp.

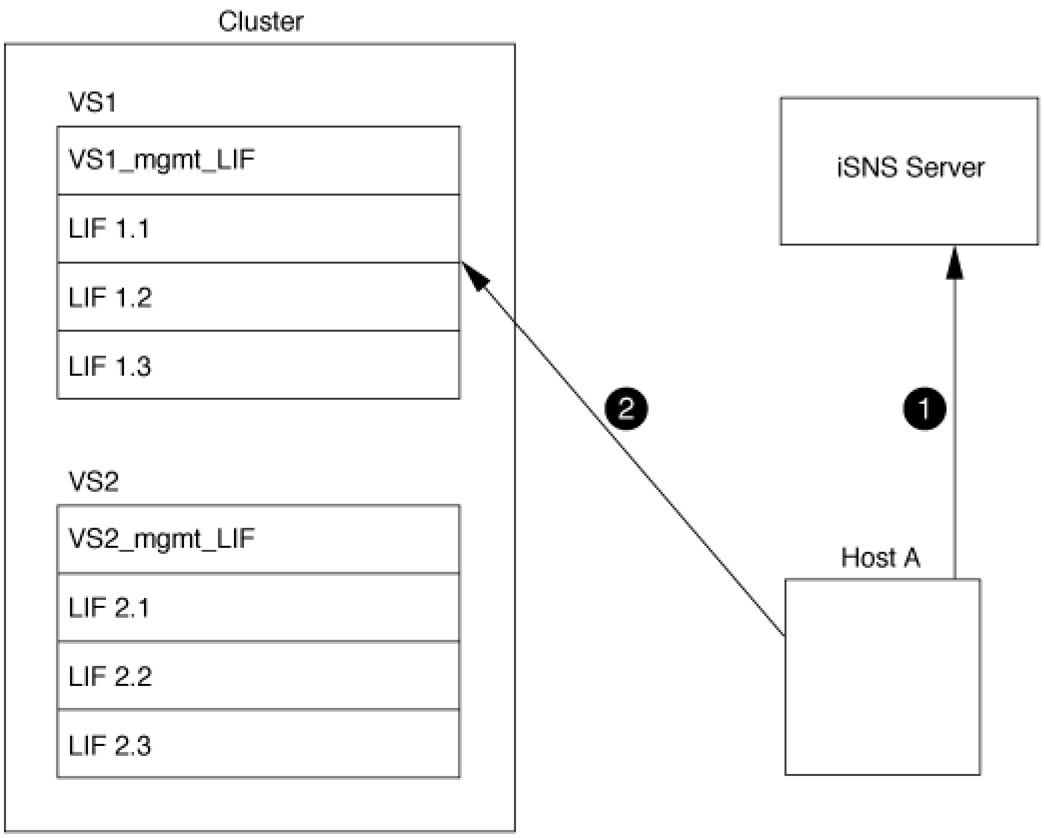
How SVMs interact with an iSNS server

The iSNS server communicates with each storage virtual machine (SVM) through the SVM management LIF. The management LIF registers all iSCSI target node name, alias, and portal information with the iSNS service for a specific SVM.

In the following example, SVM VS1 uses the SVM management LIF vs1_mgmt_lif to register with the iSNS server. During iSNS registration, an SVM sends all the iSCSI LIFs through the SVM management LIF to the iSNS Server. After the iSNS registration is complete, the iSNS server has a list of all the LIFs serving iSCSI in VS1. If a cluster contains multiple SVMs, each SVM must register individually with the iSNS server to use the iSNS service.



In the next example, after the iSNS server completes the registration with the target, Host A can discover all the LIFs for VS1 through the iSNS server as indicated in step 1. After Host A completes the discovery of the LIFs for VS1, Host A can establish a connection with any of the LIFs in VS1 as shown in step 2. Host A is not aware of any of the LIFs in VS2 until the management LIF VS2_mgmt_LIF for VS2 registers with the iSNS server.



However, if you define the interface access lists, the host can only use the defined LIFs in the interface access list to access the target.

After iSNS is initially configured, ONTAP automatically updates the iSNS server when the SVM configuration settings change.

A delay of a few minutes can occur between the time you make the configuration changes and when ONTAP sends the update to the iSNS server. Force an immediate update of the iSNS information on the iSNS server: `vserver iscsi isns update`

Commands for managing iSNS

ONTAP provides commands to manage your iSNS service.

If you want to…	Use this command…
Configure an iSNS service	`vserver iscsi isns create`
Start an iSNS service	`vserver iscsi isns start`
Modify an iSNS service	`vserver iscsi isns modify`
Display iSNS service configuration	`vserver iscsi isns show`
Force an update of registered iSNS information	`vserver iscsi isns update`

If you want to…	Use this command…
Stop an iSNS service	`vserver iscsi isns stop`
Remove an iSNS service	`vserver iscsi isns delete`
View the man page for a command	`man` *command name*

See the man page for each command for more information.

SAN provisioning with FC

You should be aware of the important concepts that are required to understand how ONTAP implements an FC SAN.

How FC target nodes connect to the network

Storage systems and hosts have adapters so that they can be connected to FC switches with cables.

When a node is connected to the FC SAN, each SVM registers the World Wide Port Name (WWPN) of its LIF with the switch Fabric Name Service. The WWNN of the SVM and the WWPN of each LIF is automatically assigned by ONTAP..

> Direct-connection to nodes from hosts with FC is not supported, NPIV is required and this requires a switch to be used.With iSCSI sessions, communication works with connections that are either network routed or direct-connect. However, both of these methods are supported with ONTAP.

How FC nodes are identified

Each SVM configured with FC is identified by a worldwide node name (WWNN).

How WWPNs are used

WWPNs identify each LIF in an SVM configured to support FC. These LIFs utilize the physical FC ports in each node in the cluster, which can be FC target cards, UTA or UTA2 configured as FC or FCoE in the nodes.

- Creating an initiator group

  The WWPNs of the host’s HBAs are used to create an initiator group (igroup). An igroup is used to control host access to specific LUNs. You can create an igroup by specifying a collection of WWPNs of initiators in an FC network. When you map a LUN on a storage system to an igroup, you can grant all the initiators in that group access to that LUN. If a host’s WWPN is not in an igroup that is mapped to a LUN, that host does not have access to the LUN. This means that the LUNs do not appear as disks on that host.

  You can also create port sets to make a LUN visible only on specific target ports. A port set consists of a group of FC target ports. You can bind an igroup to a port set. Any host in the igroup can access the LUNs only by connecting to the target ports in the port set.

- Uniquely identifying FC LIFs

WWPNs uniquely identify each FC logical interface. The host operating system uses the combination of the WWNN and WWPN to identify SVMs and FC LIFs. Some operating systems require persistent binding to ensure that the LUN appears at the same target ID on the host.

How worldwide name assignments work

Worldwide names are created sequentially in ONTAP. However, because of the way ONTAP assigns them, they might appear to be assigned in a non-sequential order.

Each adapter has a pre-configured WWPN and WWNN, but ONTAP does not use these pre-configured values. Instead, ONTAP assigns its own WWPNs or WWNNs, based on the MAC addresses of the onboard Ethernet ports.

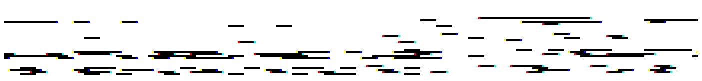
The worldwide names might appear to be non-sequential when assigned for the following reasons:

- Worldwide names are assigned across all the nodes and storage virtual machines (SVMs) in the cluster.
- Freed worldwide names are recycled and added back to the pool of available names.

How FC switches are identified

Fibre Channel switches have one worldwide node name (WWNN) for the device itself, and one worldwide port name (WWPN) for each of its ports.

For example, the following diagram shows how the WWPNs are assigned to each of the ports on a 16-port Brocade switch. For details about how the ports are numbered for a particular switch, see the vendor-supplied documentation for that switch.



Port **0**, WWPN 20:**00**:00:60:69:51:06:b4

Port **1**, WWPN 20:**01**:00:60:69:51:06:b4

Port **14**, WWPN 20:**0e**:00:60:69:51:06:b4

Port **15**, WWPN 20:**0f**:00:60:69:51:06:b4

SAN provisioning with NVMe

Beginning with ONTAP 9.4, NVMe/FC is supported in SAN environment. NVMe/FC enables storage administrators to provision namespaces and subsystems and then map the namespaces to subsystems, similar to the way LUNs are provisioned and mapped to igroups for FC and iSCSI.

An NVMe namespace is a quantity of non-volatile memory that can be formatted into logical blocks. Namespaces are the equivalent of LUNs for FC and iSCSI protocols, and an NVMe subsystem is analogous to an igroup. An NVMe subsystem can be associated with initiators so that namespaces within the subsystem can be accessed by the associated initiators.

Although analogous in function, NVMe namespaces do not support all features supported by LUNs.

Beginning with ONTAP 9.5 a license is required to support host-facing data access with NVMe. If NVMe is enabled in ONTAP 9.4, a 90 day grace period is given to acquire the license after upgrading to ONTAP 9.5. You can enable the license using the following command:

```
system license add -license-code NVMe_license_key
```

Related information

NetApp Technical Report 4684: Implementing and Configuring Modern SANs with NVMe/FC

About SAN volumes

About SAN volumes overview

ONTAP provides three basic volume provisioning options: thick provisioning, thin provisioning, and semi-thick provisioning. Each option uses different ways to manage the volume space and the space requirements for ONTAP block sharing technologies. Understanding how the options work enables you to choose the best option for your environment.

Putting SAN LUNs and NAS shares in the same FlexVol volume is not recommended. You should provision separate FlexVol volumes specifically for your SAN LUNs and you should provision separate FlexVol volumes specifically to your NAS shares. This simplifies management and replication deployments and parallels the way FlexVol volumes are supported in Active IQ Unified Manager (formerly OnCommand Unified Manager).

Thin provisioning for volumes

When a thinly provisioned volume is created, ONTAP does not reserve any extra space when the volume is created. As data is written to the volume, the volume requests the storage it needs from the aggregate to accommodate the write operation. Using thin-provisioned volumes enables you to overcommit your aggregate, which introduces the possibility of the volume not being able to secure the space it needs when the aggregate runs out of free space.

You create a thin-provisioned FlexVol volume by setting its `-space-guarantee` option to `none`.

Thick provisioning for volumes

When a thick-provisioned volume is created, ONTAP sets aside enough storage from the aggregate to ensure that any block in the volume can be written to at any time. When you configure a volume to use thick provisioning, you can employ any of the ONTAP storage efficiency capabilities, such as compression and deduplication, to offset the larger upfront storage requirements.

You create a thick-provisioned FlexVol volume by setting its `-space-slo` (service level objective) option to `thick`.

Semi-thick provisioning for volumes

When a volume using semi-thick provisioning is created, ONTAP sets aside storage space from the aggregate to account for the volume size. If the volume is running out of free space because blocks are in use by block-sharing technologies, ONTAP makes an effort to delete protection data objects (Snapshot copies and FlexClone files and LUNs) to free up the space they are holding. As long as ONTAP can delete the protection data objects fast enough to keep pace with the space required for overwrites, the write operations continue to succeed. This is called a “best effort” write guarantee.

Note: The following functionality is not supported on volumes that use semi-thick provisioning:

- storage efficiency technologies such as deduplication, compression, and compaction
- Microsoft Offloaded Data Transfer (ODX)

You create a semi-thick-provisioned FlexVol volume by setting its `-space-slo` (service level objective) option to `semi-thick`.

Use with space-reserved files and LUNs

A space-reserved file or LUN is one for which storage is allocated when it is created. Historically, NetApp has used the term “thin-provisioned LUN” to mean a LUN for which space reservation is disabled (a non-space-reserved LUN).

Note: Non-space-reserved files are not generally referred to as “thin-provisioned files”.

The following table summarizes the major differences in how the three volume provisioning options can be used with space-reserved files and LUNs:

Volume provisioning	LUN/file space reservation	Overwrites	Protection data[2]	Storage efficiency[3]
Thick	Supported	Guaranteed[1]	Guaranteed	Supported
Thin	No effect	None	Guaranteed	Supported
Semi-thick	Supported	Best effort[1]	Best effort	Not supported

Notes

1. The ability to guarantee overwrites or provide a best-effort overwrite assurance requires that space reservation is enabled on the LUN or file.
2. Protection data includes Snapshot copies, and FlexClone files and LUNs marked for automatic deletion (backup clones).
3. Storage efficiency includes deduplication, compression, any FlexClone files and LUNs not marked for automatic deletion (active clones), and FlexClone subfiles (used for Copy Offload).

Support for SCSI thin-provisioned LUNs

ONTAP supports T10 SCSI thin-provisioned LUNs as well as NetApp thin-provisioned LUNs. T10 SCSI thin provisioning enables host applications to support SCSI features including LUN space reclamation and LUN space monitoring capabilities for blocks environments. T10 SCSI thin provisioning must be supported by your SCSI host software.

You use the ONTAP `space-allocation` setting to enable/disable support for the T10 thin provisioning on a LUN. You use the ONTAP `space-allocation enable` setting to enable T10 SCSI thin provisioning on a LUN.

The `[-space-allocation {enabled|disabled}]` command in the ONTAP Command Reference Manual has more information to enable/disable support for the T10 thin provisioning and to enable T10 SCSI thin provisioning on a LUN.

ONTAP 9 commands

Configure volume provisioning options

You can configure a volume for thin provisioning, thick provisioning, or semi-thick provisioning.

About this task

Setting the `-space-slo` option to `thick` ensures the following:

- The entire volume is preallocated in the aggregate. You cannot use the `volume create` or `volume modify` command to configure the volume's `-space-guarantee` option.
- 100% of the space required for overwrites is reserved. You cannot use the `volume modify` command to configure the volume's `-fractional-reserve` option

Setting the `-space-slo` option to `semi-thick` ensures the following:

- The entire volume is preallocated in the aggregate. You cannot use the `volume create` or `volume modify` command to configure the volume's `-space-guarantee` option.
- No space is reserved for overwrites. You can use the `volume modify` command to configure the volume's `-fractional-reserve` option.
- Automatic deletion of Snapshot copies is enabled.

Step

1. Configure volume provisioning options:

   `volume create -vserver *vserver_name* -volume *volume_name* -aggregate *aggregate_name* -space-slo none|thick|semi-thick -space-guarantee none|volume`

   The `-space-guarantee` option defaults to `none` for AFF systems and for non-AFF DP volumes. Otherwise, it defaults to `volume`. For existing FlexVol volumes, use the `volume modify` command to configure provisioning options.

   The following command configures vol1 on SVM vs1 for thin provisioning:

   ```
   cluster1::> volume create –vserver vs1 -volume vol1 -space-guarantee
   none
   ```

   The following command configures vol1 on SVM vs1 for thick provisioning:

```
cluster1::> volume create –vserver vs1 -volume vol1 -space-slo thick
```

The following command configures vol1 on SVM vs1 for semi-thick provisioning:

```
cluster1::> volume create –vserver vs1 -volume vol1 -space-slo semi-
thick
```

SAN volume configuration options

You must set various options on the volume containing your LUN. The way you set the volume options determines the amount of space available to LUNs in the volume.

Autogrow

You can enable or disable Autogrow. If you enable it, autogrow allows ONTAP to automatically increase the size of the volume up to a maximum size that you predetermine. There must be space available in the containing aggregate to support the automatic growth of the volume. Therefore, if you enable autogrow, you must monitor the free space in the containing aggregate and add more when needed.

Autogrow cannot be triggered to support Snapshot creation. If you attempt to create a Snapshot copy and there is insufficient space on the volume, the Snapshot creation fails, even with autogrow enabled.

If autogrow is disabled, the size of your volume will remain the same.

Autoshrink

You can enable or disable Autoshrink. If you enable it, autoshrink allows ONTAP to automatically decrease the overall size of a volume when the amount of space consumed in the volume decreases a predetermined threshold. This increases storage efficiency by triggering volumes to automatically release unused free space.

Snapshot autodelete

Snapshot autodelete automatically deletes Snapshot copies when one of the following occurs:

- The volume is nearly full.
- The Snapshot reserve space is nearly full.
- The overwrite reserve space is full.

You can configure Snapshot autodelete to delete Snapshot copies from oldest to newest or from newest to oldest. Snapshot autodelete does not delete Snapshot copies that are linked to Snapshot copies in cloned volumes or LUNs.

If your volume needs additional space and you have enabled both autogrow and Snapshot autodelete, by default, ONTAP attempts to acquire the needed space by triggering autogrow first. If enough space is not acquired through autogrow, then Snapshot autodelete is triggered.

Snapshot reserve

Snapshot reserve defines the amount of space in the volume reserved for Snapshot copies. Space allocated to Snapshot reserve cannot be used for any other purpose. If all of the space allocated for Snapshot reserve is used, then Snapshot copies begin to consume additional space on the volume.

Requirement for moving volumes in SAN environments

Before you move a volume that contains LUNs or namespaces, you must meet certain requirements.

- For volumes containing one or more LUNs, you should have a minimum of two paths per LUN (LIFs) connecting to each node in the cluster.

  This eliminates single points of failure and enables the system to survive component failures.

- For volumes containing namespaces, the cluster must be running ONTAP 9.6 or later.

  Volume move is not supported for NVMe configurations running ONTAP 9.5.

Considerations for setting fractional reserve

Fractional reserve, also called *LUN overwrite reserve*, enables you to turn off overwrite reserve for space-reserved LUNs and files in a FlexVol volume. This can help you maximize your storage utilization, but if your environment is negatively affected by write operations failing due to lack of space, you must understand the requirements that this configuration imposes.

The fractional reserve setting is expressed as a percentage; the only valid values are `0` and `100` percent. The fractional reserve setting is an attribute of the volume.

Setting fractional reserve to `0` increases your storage utilization. However, an application accessing data residing in the volume could experience a data outage if the volume is out of free space, even with the volume guarantee set to `volume`. With proper volume configuration and use, however, you can minimize the chance of writes failing. ONTAP provides a “best effort” write guarantee for volumes with fractional reserve set to `0` when *all* of the following requirements are met:

- Deduplication is not in use
- Compression is not in use
- FlexClone sub-files are not in use
- All FlexClone files and FlexClone LUNs are enabled for automatic deletion

  This is not the default setting. You must explicitly enable automatic deletion, either at creation time or by modifying the FlexClone file or FlexClone LUN after it is created.

- ODX and FlexClone copy offload are not in use
- Volume guarantee is set to `volume`
- File or LUN space reservation is `enabled`
- Volume Snapshot reserve is set to `0`

- Volume Snapshot copy automatic deletion is `enabled` with a commitment level of `destroy`, a destroy list of `lun_clone,vol_clone,cifs_share,file_clone,sfsr`, and a trigger of `volume`

  This setting also ensures that FlexClone files and FlexClone LUNs are deleted when necessary.

Note that if your rate of change is high, in rare cases the Snapshot copy automatic deletion could fall behind, resulting in the volume running out of space, even with all of the above required configuration settings in use.

In addition, you can optionally use the volume autogrow capability to decrease the likelihood of volume Snapshot copies needing to be deleted automatically. If you enable the autogrow capability, you must monitor the free space in the associated aggregate. If the aggregate becomes full enough that the volume is prevented from growing, more Snapshot copies will probably be deleted as the free space in the volume is depleted.

If you cannot meet all of the above configuration requirements and you need to ensure that the volume does not run out of space, you must set the volume's fractional reserve setting to `100`. This requires more free space up front, but guarantees that data modification operations will succeed even when the technologies listed above are in use.

The default value and allowed values for the fractional reserve setting depend on the guarantee of the volume:

Volume guarantee	Default fractional reserve	Allowed values
Volume	100	0, 100
None	0	0, 100

About host-side space management

Host-side space management overview

In a thinly provisioned environment, host side space management completes the process of managing space from the storage system that has been freed in the host file system.

A host file system contains metadata to keep track of which blocks are available to store new data and which blocks contain valid data that must not be overwritten. This metadata is stored within the LUN. When a file is deleted in the host file system, the file system metadata is updated to mark that file's blocks as free space. Total file system free space is then recalculated to include the newly freed blocks. To the storage system, these metadata updates appear no different from any other writes being performed by the host. Therefore, the storage system is unaware that any deletions have occurred.

This creates a discrepancy between the amount of free space reported by the host and the amount of free space reported by the underlying storage system. For example, suppose you have a newly provisioned 200-GB LUN assigned to your host by your storage system. Both the host and the storage system report 200 GB of free space. Your host then writes 100 GB of data. At this point, both the host and storage system report 100 GB of used space and 100 GB of unused space.

Then you delete 50 GB of data from your host. At this point, your host will report 50 GB of used space and 150 GB of unused space. However, your storage system will report 100 GB of used space and 100 GB of unused space.

Host-side space management uses various methods to reconcile the space differential between the host and the storage system.

Automatic host-side space management with SCSI thinly provisioned LUNs

If your host supports SCSI thin provisioning, you can enable the `space-allocation` option in ONTAP to turn on automatic host-side space management.

Enabling SCSI thin provisioning enables you to do the following.

- Automatic host-side space management

  When data is deleted on a host that supports SCSI thin provisioning, host-side space management identifies the blocks of deleted data on the host file system and automatically issues one or more `SCSI UNMAP` commands to free corresponding blocks on the storage system.

- Notify the host when a LUN runs out of space while keeping the LUN online

  On hosts that do not support SCSI thin provisioning, when the volume containing LUN runs out of space and cannot automatically grow, ONTAP takes the LUN offline. However, on hosts that support SCSI thin provisioning, ONTAP does not take the LUN offline when it runs out of space. The LUN remains online in read-only mode and the host is notified that the LUN can no longer accept writes.

Related information

ONTAP SAN host configuration

Enable space allocation for SCSI thinly provisioned LUNs

If you set the `space-allocation` option to `enabled`, ONTAP notifies the host when the volume has run out of space and the LUN in the volume cannot accept writes. This option also enables ONTAP to reclaim space automatically when your host deletes data.

About this task

The `space-allocation` option is set to `disabled` by default, and you must take the LUN offline to enable space allocation. After you enable space allocation, you must perform discovery on the host before the host will recognize that space allocation has been enabled.

Steps

1. Take the LUN offline.

   ```
   lun modify -vserver vserver_name -volume volume_name -lun lun_name -state
   offline
   ```

2. Set the `-space-allocation` parameter to `enabled`:

   ```
   lun modify -vserver vserver_name -volume volume_name -lun lun_name -space
   -allocation enabled
   ```

3. Verify that space allocation is enabled:

   ```
   lun show -vserver vserver_name -volume volume_name -lun lun_name -fields
   space-allocation
   ```

4. Bring the LUN online:

```
lun modify -vserver vserver_name -volume volume_name -lun lun_name -state
online
```

5. On the host, rescan all disks to ensure that the change to the `-space-allocation` option is correctly discovered.

Host support for SCSI thin provisioning

To leverage the benefits of SCSI thin provisioning, it must be supported by your host. SCSI thin provisioning uses the Logical Block Provisioning feature as defined in the SCSI SBC-3 standard. Only hosts that support this standard can use SCSI thin provisioning in ONTAP.

The following hosts currently support SCSI thin provisioning when you enable space allocation:

- VMware ESX 5.0 and later
- Red Hat Enterprise Linux 6.2 and later
- Citrix XenServer 6.5 and later
- Microsoft Windows 2012
- Microsoft Windows 2016

When you enable the space allocation functionality in ONTAP, you turn on the following SCSI thin provisioning features:

- Unmapping and reporting space usage for space reclamation
- Reporting resource exhaustion errors

Simplified host management with SnapCenter

You can use SnapCenter software to simplify some of the management and data protection tasks associated with iSCSI and FC storage. SnapCenter is an optional management package for Windows and UNIX hosts.

You can use SnapCenter Software to easily create virtual disks from pools of storage that can be distributed among several storage systems and to automate storage provisioning tasks and simplify the process of creating Snapshot copies and clones from Snapshot copies consistent with host data.

See NetApp product documentation for more information on SnapCenter.

About igroups

Initiator groups (igroups) are tables of FC protocol host WWPNs or iSCSI host node names. You can define igroups and map them to LUNs to control which initiators have access to LUNs.

Typically, you want all of the host’s initiator ports or software initiators to have access to a LUN. If you are using multipathing software or have clustered hosts, each initiator port or software initiator of each clustered host needs redundant paths to the same LUN.

You can create igroups that specify which initiators have access to the LUNs either before or after you create LUNs, but you must create igroups before you can map a LUN to an igroup.

Initiator groups can have multiple initiators, and multiple igroups can have the same initiator. However, you cannot map a LUN to multiple igroups that have the same initiator. An initiator cannot be a member of igroups of differing ostypes.

Example of how igroups give LUN access

You can create multiple igroups to define which LUNs are available to your hosts. For example, if you have a host cluster, you can use igroups to ensure that specific LUNs are visible to only one host in the cluster or to all of the hosts in the cluster.

The following table illustrates how four igroups give access to the LUNs for four different hosts that are accessing the storage system. The clustered hosts (Host3 and Host4) are both members of the same igroup (group3) and can access the LUNs mapped to this igroup. The igroup named group4 contains the WWPNs of Host4 to store local information that is not intended to be seen by its partner.

Hosts with HBA WWPNs, IQNs, or EUIs	igroups	WWPNs, IQNs, EUIs added to igroups	LUNs mapped to igroups
Host1, single-path (iSCSI software initiator) iqn.1991-05.com.microsoft:host1	group1	iqn.1991-05.com.microsoft:host1	`/vol/vol2/lun1`
Host2, multipath (two HBAs) 10:00:00:00:c9:2b:6b:3c 10:00:00:00:c9:2b:02:3c	group2	10:00:00:00:c9:2b:6b:3c 10:00:00:00:c9:2b:02:3c	`/vol/vol2/lun2`
Host3, multipath, clustered with host 4 10:00:00:00:c9:2b:32:1b 10:00:00:00:c9:2b:41:02	group3	10:00:00:00:c9:2b:32:1b 10:00:00:00:c9:2b:41:02 10:00:00:00:c9:2b:51:2c 10:00:00:00:c9:2b:47:a2	`/vol/vol2/qtree1/lun3`
Host4, multipath, clustered (not visible to Host3) 10:00:00:00:c9:2b:51:2c 10:00:00:00:c9:2b:47:a2	group4	10:00:00:00:c9:2b:51:2c 10:00:00:00:c9:2b:47:a2	`/vol/vol2/qtree2/lun4` `/vol/vol2/qtree1/lun5`

Specify initiator WWPNs and iSCSI node names for an igroup

You can specify the iSCSI node names and WWPNs of the initiators when you create an igroup or you can add them later. If you choose to specify the initiator iSCSI node names and WWPNs when you create the LUN, they can be removed later, if needed.

Follow the instructions in your Host Utilities documentation to obtain WWPNs and to find the iSCSI node names associated with a specific host. For hosts running ESX software, use Virtual Storage Console.

Storage virtualization with VMware and Microsoft copy offload

Storage virtualization with VMware and Microsoft copy offload overview

VMware and Microsoft support copy offload operations to increase performance and network throughput. You must configure your system to meet the requirements of the VMware and Windows operating system environments to use their respective copy offload functions.

When using VMware and Microsoft copy offload in virtualized environments, your LUNs must be aligned. Unaligned LUNs can degrade performance.

Advantages of using a virtualized SAN environment

Creating a virtualized environment by using storage virtual machines (SVMs) and LIFs enables you to expand your SAN environment to all of the nodes in your cluster.

- Distributed management

  You can log in to any node in the SVM to administer all of the nodes in a cluster.

- Increased data access

  With MPIO and ALUA, you have access to your data through any active iSCSI or FC LIFs for the SVM.

- Controlled LUN access

  If you use SLM and portsets, you can limit which LIFs an initiator can use to access LUNs.

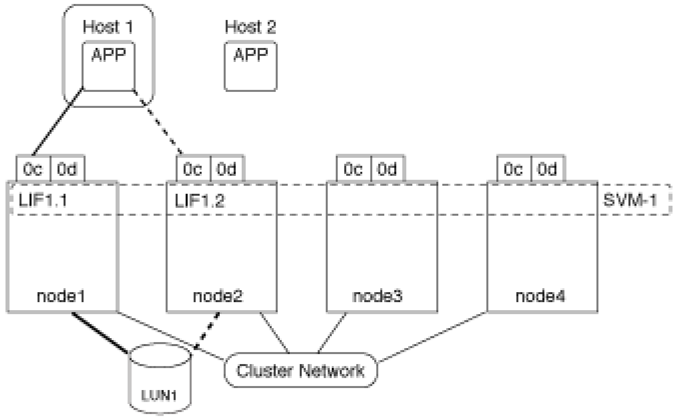
How LUN access works in a virtualized environment

In a virtualized environment, LIFs enable hosts (clients) to access LUNs through optimized and unoptimized paths.

A LIF is a logical interface that connects the SVM to a physical port. Although multiple SVMs can have multiple LIFs on the same port, a LIF belongs to one SVM. You can access LUNs through the SVMs LIFs.

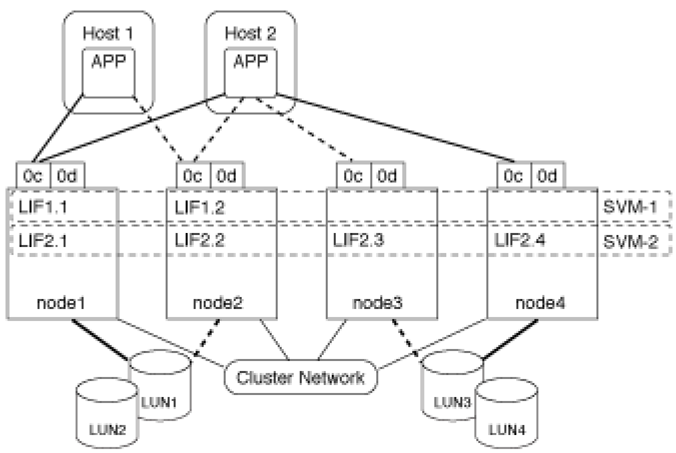
Example of LUN access with a single SVM in a cluster

In the following example, Host 1 connects to LIF1.1 and LIF1.2 in SVM-1 to access LUN1. LIF1.1 uses the physical port node1:0c and LIF1.2 uses the node2:0c. LIF1.1 and LIF1.2 belongs only to SVM-1. If a new LUN is created on node 1 or node 2, for SVM-1, then it can use these same LIFs. If a new SVM is created, then new LIFs can be created using physical ports 0c or 0d on both the nodes.



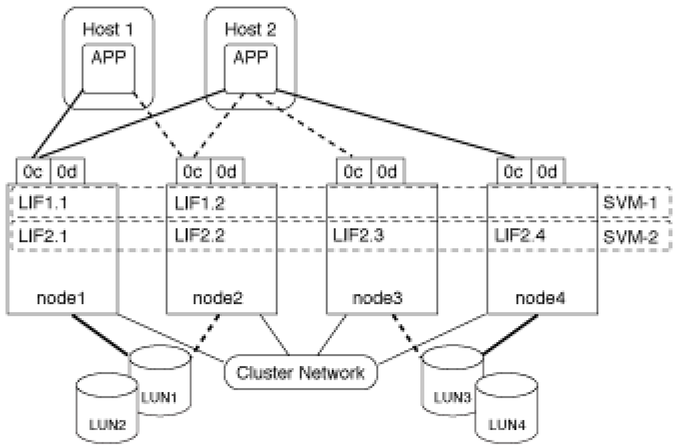
Example of LUN access with multiple SVMs in a cluster

A physical port can support multiple LIFs serving different SVMs. Because LIFs are associated with a particular SVM, the cluster nodes can send the incoming data traffic to the correct SVM. In the following example, each node from 1 through 4 has a LIF for SVM-2 using the physical port 0c on each node. Host 1 connects to LIF1.1 and LIF1.2 in SVM-1 to access LUN1. Host 2 connects to LIF2-1 and LIF2-2 in SVM-2 to access LUN2. Both SVMs are sharing the physical port 0c on the nodes 1 and 2. SVM-2 has additional LIFs that Host 2 is using to access LUNs 3 and 4. These LIFs are using physical port 0c on nodes 3 and 4. Multiple SVMs can share the physical ports on the nodes.



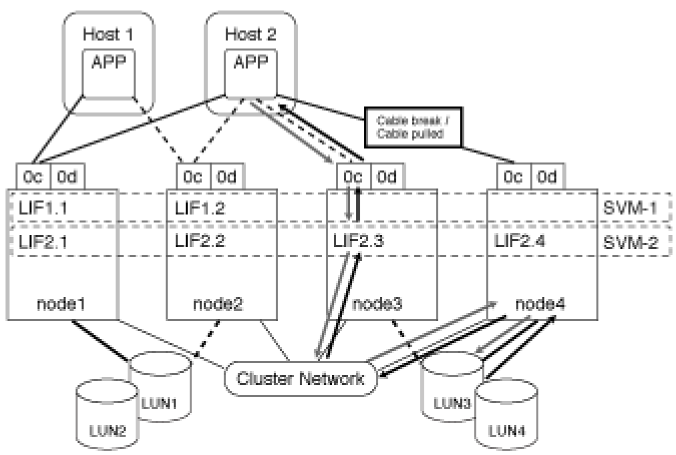
Example of an active or optimized path to a LUN from a host system

In an active or optimized path, the data traffic does not travel over the cluster network; it travels the most direct route to the LUN. The active or optimized path to LUN1 is through LIF1.1 in node1, using physical port 0c. Host 2 has two active or optimized paths, one path to node1, LIF2.1, which is sharing physical port 0c and the other path to node4, LIF2.4, which is using physical port 0c.



Example of an active or unoptimized path (indirect) path to a LUN from a host system

In an active or unoptimized path (indirect) path, the data traffic travels over the cluster network. This issue occurs only if all the active or optimized paths from a host are unavailable to handle traffic. If the path from Host 2 to SVM-2 LIF2.4 is lost, then access to LUN3 and LUN4 traverses the cluster network. Access from Host 2 uses LIF2.3 on node3. Then the traffic enters the cluster network switch and backs up to node4 for access to the LUN3 and LUN4. It will then traverse back over the cluster network switch and then back out through LIF2.3 to Host 2. This active or unoptimized path is used until the path to LIF2.4 is restored or a new LIF is established for SVM-2 on another physical port on node 4.



Considerations for LIFs in cluster SAN environments

You must be aware of certain LIF considerations in a SAN environment.

- Initiators must use Multipath I/O (MPIO) and asymmetric logical unit access(ALUA) for failover capability for clusters in a SAN iSCSI or FC environment because SAN does not support automatic failover for LIFs.
- At least one SAN LIF of the appropriate protocol must be configured on each node that hosts a mapped LUN and the node's HA partner.

  You can configure two LIFs per node, one for each fabric being used with FC and to separate Ethernet networks for iSCSI.
- Some options are not applicable for iSCSI or FC.

  For example, you cannot use IP addresses with FC.

Improve VMware VAAI performance for ESX hosts

ONTAP supports certain VMware vStorage APIs for Array Integration (VAAI) features when the ESX host is running ESX 4.1 or later. These features help offload operations from the ESX host to the storage system and increase the network throughput. The ESX host enables the features automatically in the correct environment.

The VAAI feature supports the following SCSI commands:

- `EXTENDED_COPY`

  This feature enables the host to initiate the transfer of data between the LUNs or within a LUN without involving the host in the data transfer. This results in saving ESX CPU cycles and increasing the network throughput. The extended copy feature, also known as "copy offload," is used in scenarios such as cloning a virtual machine. When invoked by the ESX host, the copy offload feature copies the data within the storage system rather than going through the host network. Copy offload transfers data in the following ways:

  - Within a LUN
  - Between LUNs within a volume
  - Between LUNs on different volumes within a storage virtual machine (SVM)
  - Between LUNs on different SVMs within a cluster If this feature cannot be invoked, the ESX host automatically uses the standard READ and WRITE commands for the copy operation.

- `WRITE_SAME`

  This feature offloads the work of writing a repeated pattern, such as all zeros, to a storage array. The ESX host uses this feature in operations such as zero-filling a file.

- `COMPARE_AND_WRITE`

  This feature bypasses certain file access concurrency limits, which speeds up operations such as booting up virtual machines.

Requirements for using the VAAI environment

The VAAI features are part of the ESX operating system and are automatically invoked by the ESX host when you have set up the correct environment.

The environment requirements are as follows:

- The ESX host must be running ESX 4.1 or later.
- The NetApp storage system that is hosting the VMware datastore must be running ONTAP.
- (Copy offload only) The source and the destination of the VMware copy operation must be hosted on the same storage system within the same cluster.

> ⓘ The copy offload feature currently does not support copying data between VMware datastores that are hosted on different storage systems.

Determine if VAAI features are supported by ESX

To confirm whether the ESX operating system supports the VAAI features, you can check the vSphere Client or use any other means of accessing the host. ONTAP supports the SCSI commands by default.

You can check your ESX host advanced settings to determine whether VAAI features are enabled. The table indicates which SCSI commands correspond to ESX control names.

SCSI command	ESX control name (VAAI feature)
EXTENDED_COPY	`HardwareAcceleratedMove`
WRITE_SAME	`HardwareAcceleratedInit`
COMPARE_AND_WRITE	`HardwareAcceleratedLocking`

Microsoft Offloaded Data Transfer (ODX)

Microsoft Offloaded Data Transfer (ODX) overview

Microsoft Offloaded Data Transfer (ODX), also known as *copy offload*, enables direct data transfers within a storage device or between compatible storage devices without transferring the data through the host computer.

ONTAP supports ODX for both the SMB and SAN protocols.

In non-ODX file transfers, the data is read from the source and is transferred across the network to the host. The host transfers the data back over the network to the destination. In ODX file transfer, the data is copied directly from the source to the destination without passing through the host.

Because ODX offloaded copies are performed directly between the source and destination, significant performance benefits are realized, including faster copy time, reduced utilization of CPU and memory on the client, and reduced network I/O bandwidth utilization.

For SAN environments, ODX is only available when it is supported by both the host and the storage system. Client computers that support ODX and have ODX enabled automatically and transparently use offloaded file transfer when moving or copying files. ODX is used regardless of whether you drag-and-drop files through Windows Explorer or use command-line file copy commands, or whether a client application initiates file copy requests.

Requirements for using ODX

If you plan to use ODX for copy offloads, you need to be familiar with volume support considerations, system requirements, and software capability requirements.

To use ODX, your system must have the following:

- ONTAP

  ODX is automatically enabled in supported versions of ONTAP.

* Minimum source volume of 2 GB

  For optimal performance, the source volume should be greater than 260 GB.

* Deduplication

  ODX uses deduplication as part of the copy process. If you do not want deduplication on your SVM, you should disable ODX on that SVM.

* ODX support on the Windows client

  ODX is supported in Windows Server 2012 or later and in Windows 8 or later. The Interoperability Matrix contains the latest information about supported Windows clients.

  NetApp Interoperability Matrix Tool

* Copy application support for ODX

  The application that performs the data transfer must support ODX. Application operations that support ODX include the following:

  * Hyper-V management operations, such as creating and converting virtual hard disks (VHDs), managing Snapshot copies, and copying files between virtual machines
  * Windows Explorer operations
  * Windows PowerShell copy commands
  * Windows command prompt copy commands The Microsoft TechNet Library contains more information about supported ODX applications on Windows servers and clients.

* If you use compressed volumes, the compression group size must be 8K.

  32K compression group size is not supported.

ODX does not work with the following volume types:

* Source volumes with capacities of less than 2 GB
* Read-only volumes
* FlexCache volumes
* Semi-thick provisioned volumes

Use cases for ODX

You should be aware of the use cases for using ODX on SVMs so that you can determine under what circumstances ODX provides you with performance benefits.

Windows servers and clients that support ODX use copy offload as the default way of copying data across remote servers. If the Windows server or client does not support ODX or the ODX copy offload fails at any point, the copy or move operation falls back to traditional reads and writes for the copy or move operation.

The following use cases support using ODX copies and moves:

* Intra-volume

The source and destination files or LUNs are within the same volume.

- Inter-volume, same node, same SVM

  The source and destination files or LUNs are on different volumes that are located on the same node. The data is owned by the same SVM.

- Inter-volume, different nodes, same SVM

  The source and destination files or LUNs are on different volumes that are located on different nodes. The data is owned by the same SVM.

- Inter-SVM, same node

  The source and destination file or LUNs are on different volumes that are located on the same node. The data is owned by different SVMs.

- Inter-SVM, different nodes

  The source and destination file or LUNs are on different volumes that are located on different nodes. The data is owned by different SVMs.

- Inter-cluster

  The source and destination LUNs are on different volumes that are located on different nodes across clusters. This is only supported for SAN and does not work for SMB.

There are some additional special use cases:

- With the ONTAP ODX implementation, you can use ODX to copy files between SMB shares and FC or iSCSI attached virtual drives.

  You can use Windows Explorer, the Windows CLI or PowerShell, Hyper-V, or other applications that support ODX to copy or move files seamlessly using ODX copy offload between SMB shares and connected LUNs, provided that the SMB shares and LUNs are on the same cluster.

- Hyper-V provides some additional use cases for ODX copy offload:

  - You can use ODX copy offload pass-through with Hyper-V to copy data within or across virtual hard disk (VHD) files or to copy data between mapped SMB shares and connected iSCSI LUNs within the same cluster.

    This allows copies from guest operating systems to pass through to the underlying storage.

  - When creating fixed-sized VHDs, ODX is used for initializing the disk with zeros, using a well-known zeroed token.
  - ODX copy offload is used for virtual machine storage migration if the source and destination storage is on the same cluster.

  > To take advantage of the use cases for ODX copy offload pass-through with Hyper-V, the guest operating system must support ODX and the guest operating system’s disks must be SCSI disks backed by storage (either SMB or SAN) that supports ODX. IDE disks on the guest operating system do not support ODX pass-through.

Special system file requirements

You can delete ODX files found in qtrees. You must not remove or modify any other ODX system files unless you are told by technical support to do so.

When using the ODX feature, there are ODX system files that exist in every volume of the system. These files enable point-in-time representation of data used during the ODX transfer. The following system files are in the root level of each volume that contains LUNs or files to which data was offloaded:

- `.copy-offload` (a hidden directory)
- `.tokens` (file under the hidden `.copy-offload` directory)

You can use the `copy-offload delete-tokens -path dir_path -node` *`node_name`* command to delete a qtree containing an ODX file.