SAN configuration reference

ONTAP 9

NetApp
February 27, 2023

Table of Contents

SAN configuration reference

SAN configuration reference

The following sections describe supported FC-NVMe, FC, iSCSI, and FCoE topologies for connecting host computers to nodes, and list supported limits for SAN components.

You should use this information in conjunction with basic SAN configuration documentation:

- SAN administration overview

Considerations for iSCSI configurations

Considerations for iSCSI configurations overview

You should consider several things when setting up your iSCSI configuration.

- You can set up your iSCSI configuration with single nodes or with HA pairs.

  Direct connect or the use of Ethernet switches is supported for connectivity. You must create LIFs for both types of connectivity

- You should configure one management LIF for every storage virtual machine (SVM) supporting SAN.
- Selective LUN mapping (SLM) limits the paths that are being utilized in accessing the LUNs owned by an HA pair.

  This is the default behavior for LUNs created with ONTAP releases.

- HA pairs are defined as the reporting nodes for the Active/Optimized and the Active/Unoptimized paths that will be used by the host in accessing the LUNs through ALUA.
- It is recommended that all SVMs in ISCSI configurations have a minimum of two LIF's per node in separate Ethernet networks for redundancy and MPIO across multiple paths.
- You need to create one or more iSCSI paths from each node in an HA pair, using logical interfaces (LIFs) to allow access to LUNs that are serviced by the HA pair.

  If a node fails, LIFs do not migrate or assume the IP addresses of the failed partner node. Instead, the MPIO software, using ALUA on the host, is responsible for selecting the appropriate paths for LUN access through LIFs.

- VLANs offer specific benefits, such as increased security and improved network reliability that you might want to leverage in iSCSI.

Ways to configure iSCSI SAN hosts with single nodes

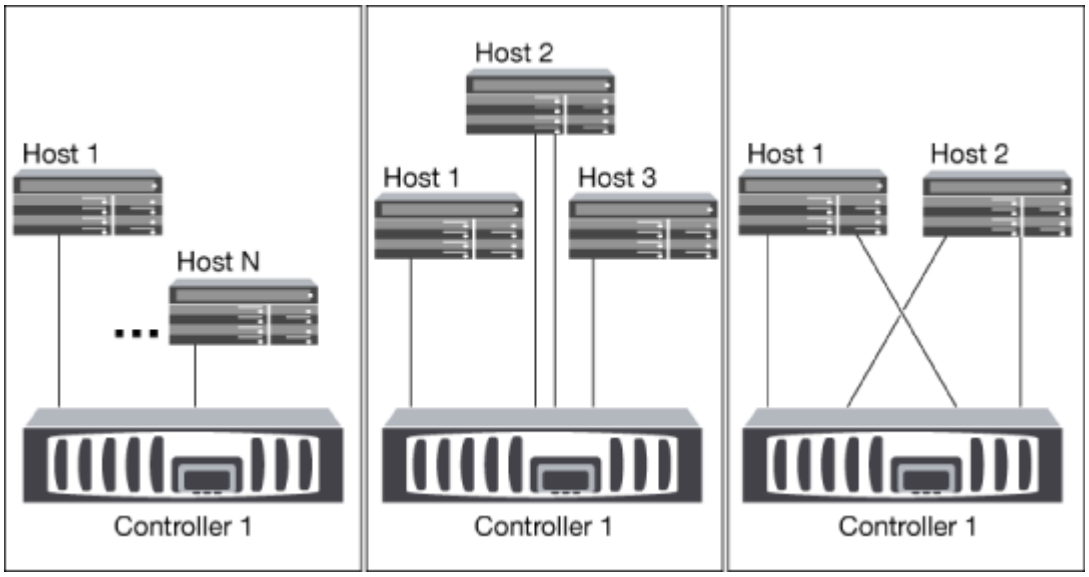
You can configure the iSCSI SAN hosts to connect directly to a single node or by using either one or multiple IP switches. You should determine whether you want a single-switch configuration that is not completely redundant or a multi-switch configuration that is completely redundant.

You can configure iSCSI SAN hosts in a direct-attached, single-switch, or multi-switch environment. If there are multiple hosts connecting to the node, each host can be configured with a different operating system. For single and multi-network configurations, the node can have multiple iSCSI connections to the switch, but multipathing software that supports ALUA is required.

If there are multiple paths from the host to the controller, then ALUA must be enabled on the host.

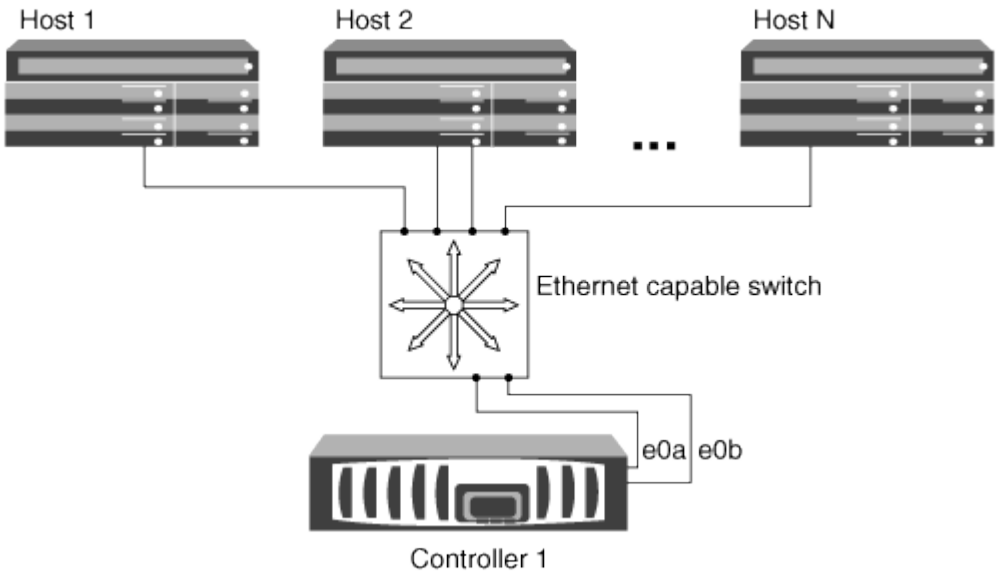
Direct-attached single-node configurations

In direct-attached configurations, one or more hosts are directly connected to the node.



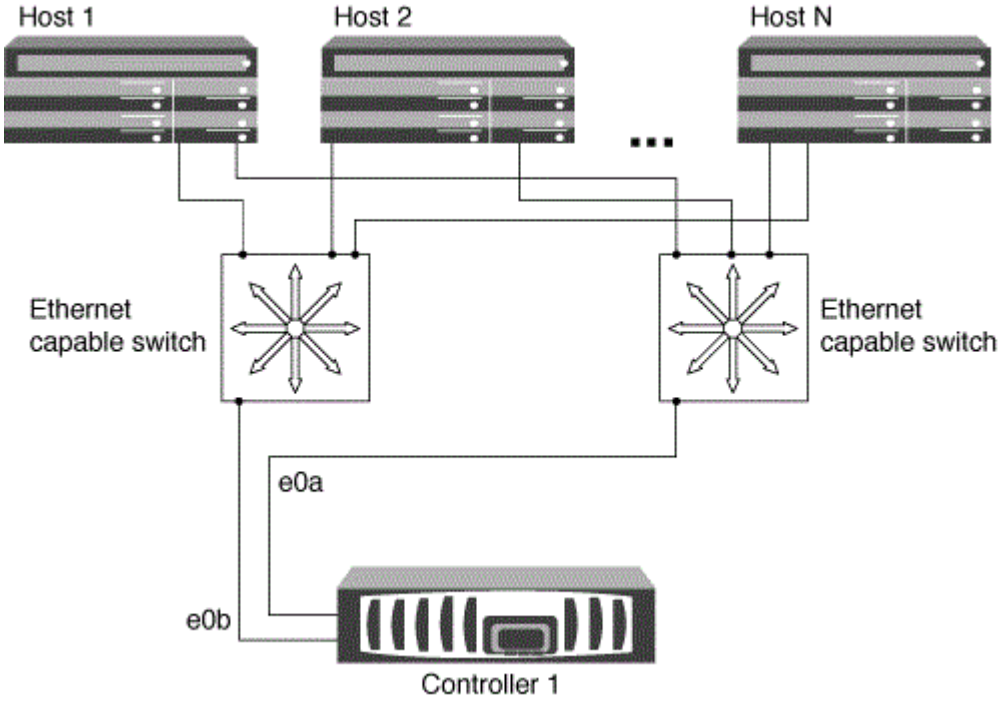
Single-network single-node configurations

In single-network single-node configurations, one switch connects a single node to one or more hosts. Because there is a single switch, this configuration is not fully redundant.



Multi-network single-node configurations

In multi-network single-node configurations, two or more switches connect a single node to one or more hosts. Because there are multiple switches, this configuration is fully redundant.



Ways to configure iSCSI SAN hosts with HA pairs

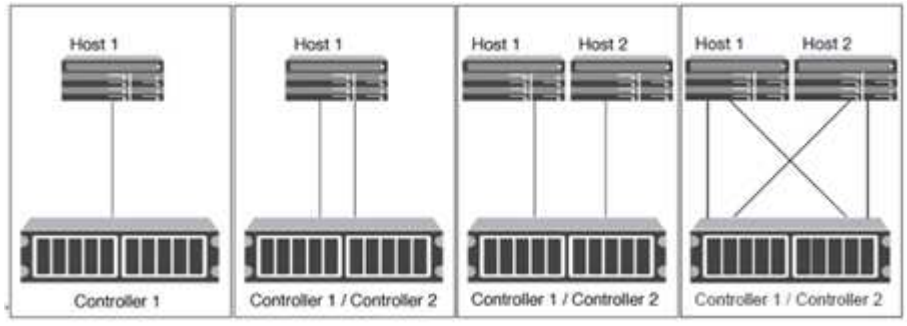
You can configure the iSCSI SAN hosts to connect to dual-node or multi-node configurations by using either one or multiple IP switches. You should determine whether you want a single-switch configuration that is not completely redundant or a multi-switch configuration that is completely redundant.

You can configure iSCSI SAN hosts with single controllers and HA pairs on direct-attached, single-network, or multi-network environments. HA pairs can have multiple iSCSI connections to each switch, but multipathing software that supports ALUA is required on each host. If there are multiple hosts, you can configure each host with a different operating system by checking the NetAppInteroperability Matrix Tool.

NetApp Interoperability Matrix Tool

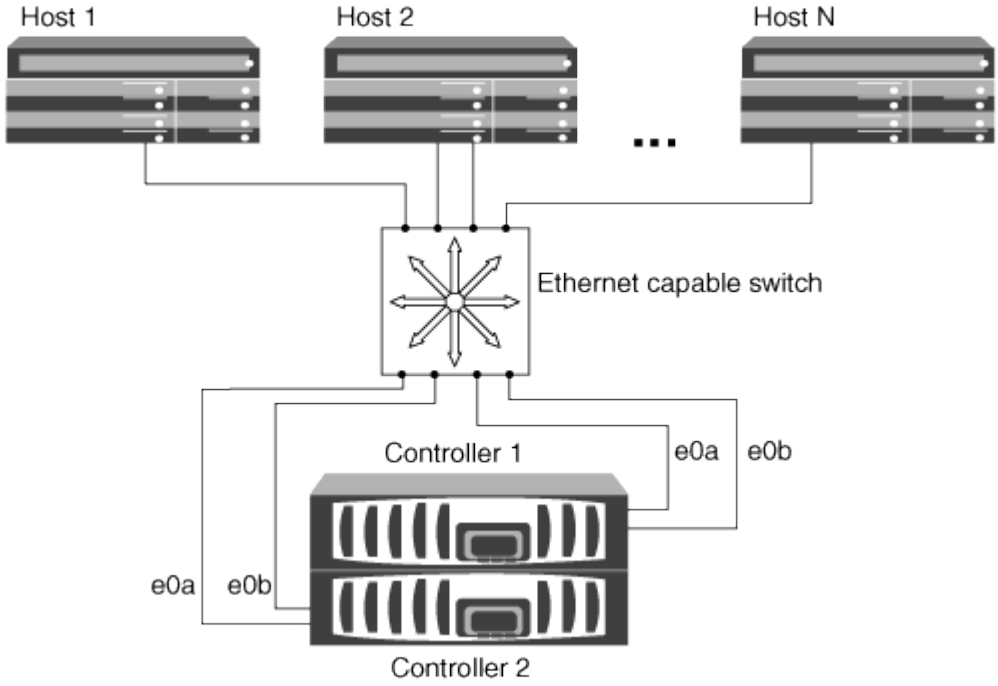
Direct-attachment

In a direct-attached configuration, one or more hosts are directly connected to the controllers.



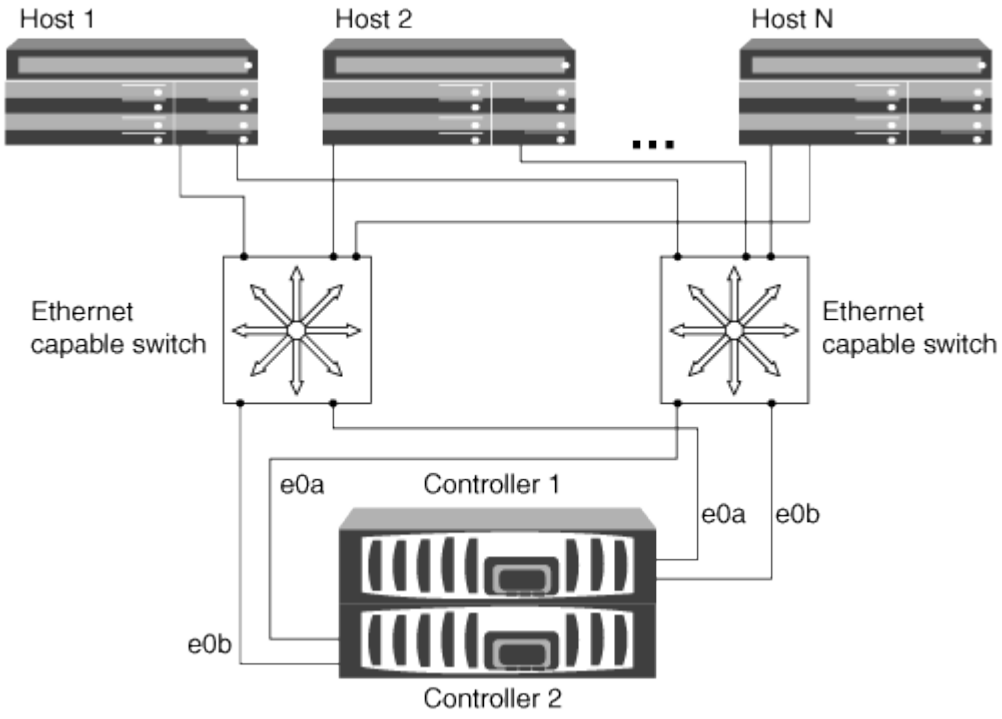
Single-network HA pairs

In single-network HA pair configurations, one switch connects the HA pair to one or more hosts. Because there is a single switch, this configuration is not fully redundant.



Multi-network HA pairs

In multi-network HA pair configurations, two or more switches connect the HA pair to one or more hosts. Because there are multiple switches, this configuration is fully redundant.



Benefits of using VLANs in iSCSI configurations

A VLAN consists of a group of switch ports grouped together into a broadcast domain. A VLAN can be on a single switch or it can span multiple switch chassis. Static and dynamic VLANs enable you to increase security, isolate problems, and limit available paths within your IP network infrastructure.

When you implement VLANs in large IP network infrastructures, you derive the following benefits:

- Increased security.

  VLANs enable you to leverage existing infrastructure while still providing enhanced security because they limit access between different nodes of an Ethernet network or an IP SAN.

- Improved Ethernet network and IP SAN reliability by isolating problems.
- Reduction of problem resolution time by limiting the problem space.
- Reduction of the number of available paths to a particular iSCSI target port.
- Reduction of the maximum number of paths used by a host.

  Having too many paths slows reconnect times. If a host does not have a multipathing solution, you can use VLANs to allow only one path.

Dynamic VLANs

Dynamic VLANs are MAC address-based. You can define a VLAN by specifying the MAC address of the members you want to include.

Dynamic VLANs provide flexibility and do not require mapping to the physical ports where the device is physically connected to the switch. You can move a cable from one port to another without reconfiguring the VLAN.

Static VLANs

Static VLANs are port-based. The switch and switch port are used to define the VLAN and its members.

Static VLANs offer improved security because it is not possible to breach VLANs using media access control (MAC) spoofing. However, if someone has physical access to the switch, replacing a cable and reconfiguring the network address can allow access.

In some environments, it is easier to create and manage static VLANs than dynamic VLANs. This is because static VLANs require only the switch and port identifier to be specified, instead of the 48-bit MAC address. In addition, you can label switch port ranges with the VLAN identifier.

Considerations for FC-NVMe configurations

Beginning with ONTAP 9.4, the non-volatile memory express (NVMe) protocol is available for SAN environments. FC-NVMe allows you to run NVMe over an existing FC network with an AFF system. FC-NVMe uses the same physical setup and zoning practice as traditional FC networks but allows for greater bandwidth, increased IOPs and reduced latency than FC-SCSI.

Supported configurations:

- NVMe is supported on AFF platforms that have 32G FC ports.
- You can set up your FC-NVMe configuration with single nodes or HA pairs using a single fabric or multifabric.
- NVMe is supported on 4-node clusters or smaller.
- NVMe can be the only data protocol on the storage virtual machine (SVM).
- Up to 8 NVMe SVMs are supported per cluster.
- FC-NVMe can be the only data protocol on data LIFs.
- LUNs and namespaces cannot be mixed on the same volume.
- You should configure one management LIF for every SVM supporting SAN.
- The use of heterogeneous FC switch fabrics is not supported, except in the case of embedded blade switches.

  Specific exceptions are listed on the NetApp Interoperability Matrix Tool.

- Cascade, partial mesh, full mesh, core-edge, and director fabrics are all industry-standard methods of connecting FC switches to a fabric, and all are supported.

  A fabric can consist of one or multiple switches, and the storage controllers can be connected to multiple switches.

Functionality enhancements:

This functionality is supported…	**Starting with…**
volume move with mapped namespaces	ONTAP 9.6
Namespaces support 512 byte blocks and 4096 byte blocks. 4096 is the default value. 512 should only be used if the host operating system does not support 4096 byte blocks.	ONTAP 9.6
Multipath HA pair failover/giveback	ONTAP 9.5

The following applies only to nodes running ONTAP 9.4:

- NVMe LIFs and namespaces must be hosted on the same node.
- The NVMe service must be created before the NVMe LIF is created.

The following ONTAP features are not supported by NVMe configurations:

- NVMe namespace move
- NVMe namespaces (Copy on Demand)
- Creating namespaces on a volume transitioned from Data ONTAP operating in 7-mode.
- Sync
- Virtual Storage Console

See the NetApp Hardware Universe for a complete list of NVMe limits.

Related information

How to configure and Connect SUSE Enterprise Linux to ONTAP NVMe/FC namespaces

Licensing information for NVMe protocol on ONTAP

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

Considerations for FC configurations

Considerations for FC configurations overview

You should be aware of several things when setting up your FC configuration.

- You can set up your FC configuration with single nodes or HA pairs using a single fabric or multifabric.
- You should configure two FC data LIFs per node.

  This creates redundancy and protects against loss of data access.

- You should configure one management LIF for every storage virtual machine (SVM) supporting SAN.
- Multiple hosts, using different operating systems, such as Windows, Linux, or UNIX, can access the storage solution at the same time.

  Hosts require that a supported multipathing solution be installed and configured. Supported operating systems and multipathing solutions can be verified on the Interoperability Matrix.

- ONTAP supports single, dual, or multiple node solutions that are connected to multiple physically independent storage fabrics; a minimum of two are recommended for SAN solutions.

  This provides redundancy at the fabric and storage system layers. Redundancy is particularly important because these layers typically support many hosts.

- The use of heterogeneous FC switch fabrics is not supported, except in the case of embedded blade switches.

  Specific exceptions are listed on the Interoperability Matrix.

- Cascade, partial mesh, full mesh, core-edge, and director fabrics are all industry-standard methods of connecting FC switches to a fabric, and all are supported.

  A fabric can consist of one or multiple switches, and the storage controllers can be connected to multiple switches.

Related information

NetApp Interoperability Matrix Tool

Ways to configure FC and FC-NVMe SAN hosts with single nodes

You can configure FC and FC-NVMe SAN hosts with single nodes through one or more fabrics. N-Port ID Virtualization (NPIV) is required and must be enabled on all FC

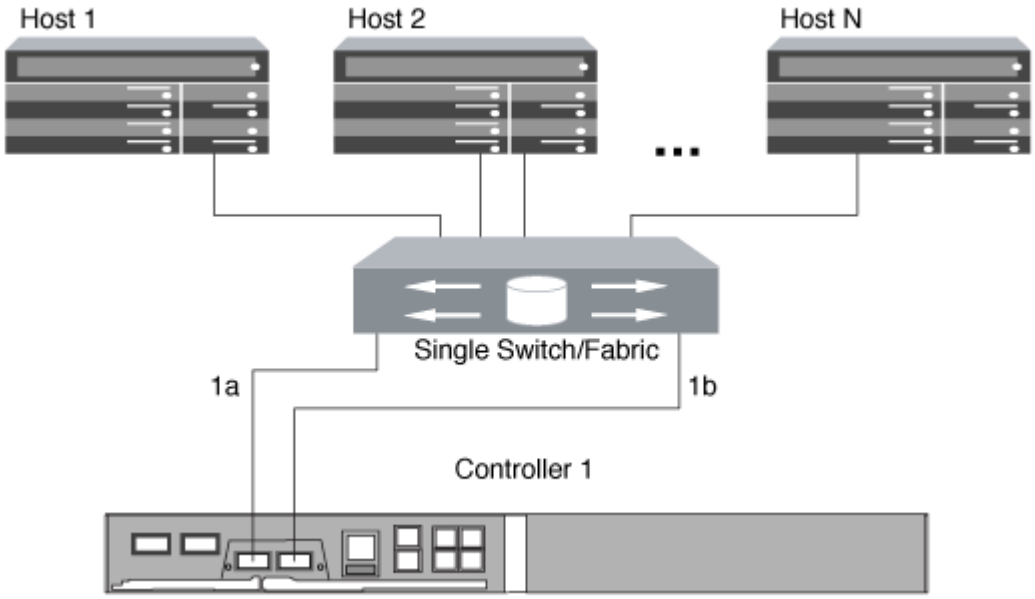
switches in the fabric. You cannot directly attach FC or FC-NMVE SAN hosts to single nodes without using an FC switch.

You can configure FC or FC-NVMe SAN hosts with single nodes through a single fabric or multifabrics. The FC target ports (0a, 0c, 0b, 0d) in the illustrations are examples. The actual port numbers vary depending on the model of your storage node and whether you are using expansion adapters.

Single-fabric single-node configurations

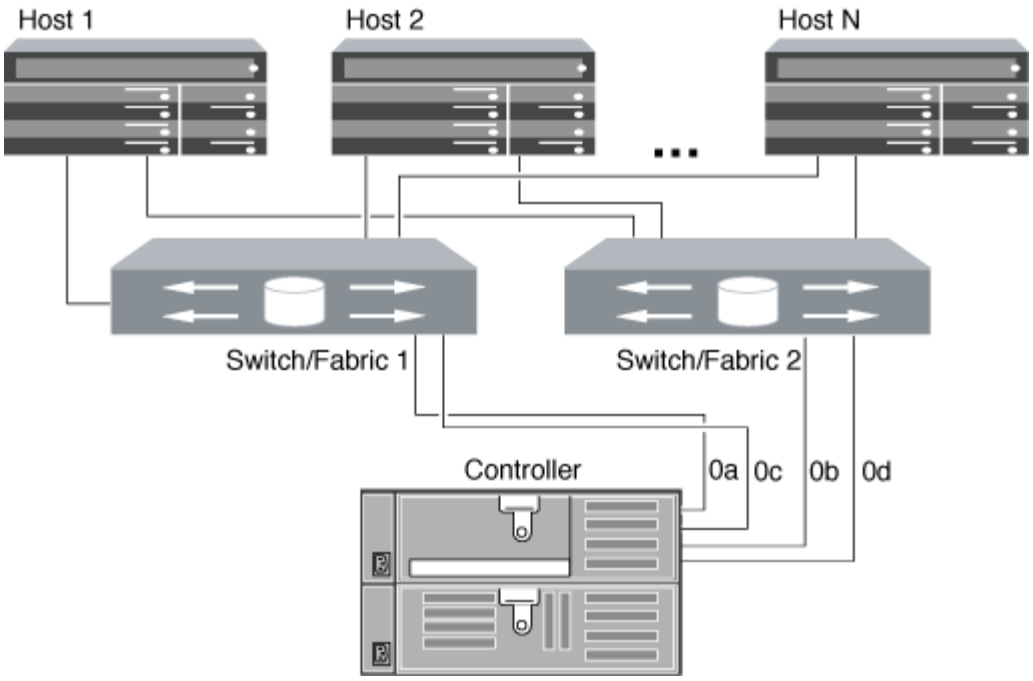
In single-fabric single-node configurations, there is one switch connecting a single node to one or more hosts. Because there is a single switch, this configuration is not fully redundant. All hardware platforms that support FC and FC-NVMe support single-fabric single-node configurations. However, the FAS2240 platform requires the X1150A-R6 expansion adapter to support a single-fabric single-node configuration.

The following figure shows a FAS2240 single-fabric single-node configuration. It shows the storage controllers side by side, which is how they are mounted in the FAS2240-2. For the FAS2240-4, the controllers are mounted one above the other. There is no difference in the SAN configuration for the two models.



Multifabric single-node configurations

In multifabric single-node configurations, there are two or more switches connecting a single node to one or more hosts. For simplicity, the following figure shows a multifabric single-node configuration with only two fabrics, but you can have two or more fabrics in any multifabric configuration. In this figure, the storage controller is mounted in the top chassis and the bottom chassis can be empty or can have an IOMX module, as it does in this example.



Related information

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

Ways to configure FC & FC-NVMe SAN hosts with HA pairs

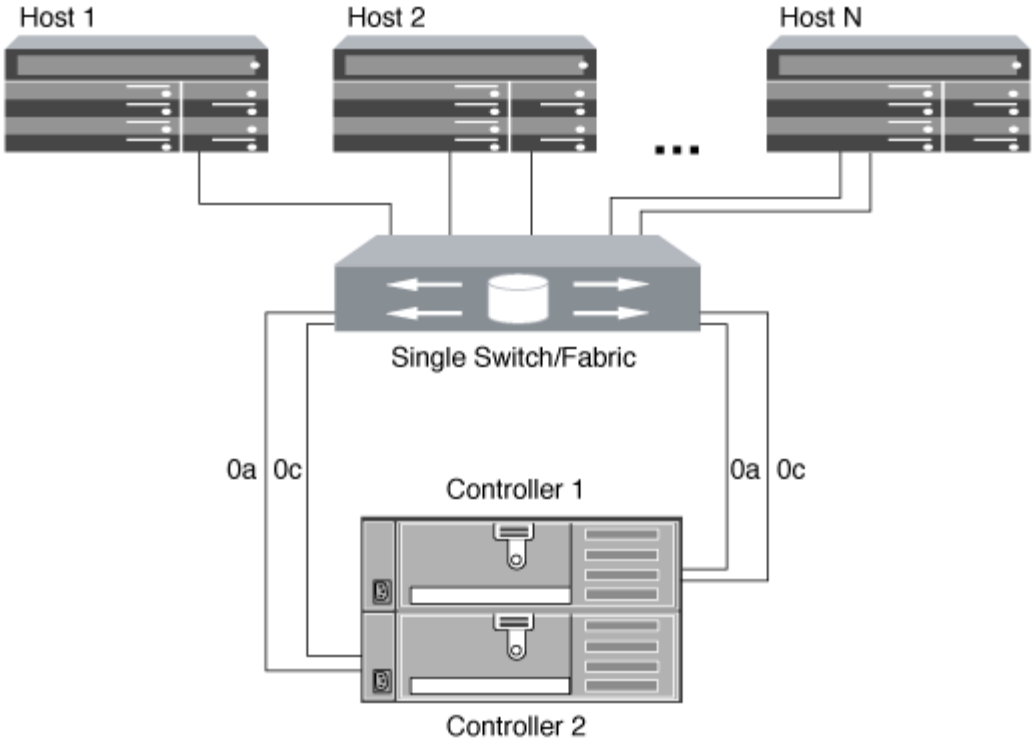
You can configure FC and FC-NVMe SAN hosts to connect to HA pairs through one or more fabrics. You cannot directly attach FC or FC-NVMe SAN hosts to HA pairs without using a switch.

You can configure FC and FC-NVMe SAN hosts with single fabric HA pairs or with multifabric HA pairs. The FC target port numbers (0a, 0c, 0d, 1a, 1b) in the illustrations are examples. The actual port numbers vary depending on the model of your storage node and whether you are using expansion adapters.

Single-fabric HA pairs

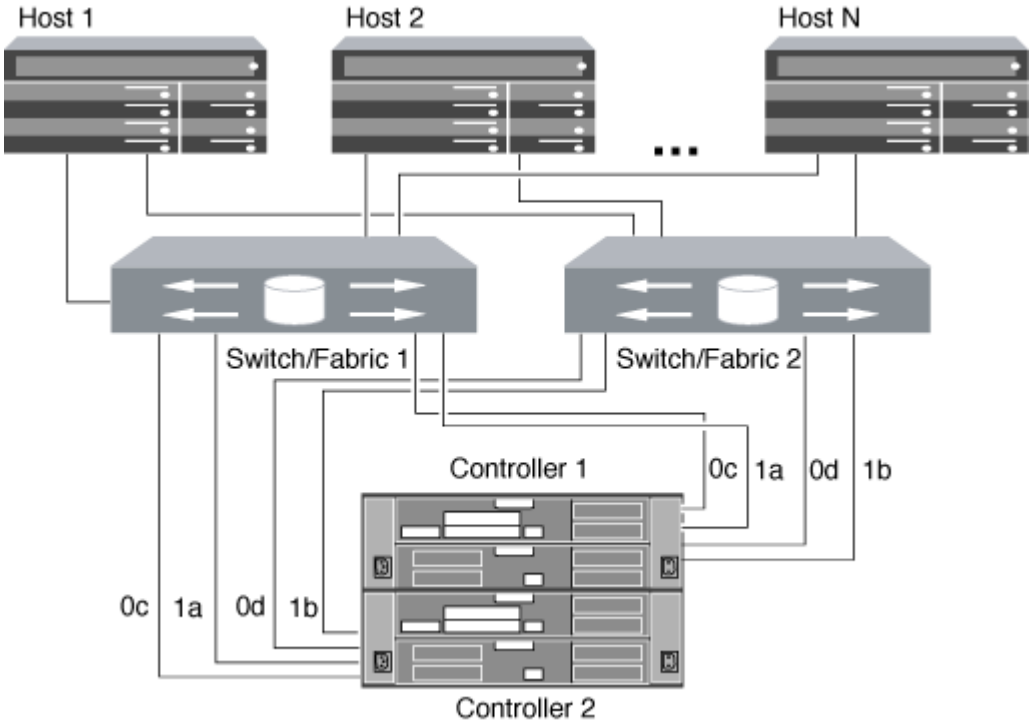
In single-fabric HA pair configurations, there is one fabric connecting both controllers in the HA pair to one or more hosts. Because the hosts and controllers are connected through a single switch, single-fabric HA pairs are not fully redundant.

All platforms that support FC configurations support single-fabric HA pair configurations, except the FAS2240 platform. The FAS2240 platform only supports single-fabric single-node configurations.



Multifabric HA pairs

In multifabric HA pairs, there are two or more switches connecting HA pairs to one or more hosts. For simplicity, the following multifabric HA pair figure shows only two fabrics, but you can have two or more fabrics in any multifabric configuration:



FC switch configuration best practices

For best performance, you should consider certain best practices when configuring your

FC switch.

A fixed link speed setting is the best practice for FC switch configurations, especially for large fabrics because it provides the best performance for fabric rebuilds and can significantly save time. Although autonegotiation provides the greatest flexibility, FC switch configuration does not always perform as expected, and it adds time to the overall fabric-build sequence.

All of the switches that are connected to the fabric must support N_Port ID virtualization (NPIV) and must have NPIV enabled. ONTAP uses NPIV to present FC targets to a fabric.

For details about which environments are supported, see the NetApp Interoperability Matrix Tool.

For FC and iSCSI best practices, see Best Practices for Scalable SAN - ONTAP 9.

Supported number of FC hop counts

The maximum supported FC hop count between a host and storage system depends on the switch supplier and storage system support for FC configurations.

The hop count is defined as the number of switches in the path between the initiator (host) and target (storage system). Cisco also refers to this value as the *diameter of the SAN fabric*.

Switch supplier	Supported hop count
Brocade	7 for FC5 for FCoE
Cisco	7 for FCUp to 3 of the switches can be FCoE switches.

Related information

NetApp Downloads: Brocade Scalability Matrix Documents

NetApp Downloads: Cisco Scalability Matrix Documents

FC target port supported speeds

FC target ports can be configured to run at different speeds. You should set the target port speed to match the speed of the device to which it connects. All target ports used by a given host should be set to the same speed.

FC target ports can be used for FC-NVMe configurations in the exact same way they are used for FC configurations.

You should set the target port speed to match the speed of the device to which it connects instead of using autonegotiation. A port that is set to autonegotiation can take longer to reconnect after a takeover/giveback or other interruption.

You can configure onboard ports and expansion adapters to run at the following speeds. Each controller and expansion adapter port can be configured individually for different speeds as needed.

4 Gb ports	8 Gb ports	16 Gb ports	32 Gb ports
• 4 Gb • 2 Gb • 1 Gb	• 8 Gb • 4 Gb • 2 Gb	• 16 Gb • 8 Gb • 4 Gb	• 32 Gb • 16 Gb • 8 Gb

UTA2 ports can use an 8 Gb SFP+ adapter to support 8, 4, and 2 Gb speeds, if required.

FC Target port configuration recommendations

For best performance and highest availability, you should use the recommended FC target port configuration.

The following table shows the preferred port usage order for onboard FC and FC-NVMe target ports. For expansion adapters, the FC ports should be spread so that they do not use the same ASIC for connectivity. The preferred slot order is listed in NetApp Hardware Universe for the version of ONTAP software used by your controller.

FC-NVMe is supported on the following models:

- AFF A300

  The AFF A300 onboard ports do not support FC-NVMe.

- AFF A700
- AFF A700s
- AFF A800

The FAS22xx and FAS2520 systems do not have onboard FC ports and do not support add-on adapters.

Controller	Port pairs with shared ASIC	Number of target ports: Preferred ports
FAS9000, AFF A700, AFF A700s and AFF A800	None	All data ports are on expansion adapters. See NetApp Hardware Universe for more information.
8080, 8060 and 8040	0e+0f 0g+0h	1: 0e 2: 0e, 0g 3: 0e, 0g, 0h 4: 0e, 0g, 0f, 0h

Controller	Port pairs with shared ASIC	Number of target ports: Preferred ports
FAS8200 and AFF A300	0g+0h	1: 0g 2: 0g, 0h
8020	0c+0d	1: 0c 2: 0c, 0d
62xx	0a+0b 0c+0d	1: 0a 2: 0a, 0c 3: 0a, 0c, 0b 4: 0a, 0c, 0b, 0d
32xx	0c+0d	1: 0c 2: 0c, 0d
FAS2554, FAS2552, FAS2600 series,FAS2720,FAS2750, AFF A200 and AFF A220	0c+0d 0e+0f	1: 0c 2: 0c, 0e 3: 0c, 0e, 0d 4: 0c, 0e, 0d, 0f

Manage systems with FC adapters

Managing systems with FC adapters overview

Commands are available to manage onboard FC adapters and FC adapter cards. These commands can be used to configure the adapter mode, display adapter information, and change the speed.

Most storage systems have onboard FC adapters that can be configured as initiators or targets. You can also use FC adapter cards configured as initiators or targets. Initiators connect to back-end disk shelves, and possibly foreign storage arrays (FlexArray). Targets connect only to FC switches. Both the FC target HBA ports and the switch port speed should be set to the same value and should not be set to auto.

Commands for managing FC adapters

You can use FC commands to manage FC target adapters, FC initiator adapters, and onboard FC adapters for your storage controller. The same commands are used to manage FC adapters for the FC protocol and the FC-NVMe protocol.

FC initiator adapter commands work only at the node level. You must use the `run -node node_name` command before you can use the FC initiator adapter commands.

Commands for managing FC target adapters

If you want to…	Use this command…
Display FC adapter information on a node	`network fcp adapter show`
Modify FC target adapter parameters	`network fcp adapter modify`
Display FC protocol traffic information	`run -node node_name sysstat -f`
Display how long the FC protocol has been running	`run -node node_name uptime`
Display adapter configuration and status	`run -node node_name sysconfig -v adapter`
Verify which expansion cards are installed and whether there are any configuration errors	`run -node node_name sysconfig -ac`
View a man page for a command	`man command_name`

Commands for managing FC initiator adapters

If you want to…	Use this command…
Display information for all initiators and their adapters in a node	`run -node node_name storage show adapter`
Display adapter configuration and status	`run -node node_name sysconfig -v adapter`
Verify which expansion cards are installed and whether there are any configuration errors	`run -node node_name sysconfig -ac`

Commands for managing onboard FC adapters

If you want to…	Use this command…
Display the status of the onboard FC ports	`system node hardware unified-connect show`

Configure FC adapters for initiator mode

You can configure individual FC ports of onboard adapters and certain FC adapter cards

for initiator mode. Initiator mode is used to connect the ports to tape drives, tape libraries, or third-party storage with FlexArray Virtualization or Foreign LUN Import (FLI).

What you'll need

- LIFs on the adapter must be removed from any port sets of which they are members.
- All LIF's from every storage virtual machine (SVM) using the physical port to be modified must be migrated or destroyed before changing the personality of the physical port from target to initiator.

About this task

Each onboard FC port can be individually configured as an initiator or a target. Ports on certain FC adapters can also be individually configured as either a target port or an initiator port, just like the onboard FC ports. A list of adapters that can be configured for target mode is available in NetApp Hardware Universe.

NVMe/FC does support initiator mode.

Steps

1. Remove all LIFs from the adapter:

   ```
   network interface delete -vserver SVM_name -lif lif_name,lif_name
   ```

2. Take your adapter offline:

   ```
   network fcp adapter modify -node node_name -adapter adapter_port -status-admin down
   ```

   If the adapter does not go offline, you can also remove the cable from the appropriate adapter port on the system.

3. Change the adapter from target to initiator:

   ```
   system hardware unified-connect modify -t initiator adapter_port
   ```

4. Reboot the node hosting the adapter you changed.
5. Verify that the FC ports are configured in the correct state for your configuration:

   ```
   system hardware unified-connect show
   ```

6. Bring the adapter back online:

   ```
   node run -node node_name storage enable adapter adapter_port
   ```

Configure FC adapters for target mode

You can configure individual FC ports of onboard adapters and certain FC adapter cards for target mode. Target mode is used to connect the ports to FC initiators.

About this task

Each onboard FC port can be individually configured as an initiator or a target. Ports on certain FC adapters can also be individually configured as either a target port or an initiator port, just like the onboard FC ports. A list of adapters that can be configured for target mode is available in the NetApp Hardware Universe.

The same steps are used when configuring FC adapters for the FC protocol and the FC-NVMe protocol. However, only certain FC adapters support FC-NVMe. See the NetApp Hardware Universe for a list of adapters that support the FC-NVMe protocol.

Steps

1. Take the adapter offline:

   `node run -node node_name storage disable adapter adapter_name`

   If the adapter does not go offline, you can also remove the cable from the appropriate adapter port on the system.

2. Change the adapter from initiator to target:

   `system node hardware unified-connect modify -t target -node node_name adapter adapter_name`

3. Reboot the node hosting the adapter you changed.
4. Verify that the target port has the correct configuration:

   `network fcp adapter show -node node_name`

5. Bring your adapter online:

   `network fcp adapter modify -node node_name -adapter adapter_port -state up`

Display information about an FC target adapter

You can use the `network fcp adapter show` command to display system configuration and adapter information for any FC adapter in the system.

Step

1. Display information about the FC adapter by using the `network fcp adapter show` command.

   The output displays system configuration information and adapter information for each slot that is used.

   `network fcp adapter show -instance -node node1 -adapter 0a`

Change the FC adapter speed

You should set your adapter target port speed to match the speed of the device to which it connects, instead of using autonegotiation. A port that is set to autonegotiation can take longer time to reconnect after a takeover/giveback or other interruption.

What you'll need

All LIFs that use this adapter as their home port must be offline.

About this task

Because this task encompasses all storage virtual machines (SVMs) and all LIFs in a cluster, you must use the `-home-port` and `-home-lif` parameters to limit the scope of this operation. If you do not use these

parameters, the operation applies to all LIFs in the cluster, which might not be desirable.

Steps

1. Take all of the LIFs on this adapter offline:

   ```
   network interface modify -vserver * -lif * { -home-node node1 -home-port 0c }
   -status-admin down
   ```

2. Take the adapter offline:

   ```
   network fcp adapter modify -node node1 -adapter 0c -state down
   ```

   If the adapter does not go offline, you can also remove the cable from the appropriate adapter port on the system.

3. Determine the maximum speed for the port adapter:

   ```
   fcp adapter show -instance
   ```

   You cannot modify the adapter speed beyond the maximum speed.

4. Change the adapter speed:

   ```
   network fcp adapter modify -node node1 -adapter 0c -speed 16
   ```

5. Bring the adapter online:

   ```
   network fcp adapter modify -node node1 -adapter 0c -state up
   ```

6. Bring all of the LIFs on the adapter online:

   ```
   network interface modify -vserver * -lif * { -home-node node1 -home-port 0c }
   -status-admin up
   ```

Supported FC ports

The number of onboard FC ports and CNA/UTA2 ports configured for FC varies based on the model of the controller. FC ports are also available through supported FC target expansion adapters or additional UTA2 cards configured with FC SFP+ adapters.

Onboard FC, UTA, and UTA2 ports

- Onboard ports can be individually configured as either target or initiator FC ports.
- The number of onboard FC ports differs depending on controller model.

  The NetApp Hardware Universe contains a complete list of onboard FC ports on each controller model.

- FC ports are only available on FAS2240 systems through the X1150A-R6 expansion adapter.

  FAS2220 and FAS2520 systems do not support FC.

Target expansion adapter FC ports

- Available target expansion adapters differ depending on controller model.

  The NetApp Hardware Universe contains a complete list of target expansion adapters for each controller model.

- The ports on some FC expansion adapters are configured as initiators or targets at the factory and cannot be changed.

  Others can be individually configured as either target or initiator FC ports, just like the onboard FC ports. A complete list is available in NetApp Hardware Universe.

Prevent loss of connectivity when using the X1133A-R6 adapter

You can prevent loss of connectivity during a port failure by configuring your system with redundant paths to separate X1133A-R6 HBAs.

The X1133A-R6 HBA is a 4-port, 16 Gb FC adapter consisting of two 2-port pairs. The X1133A-R6 adapter can be configured as target mode or initiator mode. Each 2-port pair is supported by a single ASIC (for example, Port 1 and Port 2 on ASIC 1 and Port 3 and Port 4 on ASIC 2). Both ports on a single ASIC must be configured to operate in the same mode, either target mode or initiator mode. If an error occurs with the ASIC supporting a pair, both ports in the pair go offline.

To prevent this loss of connectivity, you configure your system with redundant paths to separate X1133A-R6 HBAs, or with redundant paths to ports supported by different ASICs on the HBA.

Manage X1143A-R6 adapters

Supported port configurations for X1143A-R6 adapters overview

By default the X1143A-R6 adapter is configured in FC target mode, but you can configure its ports as either 10 Gb Ethernet and FCoE (CNA) ports or as 16 Gb FC initiator or target ports. This requires different SFP+ adapters.

When configured for Ethernet and FCoE, X1143A-R6 adapters support concurrent NIC and FCoE target traffic on the same 10-GBE port. When configured for FC, each two-port pair that shares the same ASIC can be individually configured for FC target or FC initiator mode. This means that a single X1143A-R6 adapter can support FC target mode on one two-port pair and FC initiator mode on another two-port pair. Port pairs connected to the same ASIC must be configured in the same mode.

In FC mode, the X1143A-R6 adapter behaves just like any existing FC device with speeds up to 16 Gbps. In CNA mode, you can use the X1143A-R6 adapter for concurrent NIC and FCoE traffic sharing the same 10 GbE port. CNA mode only supports FC target mode for the FCoE function.

Configure the ports

To configure the unified target adapter (X1143A-R6), you must configure the two adjacent ports on the same chip in the same personality mode.

Steps

1. Configure the ports as needed for Fibre Channel (FC) or Converged Network Adapter (CNA) using the

`system node hardware unified-connect modify` command.

2. Attach the appropriate cables for FC or 10 Gb Ethernet.
3. Verify that you have the correct SFP+ installed:

   `network fcp adapter show -instance -node -adapter`

   For CNA, you should use a 10Gb Ethernet SFP. For FC, you should either use an 8 Gb SFP or a 16 Gb SFP, based on the FC fabric being connected to.

Change the UTA2 port from CNA mode to FC mode

You should change the UTA2 port from Converged Network Adapter (CNA) mode to Fibre Channel (FC) mode to support the FC initiator and FC target mode. You should change the personality from CNA mode to FC mode when you need to change the physical medium that connects the port to its network.

Steps

1. Take the adapter offline:

   `network fcp adapter modify -node node_name -adapter adapter_name -status-admin down`

2. Change the port mode:

   `ucadmin modify -node node_name -adapter adapter_name -mode fcp`

3. Reboot the node, and then bring the adapter online:

   `network fcp adapter modify -node node_name -adapter adapter_name -status-admin up`

4. Notify your admin or VIF manager to delete or remove the port, as applicable:

   - If the port is used as a home port of a LIF, is a member of an interface group (ifgrp), or hosts VLANs, then an admin should do the following:
     i. Move the LIFs, remove the port from the ifgrp, or delete the VLANs, respectively.
     ii. Manually delete the port by running the `network port delete` command.

     If the `network port delete` command fails, the admin should address the errors, and then run the command again.

   - If the port is not used as the home port of a LIF, is not a member of an ifgrp, and does not host VLANs, then the VIF manager should remove the port from its records at the time of reboot.

     If the VIF manager does not remove the port, then the admin must remove it manually after the reboot by using the `network port delete` command.

```
net-f8040-34::> network port show
```

```
    Node: net-f8040-34-01
                                                          Speed(Mbps) Health
    Port      IPspace      Broadcast Domain Link MTU  Admin/Oper  Status
    --------- ------------ ---------------- ---- ---- -----------
--------
    ...
    e0i       Default      Default          down 1500  auto/10     -
    e0f       Default      Default          down 1500  auto/10     -
    ...

    net-f8040-34::> ucadmin show
                           Current  Current    Pending  Pending    Admin
    Node          Adapter  Mode     Type       Mode     Type
Status
    ------------  -------  -------  ---------  -------  ---------
-----------
    net-f8040-34-01
                  0e       cna      target     -        -
offline
    net-f8040-34-01
                  0f       cna      target     -        -
offline
    ...

    net-f8040-34::> network interface create -vs net-f8040-34 -lif m
-role
node-mgmt-home-node net-f8040-34-01 -home-port e0e -address 10.1.1.1
-netmask 255.255.255.0


    net-f8040-34::> network interface show -fields home-port, curr-port

    vserver lif                    home-port curr-port
    ------- ---------------------- --------- ---------
    Cluster net-f8040-34-01_clus1 e0a       e0a
    Cluster net-f8040-34-01_clus2 e0b       e0b
    Cluster net-f8040-34-01_clus3 e0c       e0c
    Cluster net-f8040-34-01_clus4 e0d       e0d
    net-f8040-34
            cluster_mgmt           e0M       e0M
    net-f8040-34
            m                      e0e       e0i
    net-f8040-34
            net-f8040-34-01_mgmt1 e0M       e0M
    7 entries were displayed.
```

```
    net-f8040-34::> ucadmin modify local 0e fc

    Warning: Mode on adapter 0e and also adapter 0f will be changed to
fc.
    Do you want to continue? {y|n}: y
    Any changes will take effect after rebooting the system. Use the
"system node reboot" command to reboot.

    net-f8040-34::> reboot local
      (system node reboot)

    Warning: Are you sure you want to reboot node "net-f8040-34-01"?
    {y|n}: y
```

5. Verify that you have the correct SFP+ installed:

   `network fcp adapter show -instance -node -adapter`

   For CNA, you should use a 10Gb Ethernet SFP. For FC, you should either use an 8 Gb SFP or a 16 Gb SFP, before changing the configuration on the node.

Change the CNA/UTA2 target adapter optical modules

You should change the optical modules on the unified target adapter (CNA/UTA2) to support the personality mode you have selected for the adapter.

Steps

1. Verify the current SFP+ used in the card. Then, replace the current SFP+ with the appropriate SFP+ for the preferred personality (FC or CNA).
2. Remove the current optical modules from the X1143A-R6 adapter.
3. Insert the correct modules for your preferred personality mode (FC or CNA) optics.
4. Verify that you have the correct SFP+ installed:

   `network fcp adapter show -instance -node -adapter`

   Supported SFP+ modules and Cisco-branded Copper (Twinax) cables are listed in the NetApp Hardware Universe.

View adapter settings

To view the settings for your unified target adapter (X1143A-R6), you must run the `system hardware unified-connect show` command to display all modules on your controller.

Steps

1. Boot your controller without the cables attached.
2. Run the `system hardware unified-connect show` command to see the port configuration and

modules.

3. View the port information before configuring the CNA and ports.

Ways to Configure FCoE

Ways to Configure FCoE overview

FCoE can be configured in various ways using FCoE switches. Direct-attached configurations are not supported in FCoE.

All FCoE configurations are dual-fabric, fully redundant, and require host-side multipathing software. In all FCoE configurations, you can have multiple FCoE and FC switches in the path between the initiator and target, up to the maximum hop count limit. To connect switches to each other, the switches must run a firmware version that supports Ethernet ISLs. Each host in any FCoE configuration can be configured with a different operating system.

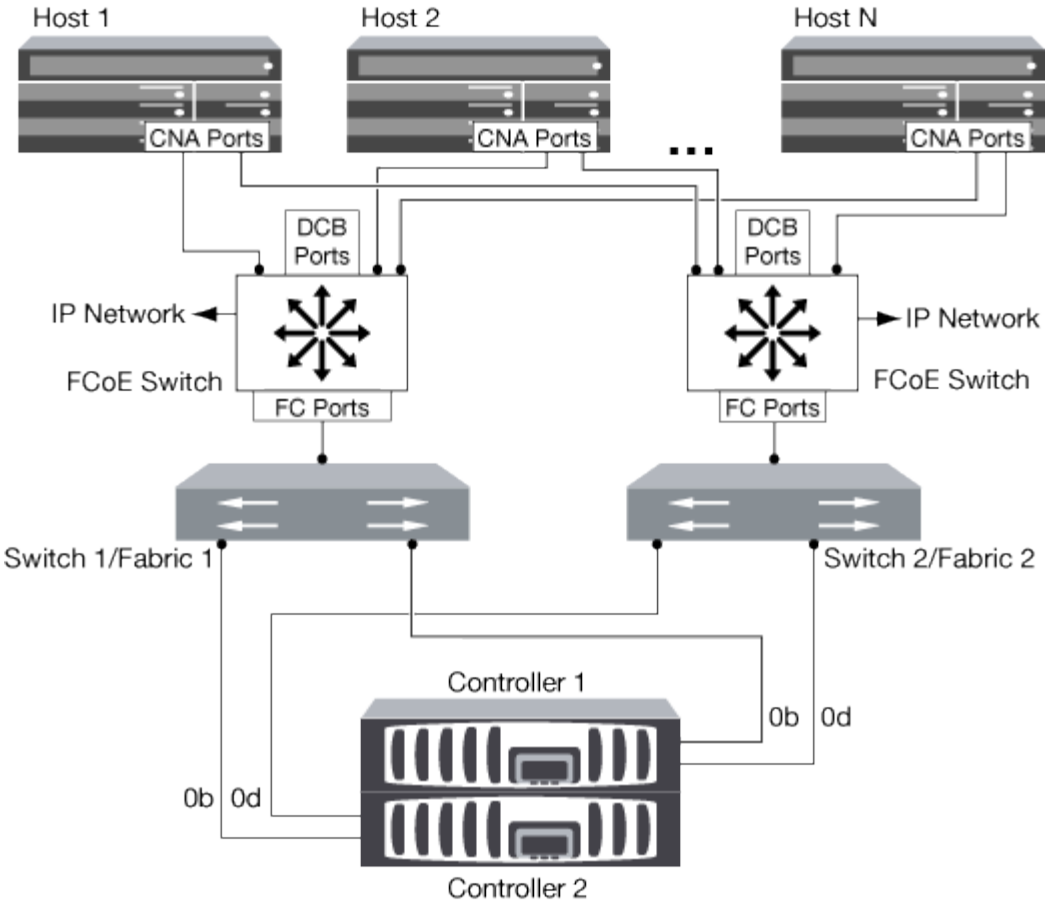
FCoE configurations require Ethernet switches that explicitly support FCoE features. FCoE configurations are validated through the same interoperability and quality assurance process as FC switches. Supported configurations are listed in the Interoperability Matrix. Some of the parameters included in these supported configurations are the switch model, the number of switches that can be deployed in a single fabric, and the supported switch firmware version.

The FC target expansion adapter port numbers in the illustrations are examples. The actual port numbers might vary, depending on the expansion slots in which the FCoE target expansion adapters are installed.

FCoE initiator to FC target

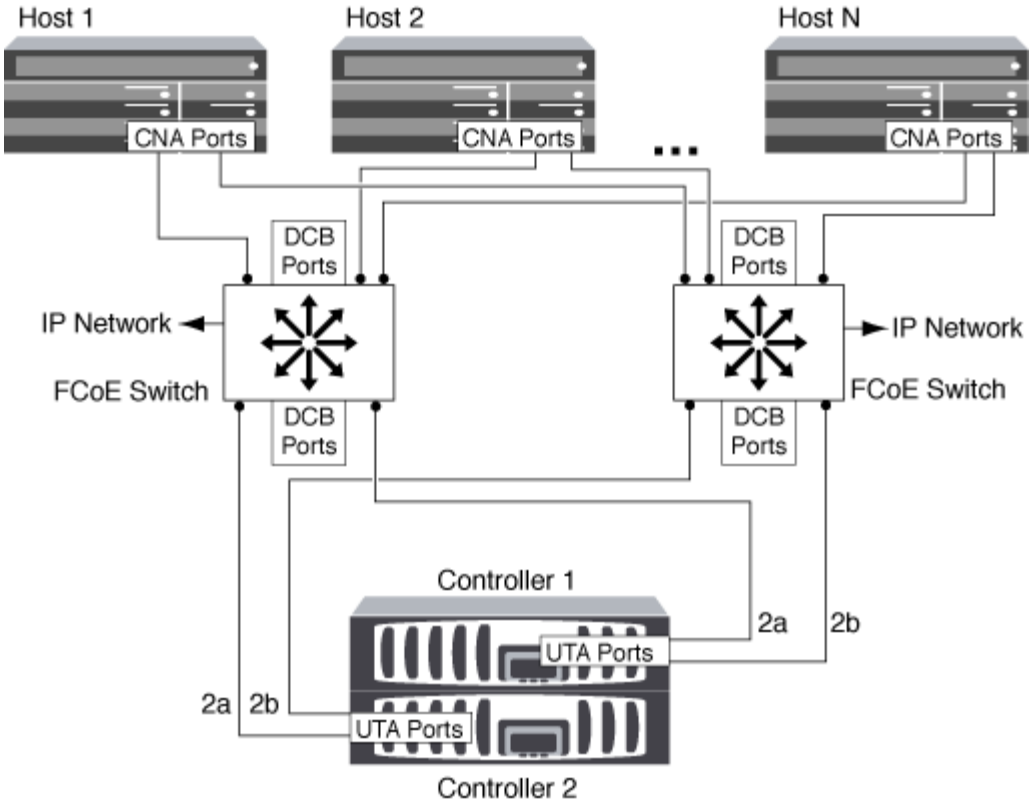
Using FCoE initiators (CNAs), you can connect hosts to both controllers in an HA pair through FCoE switches to FC target ports. The FCoE switch must also have FC ports. The host FCoE initiator always connects to the FCoE switch. The FCoE switch can connect directly to the FC target or can connect to the FC target through FC switches.

The following illustration shows host CNAs connecting to an FCoE switch, and then to an FC switch before connecting to the HA pair:



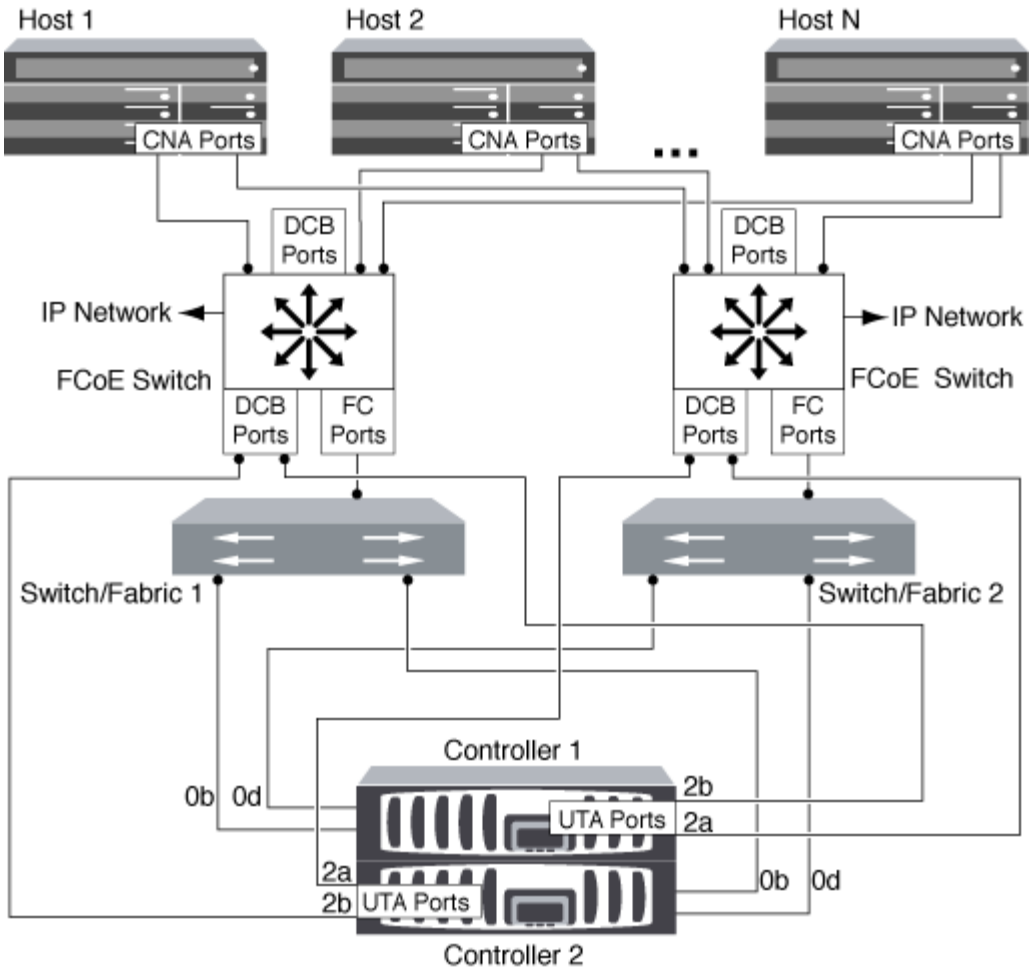
FCoE initiator to FCoE target

Using host FCoE initiators (CNAs), you can connect hosts to both controllers in an HA pair to FCoE target ports (also called UTAs or UTA2s) through FCoE switches.



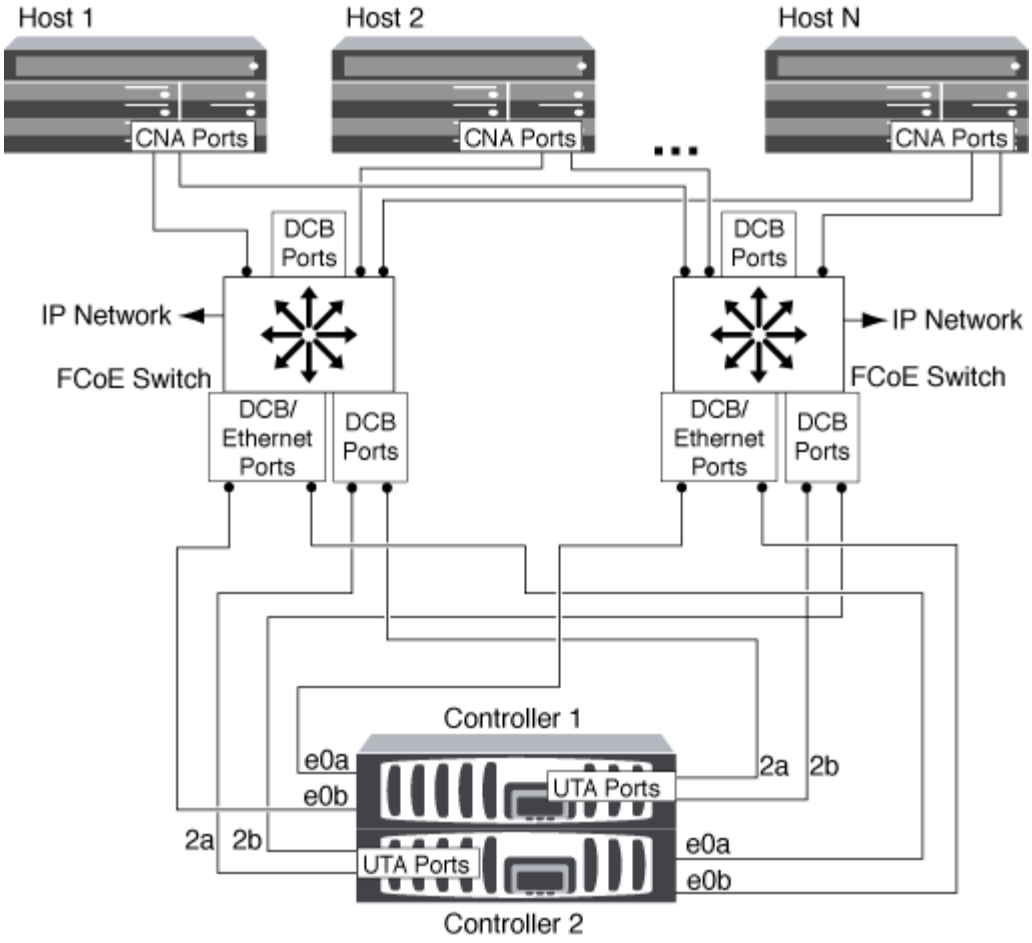
FCoE initiator to FCoE and FC targets

Using host FCoE initiators (CNAs), you can connect hosts to both controllers in an HA pair to FCoE and FC target ports (also called UTAs or UTA2s) through FCoE switches.



FCoE mixed with IP storage protocols

Using host FCoE initiators (CNAs), you can connect hosts to both controllers in an HA pair to FCoE target ports (also called UTAs or UTA2s) through FCoE switches. FCoE ports cannot use traditional link aggregation to a single switch. Cisco switches support a special type of link aggregation (Virtual Port Channel) that does support FCoE. A Virtual Port Channel aggregates individual links to two switches. You can also use Virtual Port Channels for other Ethernet traffic. Ports used for traffic other than FCoE, including NFS, SMB, iSCSI, and other Ethernet traffic, can use regular Ethernet ports on the FCoE switches.



FCoE initiator and target combinations

Certain combinations of FCoE and traditional FC initiators and targets are supported.

FCoE initiators

You can use FCoE initiators in host computers with both FCoE and traditional FC targets in storage controllers. The host FCoE initiator must connect to an FCoE DCB (data center bridging) switch; direct connection to a target is not supported.

The following table lists the supported combinations:

Initiator	Target	Supported?
FC	FC	Yes
FC	FCoE	Yes
FCoE	FC	Yes
FCoE	FCoE	Yes

FCoE targets

You can mix FCoE target ports with 4-Gb, 8-Gb, or 16-Gb FC ports on the storage controller regardless of whether the FC ports are add-in target adapters or onboard ports. You can have both FCoE and FC target adapters in the same storage controller.

The rules for combining onboard and expansion FC ports still apply.

FCoE supported hop count

The maximum supported Fibre Channel over Ethernet (FCoE) hop count between a host and storage system depends on the switch supplier and storage system support for FCoE configurations.

The hop count is defined as the number of switches in the path between the initiator (host) and target (storage system). Documentation from Cisco Systems also refers to this value as the *diameter of the SAN fabric*.

For FCoE, you can have FCoE switches connected to FC switches.

For end-to-end FCoE connections, the FCoE switches must be running a firmware version that supports Ethernet inter-switch links (ISLs).

The following table lists the maximum supported hop counts:

Switch supplier	Supported hop count
Brocade	7 for FC 5 for FCoE
Cisco	7 Up to 3 of the switches can be FCoE switches.

Fibre Channel and FCoE zoning

Fibre Channel and FCoE zoning overview

An FC, FC-NVMe or FCoE zone is a logical grouping of one or more ports within a fabric. For devices to be able see each other, connect, create sessions with one another, and communicate, both ports need to have a common zone membership. Single initiator zoning is recommended.

Reasons for zoning

- Zoning reduces or eliminates *crosstalk* between initiator HBAs.

  This occurs even in small environments and is one of the best arguments for implementing zoning. The logical fabric subsets created by zoning eliminate crosstalk problems.

- Zoning reduces the number of available paths to a particular FC, FC-NVMe, or FCoE port and reduces the number of paths between a host and a particular LUN that is visible.

  For example, some host OS multipathing solutions have a limit on the number of paths they can manage. Zoning can reduce the number of paths that an OS multipathing driver sees. If a host does not have a multipathing solution installed, you need to verify that only one path to a LUN is visible by using either zoning in the fabric or a combination of Selective LUN Mapping (SLM) and portsets in the SVM.

- Zoning increases security by limiting access and connectivity to end-points that share a common zone.

  Ports that have no zones in common cannot communicate with one another.

- Zoning improves SAN reliability by isolating problems that occur and helps to reduce problem resolution time by limiting the problem space.

Recommendations for zoning

- You should implement zoning any time, if four or more hosts are connected to a SAN or if SLM is not implemented on the nodes to a SAN.
- Although World Wide Node Name zoning is possible with some switch vendors, World Wide Port Name zoning is required to properly define a specific port and to use NPIV effectively.
- You should limit the zone size while still maintaining manageability.

  Multiple zones can overlap to limit size. Ideally, a zone is defined for each host or host cluster.

- You should use single-initiator zoning to eliminate crosstalk between initiator HBAs.

World Wide Name-based zoning

Zoning based on World Wide Name (WWN) specifies the WWN of the members to be included within the zone. When zoning in ONTAP, you must use World Wide Port Name (WWPN) zoning.

WWPN zoning provides flexibility because access is not determined by where the device is physically connected to the fabric. You can move a cable from one port to another without reconfiguring zones.

For Fibre Channel paths to storage controllers running ONTAP, be sure the FC switches are zoned using the WWPNs of the target logical interfaces (LIFs), not the WWPNs of the physical ports on the node. For more information on LIFs, see the *ONTAP Network Management Guide*.

Network management

Individual zones

In the recommended zoning configuration, there is one host initiator per zone. The zone consists of the host initiator port and one or more target LIFs on the storage nodes that are providing access to the LUNs up to the desired number of paths per target. This means that hosts accessing the same nodes cannot see each other's ports, but each initiator can access any node.

You should add all LIF's from the storage virtual machine (SVM) into the zone with the host initiator. This allows you to move volumes or LUNs without editing your existing zones or creating new zones.

For Fibre Channel paths to nodes running ONTAP, be sure that the FC switches are zoned using the WWPNs of the target logical interfaces (LIFs), not the WWPNs of the physical ports on the node. The WWPNs of the physical ports start with “50” and the WWPNs of the LIFs start with “20”.

Single-fabric zoning

In a single-fabric configuration, you can still connect each host initiator to each storage node. Multipathing software is required on the host to manage multiple paths. Each host should have two initiators for multipathing to provide resiliency in the solution.

Each initiator should have a minimum of one LIF from each node that the initiator can access. The zoning should allow at least one path from the host initiator to the HA pair of nodes in the cluster to provide a path for LUN connectivity. This means that each initiator on the host might only have one target LIF per node in its zone configuration. If there is a requirement for multipathing to the same node or multiple nodes in the cluster, then each node will have multiple LIFs per node in its zone configuration. This enables the host to still access its LUNs if a node fails or a volume containing the LUN is moved to a different node. This also requires the reporting nodes to be set appropriately.

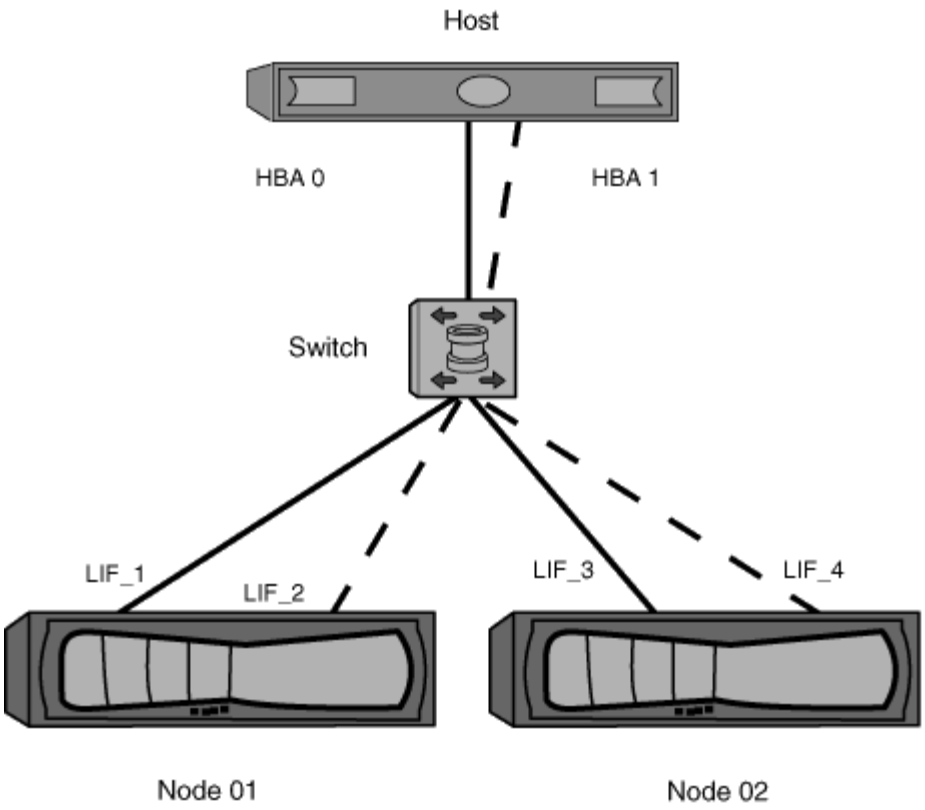
Single-fabric configurations are supported, but are not considered highly available. The failure of a single component can cause loss of access to data.

In the following figure, the host has two initiators and is running multipathing software. There are two zones:

> The naming convention used in this figure is just a recommendation of one possible naming convention that you can choose to use for your ONTAP solution.

- Zone 1: HBA 0, LIF_1, and LIF_3
- Zone 2: HBA 1, LIF_2, and LIF_4

If the configuration included more nodes, the LIFs for the additional nodes would be included in these zones.



In this example, you could also have all four LIFs in each zone. In that case, the zones would be as follows:

- Zone 1: HBA 0, LIF_1, LIF_2, LIF_3, and LIF_4
- Zone 2: HBA 1, LIF_1, LIF_2, LIF_3, and LIF_4

The host operating system and multipathing software have to support the number of supported paths that are being used to access the LUNs on the nodes. To determine the number of paths used to access the LUNs on nodes, see the SAN configuration limits section.

Related information

NetApp Hardware Universe

Dual-fabric HA pair zoning

In dual-fabric configurations, you can connect each host initiator to each cluster node. Each host initiator uses a different switch to access the cluster nodes. Multipathing software is required on the host to manage multiple paths.

Dual-fabric configurations are considered high availability because access to data is maintained if a single component fails.

In the following figure, the host has two initiators and is running multipathing software. There are two zones. SLM is configured so that all nodes are considered as reporting nodes.

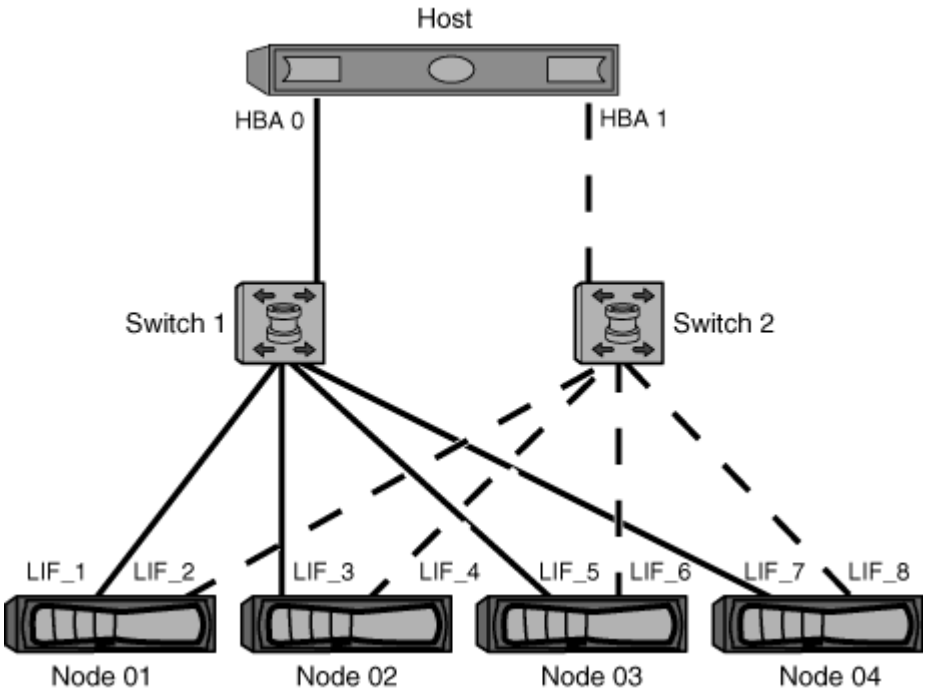
The naming convention used in this figure is just a recommendation of one possible naming convention that you can choose to use for your ONTAP solution.

- Zone 1: HBA 0, LIF_1, LIF_3, LIF_5, and LIF_7
- Zone 2: HBA 1, LIF_2, LIF_4, LIF_6, and LIF_8

Each host initiator is zoned through a different switch. Zone 1 is accessed through Switch 1. Zone 2 is accessed through Switch 2.

Each initiator can access a LIF on every node. This enables the host to still access its LUNs if a node fails. SVMs have access to all iSCSI and FC LIFs on every node in a clustered solution based on the setting for Selective LUN Map (SLM) and the reporting node configuration. You can use SLM, portsets, or FC switch zoning to reduce the number of paths from an SVM to the host and the number of paths from an SVM to a LUN.

If the configuration included more nodes, the LIFs for the additional nodes would be included in these zones.



The host operating system and multipathing software have to support the number of paths that is being used to access the LUNs on the nodes.

Related information

NetApp Hardware Universe

Zoning restrictions for Cisco FC and FCoE switches

When using Cisco FC and FCoE switches, a single fabric zone must not contain more than one target LIF for the same physical port. If multiple LIFs on the same port are in the same zone, then the LIF ports might fail to recover from a connection loss.

Regular FC switches are used for the FC-NVMe protocol in the exact same way they are used for the FC protocol.

- Multiple LIFs for the FC and FCoE protocols, can share physical ports on a node as long as they are in different zones.

- FC-NVMe and FCoE cannot share the same physical port.
- FC and FC-NVMe can share the same 32 Gb physical port.
- Cisco FC and FCoE switches require each LIF on a given port to be in a separate zone from the other LIFs on that port.
- A single zone can have both FC and FCoE LIFs. A zone can contain a LIF from every target port in the cluster, but be careful to not exceed the host's path limits and verify the SLM configuration.
- LIFs on different physical ports can be in the same zone.
- Cisco switches require that LIFs be separated.

  Though not required, separating LIFs is recommended for all switches

Requirements for shared SAN configurations

Shared SAN configurations are defined as hosts that are attached to both ONTAP storage systems and other vendors' storage systems. Accessing ONTAP storage systems and other vendors' storage systems from a single host is supported as long as several requirements are met.

For all of the host operating systems, it is a best practice to use separate adapters to connect to each vendor's storage systems. Using separate adapters reduces the chances of conflicting drivers and settings. For connections to an ONTAP storage system, the adapter model, BIOS, firmware, and driver must be listed as supported in the NetApp Interoperability Matrix Tool.

You should set the required or recommended timeout values and other storage parameters for the host. You must always install the NetApp software or apply the NetApp settings last.

- For AIX, you should apply the values from the AIX Host Utilities version that is listed in the Interoperability Matrix Tool for your configuration.
- For ESX, you should apply host settings by using Virtual Storage Console for VMware vSphere.
- For HP-UX, you should use the HP-UX default storage settings.
- For Linux, you should apply the values from the Linux Host Utilities version that is listed in the Interoperability Matrix Tool for your configuration.
- For Solaris, you should apply the values from the Solaris Host Utilities version that is listed in the Interoperability Matrix Tool for your configuration.
- For Windows, you should install the Windows Host Utilities version that is listed in the Interoperability Matrix Tool for your configuration.

Related information

NetApp Interoperability Matrix Tool

Host support for multipathing

Host support for multipathing overview

ONTAP always uses Asymmetric Logical Unit Access (ALUA) for both FC and iSCSI paths. Be sure to use host configurations that support ALUA for FC and iSCSi protocols.

Beginning with ONTAP 9.5 multipath HA pair failover/giveback is supported for NVMe configurations using Asynchronous Namespace Access (ANA). In ONTAP 9.4, NVMe only supports one path from host to target. The application host needs to manage path failover to its high availability (HA) partner.

For information about which specific host configurations support ALUA or ANA, see the NetApp Interoperability Matrix Tool and ONTAP SAN Host Configuration for your host operating system.

When host multipathing software is required

If there is more than one path from the storage virtual machine (SVM) logical interfaces (LIFs) to the fabric, multipathing software is required. Multipathing software is required on the host any time the host can access a LUN through more than one path.

The multipathing software presents a single disk to the operating system for all paths to a LUN. Without multipathing software, the operating system could treat each path as a separate disk, which can lead to data corruption.

Your solution is considered to have multiple paths if you have any of the following:

- A single initiator port in the host attaching to multiple SAN LIFs in the SVM
- Multiple initiator ports attaching to a single SAN LIF in the SVM
- Multiple initiator ports attaching to multiple SAN LIFs in the SVM

In single-fabric single-node configurations, multipathing software is not required if you only have a single path from the host to the node.

Multipathing software is recommended in HA configurations. In addition to Selective LUN Map, using FC switch zoning or portsets to limit the paths used to access LUNs is recommended.

Multipathing software is also known as MPIO (multipath I/O) software.

Recommended number of paths from host to nodes in cluster

You should not exceed more than eight paths from your host to each node in your cluster, paying attention to the total number of paths that can be supported for the host OS and the multipathing used on the host.

You should have a minimum of two paths per LUN connecting to each reporting node through Selective LUN Map (SLM) being used by the storage virtual machine (SVM) in your cluster. This eliminates single points of failure and enables the system to survive component failures.

If you have four or more nodes in your cluster or more than four target ports being used by the SVMs in any of your nodes, you can use the following methods to limit the number of paths that can be used to access LUNs on your nodes so that you do not exceed the recommended maximum of eight paths.

- SLM

  SLM reduces the number of paths from the host to LUN to only paths on the node owning the LUN and the owning node's HA partner. SLM is enabled by default.

- Portsets for iSCSI
- FC igroup mappings from your host

- FC switch zoning

Related information

SAN administration

Configuration limits

Determine the number of supported nodes for SAN configurations

The number of nodes per cluster supported by ONTAP varies depending on your version of ONTAP, the storage controller models in your cluster, and the protocol of your cluster nodes.

About this task

If any node in the cluster is configured for FC, FC-NVMe, FCoE, or iSCSI, that cluster is limited to the SAN node limits. Node limits based on the controllers in your cluster are listed in the *Hardware Universe*.

Steps

1. Go to NetApp Hardware Universe.
2. Click **Platforms** in the upper left (next to the **Home** button) and select the platform type.
3. Select the check box next to your version of ONTAP.

   A new column is displayed for you to choose your platforms.

4. Select the check boxes next to the platforms used in your solution.
5. Unselect the **Select All** check box in the **Choose Your Specifications** column.
6. Select the **Max Nodes per Cluster (NAS/SAN)** check box.
7. Click **Show Results**.

Related information

NetApp Hardware Universe

Determine the number of supported hosts per cluster in FC and FC-NVMe configurations

The maximum number of SAN hosts that can be connected to a cluster varies greatly based upon your specific combination of multiple cluster attributes, such as the number of hosts connected to each cluster node, initiators per host, sessions per host, and nodes in the cluster.

About this task

For FC and FC-NVMe configurations, you should use the number of initiator-target nexuses (ITNs) in your system to determine whether you can add more hosts to your cluster.

An ITN represents one path from the host's initiator to the storage system's target. The maximum number of ITNs per node in FC and FC-NVMe configurations is 2,048. As long as you are below the maximum number of ITNs, you can continue to add hosts to your cluster.

To determine the number of ITNs used in your cluster, perform the following steps for each node in the cluster.

Steps

1. Identify all the LIFs on a given node.
2. Run the following command for every LIF on the node:

   ```
   fcp initiator show -fields wwpn, lif
   ```

   The number of entries displayed at the bottom of the command output represents your number of ITNs for that LIF.

3. Record the number of ITNs displayed for each LIF.
4. Add the number of ITNs for each LIF on every node in your cluster.

   This total represents the number of ITNs in your cluster.

Determine the supported number of hosts in iSCSI configurations

The maximum number of SAN hosts that can be connected in iSCSI configurations varies greatly based on your specific combination of multiple cluster attributes, such as the number of hosts connected to each cluster node, initiators per host, logins per host, and nodes in the cluster.

About this task

The number of hosts that can be directly connected to a node or that can be connected through one or more switches depends on the number of available Ethernet ports. The number of available Ethernet ports is determined by the model of the controller and the number and type of adapters installed in the controller. The number of supported Ethernet ports for controllers and adapters is available in the *Hardware Universe*.

For all multi-node cluster configurations, you must determine the number of iSCSI sessions per node to know whether you can add more hosts to your cluster. As long as your cluster is below the maximum number of iSCSI sessions per node, you can continue to add hosts to your cluster. The maximum number of iSCSI sessions per node varies based on the types of controllers in your cluster.

Steps

1. Identify all of the target portal groups on the node.
2. Check the number of iSCSI sessions for every target portal group on the node:

   ```
   iscsi session show -tpgroup tpgroup
   ```

   The number of entries displayed at the bottom of the command output represents your number of iSCSI sessions for that target portal group.

3. Record the number of iSCSI sessions displayed for each target portal group.
4. Add the number of iSCSI sessions for each target portal group on the node.

   The total represents the number of iSCSI sessions on your node.

FC switch configuration limits

Fibre Channel switches have maximum configuration limits, including the number of logins supported per port, port group, blade, and switch. The switch vendors document their supported limits.

Each FC logical interface (LIF) logs into an FC switch port. The total number of logins from a single target on the node equals the number of LIFs plus one login for the underlying physical port. Do not exceed the switch vendor's configuration limits for logins or other configuration values. This also holds true for the initiators being used on the host side in virtualized environments with NPIV enabled. Do not exceed the switch vendor's configuration limits for logins for either the target or the initiators being used in the solution.

Brocade switch limits

You can find the configuration limits for Brocade switches in the *Brocade Scalability Guidelines*.

Cisco Systems switch limits

You can find the configuration limits for Cisco switches in the Cisco Configuration Limits guide for your version of Cisco switch software.

Calculate queue depth overview

You might need to tune your FC queue depth on the host to achieve the maximum values for ITNs per node and FC port fan-in. The maximum number of LUNs and the number of HBAs that can connect to an FC port are limited by the available queue depth on the FC target ports.

About this task

Queue depth is the number of I/O requests (SCSI commands) that can be queued at one time on a storage controller. Each I/O request from the host's initiator HBA to the storage controller's target adapter consumes a queue entry. Typically, a higher queue depth equates to better performance. However, if the storage controller's maximum queue depth is reached, that storage controller rejects incoming commands by returning a QFULL response to them. If a large number of hosts are accessing a storage controller, you should plan carefully to avoid QFULL conditions, which significantly degrade system performance and can lead to errors on some systems.

In a configuration with multiple initiators (hosts), all hosts should have similar queue depths. Because of the inequality in queue depth between hosts connected to the storage controller through the same target port, hosts with smaller queue depths are being deprived of access to resources by hosts with larger queue depths.

The following general recommendations can be made about "tuning" queue depths:

- For small to mid-size systems, use an HBA queue depth of 32.
- For large systems, use an HBA queue depth of 128.
- For exception cases or performance testing, use a queue depth of 256 to avoid possible queuing problems.
- All hosts should have the queue depths set to similar values to give equal access to all hosts.
- To avoid performance penalties or errors, the storage controller target FC port queue depth must not be exceeded.

Steps

1. Count the total number of FC initiators in all of the hosts that connect to one FC target port.
2. Multiply by 128.
   - If the result is less than 2,048, set the queue depth for all initiators to 128. You have 15 hosts with one initiator connected to each of two target ports on the storage controller. 15 × 128 = 1,920. Because 1,920 is less than the total queue depth limit of 2,048, you can set the queue depth for all of your initiators to 128.
   - If the result is greater than 2,048, go to step 3. You have 30 hosts with one initiator connected to each of two target ports on the storage controller. 30 × 128 = 3,840. Because 3,840 is greater than the total queue depth limit of 2,048, you should choose one of the options under step 3 for remediation.
3. Choose one of the following options to add more hosts to the storage controller.
   - Option 1:
     i. Add more FC target ports.
     ii. Redistribute your FC initiators.
     iii. Repeat steps 1 and 2.
     The desired queue depth of 3,840 exceeds the available queue depth per port. To remedy this, you can add a two-port FC target adapter to each controller, then rezone your FC switches so that 15 of your 30 hosts connect to one set of ports, and the remaining 15 hosts connect to a second set of ports. The queue depth per port is then reduced to 15 × 128 = 1,920.
   - Option 2:
     i. Designate each host as “large” or “small” based on its expected I/O need.
     ii. Multiply the number of large initiators by 128.
     iii. Multiply the number of small initiators by 32.
     iv. Add the two results together.
     v. If the result is less than 2,048, set the queue depth for large hosts to 128 and the queue depth for small hosts to 32.
     vi. If the result is still greater than 2,048 per port, reduce the queue depth per initiator until the total queue depth is less than or equal to 2,048.

> To estimate the queue depth needed to achieve a certain I/O per second throughput, use this formula:
>
> Needed queue depth = (Number of I/O per second) × (Response time)
>
> For example, if you need 40,000 I/O per second with a response time of 3 milliseconds, the needed queue depth = 40,000 × (.003) = 120.

The maximum number of hosts that you can connect to a target port is 64, if you decide to limit the queue depth to the basic recommendation of 32. However, if you decide to have a queue depth of 128, then you can have a maximum of 16 hosts connected to one target port. The larger the queue depth, the fewer hosts that a single target port can support. If your requirement is such that you cannot compromise on the queue depth, then you should get more target ports.

The desired queue depth of 3,840 exceeds the available queue depth per port. You have 10 “large” hosts that have high storage I/O needs, and 20 “small” hosts that have low I/O needs. Set the initiator queue depth on the large hosts to 128 and the initiator queue depth on the small hosts to 32.

Your resulting total queue depth is (10 × 128) + (20 × 32) = 1,920.

You can spread the available queue depth equally across each initiator.

Your resulting queue depth per initiator is 2,048 ÷ 30 = 68.

Set queue depths on SAN hosts

You might need to change the queue depths on your host to achieve the maximum values for ITNs per node and FC port fan-in.

AIX hosts

You can change the queue depth on AIX hosts using the `chdev` command. Changes made using the `chdev` command persist across reboots.

Examples:

- To change the queue depth for the hdisk7 device, use the following command:

  ```
  chdev -l hdisk7 -a queue_depth=32
  ```

- To change the queue depth for the fcs0 HBA, use the following command:

  ```
  chdev -l fcs0 -a num_cmd_elems=128
  ```

  The default value for `num_cmd_elems` is 200. The maximum value is 2,048.

  > ⓘ It might be necessary to take the HBA offline to change `num_cmd_elems` and then bring it back online using the `rmdev -l fcs0 -R` and `makdev -l fcs0 -P` commands.

HP-UX hosts

You can change the LUN or device queue depth on HP-UX hosts using the kernel parameter `scsi_max_qdepth`. You can change the HBA queue depth using the kernel parameter `max_fcp_reqs`.

- The default value for `scsi_max_qdepth` is 8. The maximum value is 255.

  `scsi_max_qdepth` can be dynamically changed on a running system using the `-u` option on the `kmtune` command. The change will be effective for all devices on the system. For example, use the following command to increase the LUN queue depth to 64:

  ```
  kmtune -u -s scsi_max_qdepth=64
  ```

  It is possible to change queue depth for individual device files using the `scsictl` command. Changes using the `scsictl` command are not persistent across system reboots. To view and change the queue depth for a particular device file, execute the following command:

  ```
  scsictl -a /dev/rdsk/c2t2d0

  scsictl -m queue_depth=16 /dev/rdsk/c2t2d0
  ```

- The default value for `max_fcp_reqs` is 512. The maximum value is 1024.

The kernel must be rebuilt and the system must be rebooted for changes to `max_fcp_reqs` to take effect. To change the HBA queue depth to 256, for example, use the following command:

```
kmtune -u -s max_fcp_reqs=256
```

Solaris hosts

You can set the LUN and HBA queue depth for your Solaris hosts.

- For LUN queue depth: The number of LUNs in use on a host multiplied by the per-LUN throttle (lun-queue-depth) must be less than or equal to the tgt-queue-depth value on the host.
- For queue depth in a Sun stack: The native drivers do not allow for per LUN or per target `max_throttle` settings at the HBA level. The recommended method for setting the `max_throttle` value for native drivers is on a per-device type (VID_PID) level in the `/kernel/drv/sd.conf` and `/kernel/drv/ssd.conf` files. The host utility sets this value to 64 for MPxIO configurations and 8 for Veritas DMP configurations.

Steps

1. `# cd/kernel/drv`
2. `# vi lpfc.conf`
3. Search for `/tft-queue (/tgt-queue)`

   ```
   tgt-queue-depth=32
   ```

   The default value is set to 32 at installation.

4. Set the desired value based on the configuration of your environment.
5. Save the file.
6. Reboot the host using the `sync; sync; sync; reboot -- -r` command.

VMware hosts for a QLogic HBA

Use the `esxcfg-module` command to change the HBA timeout settings. Manually updating the `esx.conf` file is not recommended.

Steps

1. Log on to the service console as the root user.
2. Use the `#vmkload_mod -l` command to verify which Qlogic HBA module is currently loaded.
3. For a single instance of a Qlogic HBA, run the following command:

   ```
   #esxcfg-module -s ql2xmaxqdepth=64 qla2300_707
   ```

   This example uses qla2300_707 module. Use the appropriate module based on the output of `vmkload_mod -l`.

4. Save your changes using the following command:

   ```
   #/usr/sbin/esxcfg-boot -b
   ```

5. Reboot the server using the following command:

```
#reboot
```

6. Confirm the changes using the following commands:
   a. `#esxcfg-module -g qla2300_707`
   b. `qla2300_707 enabled = 1 options = 'ql2xmaxqdepth=64'`

VMware hosts for an Emulex HBA

Use the `esxcfg-module` command to change the HBA timeout settings. Manually updating the `esx.conf` file is not recommended.

Steps

1. Log on to the service console as the root user.
2. Use the `#vmkload_mod -l grep lpfc` command to verify which Emulex HBA is currently loaded.
3. For a single instance of an Emulex HBA, enter the following command:

```
#esxcfg-module -s lpfc0_lun_queue_depth=16 lpfcdd_7xx
```

> Depending on the model of the HBA, the module can be either lpfcdd_7xx or lpfcdd_732. The above command uses the lpfcdd_7xx module. You should use the appropriate module based on the outcome of `vmkload_mod -l`.

   Running this command will set the LUN queue depth to 16 for the HBA represented by lpfc0.

4. For multiple instances of an Emulex HBA, run the following command:

```
a esxcfg-module -s "lpfc0_lun_queue_depth=16 lpfc1_lun_queue_depth=16"
lpfcdd_7xx
```

   The LUN queue depth for lpfc0 and the LUN queue depth for lpfc1 is set to 16.

5. Enter the following command:

```
#esxcfg-boot -b
```

6. Reboot using `#reboot`.

Windows hosts for an Emulex HBA

On Windows hosts, you can use the `LPUTILNT` utility to update the queue depth for Emulex HBAs.

Steps

1. Run the `LPUTILNT` utility located in the `C:\WINNT\system32` directory.
2. Select **Drive Parameters** from the menu on the right side.
3. Scroll down and double-click **QueueDepth**.

If you are setting **QueueDepth** greater than 150, the following Windows Registry value also need to be increased appropriately:

`HKEY_LOCAL_MACHINE\System\CurrentControlSet\Services\lpxnds\Parameters\Device\NumberOfRequests`

Windows hosts for a Qlogic HBA

On Windows hosts, you can use theand the `SANsurfer` HBA manager utility to update the queue depths for Qlogic HBAs.

Steps

1. Run the `SANsurfer` HBA manager utility.
2. Click on **HBA port** > **Settings**.
3. Click **Advanced HBA port settings** in the list box.
4. Update the `Execution Throttle` parameter.

Linux hosts for Emulex HBA

You can update the queue depths of an Emulex HBA on a Linux host. To make the updates persistent across reboots, you must then create a new RAM disk image and reboot the host.

Steps

1. Identify the queue depth parameters to be modified:

   ```
   modinfo lpfc|grep queue_depth
   ```

   The list of queue depth parameters with their description is displayed. Depending on your operating system version, you can modify one or more of the following queue depth parameters:

   - `lpfc_lun_queue_depth`: Maximum number of FC commands that can be queued to a specific LUN (uint)
   - `lpfc_hba_queue_depth`: Maximum number of FC commands that can be queued to an lpfc HBA (uint)
   - `lpfc_tgt_queue_depth`: Maximum number of FC commands that can be queued to a specific target port (uint)

     The `lpfc_tgt_queue_depth` parameter is applicable only for Red Hat Enterprise Linux 7.x systems, SUSE Linux Enterprise Server 11 SP4 systems and 12.x systems.

2. Update the queue depths by adding the queue depth parameters to the `/etc/modprobe.conf` file for a Red Hat Enterprise Linux 5.x system and to the `/etc/modprobe.d/scsi.conf` file for a Red Hat Enterprise Linux 6.x or 7.x system, or a SUSE Linux Enterprise Server 11.x or 12.x system.

   Depending on your operating system version, you can add one or more of the following commands:

   - `options lpfc lpfc_hba_queue_depth=new_queue_depth`
   - `options lpfc lpfc_lun_queue_depth=new_queue_depth`

- `options lpfc_tgt_queue_depth=new_queue_depth`

3. Create a new RAM disk image, and then reboot the host to make the updates persistent across reboots.

   For more information, see the System administration for your version of Linux operating system.

4. Verify that the queue depth values are updated for each of the queue depth parameter that you have modified:

   ```
   cat /sys/class/scsi_host/host_number/lpfc_lun_queue_depthcat
   /sys/class/scsi_host/host_number/lpfc_tgt_queue_depthcat
   /sys/class/scsi_host/host_number/lpfc_hba_queue_depth
   ```

   ```
   root@localhost ~]#cat /sys/class/scsi_host/host5/lpfc_lun_queue_depth
         30
   ```

   The current value of the queue depth is displayed.

Linux hosts for QLogic HBA

You can update the device queue depth of a QLogic driver on a Linux host. To make the updates persistent across reboots, you must then create a new RAM disk image and reboot the host. You can use the QLogic HBA management GUI or command-line interface (CLI) to modify the QLogic HBA queue depth.

This task shows how to use the QLogic HBA CLI to modify the QLogic HBA queue depth

Steps

1. Identify the device queue depth parameter to be modified:

   ```
   modinfo qla2xxx | grep ql2xmaxqdepth
   ```

   You can modify only the `ql2xmaxqdepth` queue depth parameter, which denotes the maximum queue depth that can be set for each LUN. The default value is 64 for RHEL 7.5 and later. The default value is 32 for RHEL 7.4 and earlier.

   ```
   root@localhost ~]# modinfo qla2xxx|grep ql2xmaxqdepth
   parm:       ql2xmaxqdepth:Maximum queue depth to set for each LUN.
   Default is 64. (int)
   ```

2. Update the device queue depth value:
   - If you want to make the modifications persistent, perform the following steps:
     i. Update the queue depths by adding the queue depth parameter to the `/etc/modprobe.conf` file for a Red Hat Enterprise Linux 5.x system and to the `/etc/modprobe.d/scsi.conf` file for a Red Hat Enterprise Linux 6.x or 7.x system, or a SUSE Linux Enterprise Server 11.x or 12.x system: `options qla2xxx ql2xmaxqdepth=new_queue_depth`
     ii. Create a new RAM disk image, and then reboot the host to make the updates persistent across reboots.

For more information, see the System administration for your version of Linux operating system.

- If you want to modify the parameter only for the current session, run the following command:

  `echo new_queue_depth > /sys/module/qla2xxx/parameters/ql2xmaxqdepth`

  In the following example, the queue depth is set to 128.

  ```
  echo 128 > /sys/module/qla2xxx/parameters/ql2xmaxqdepth
  ```

3. Verify that the queue depth values are updated:

   `cat /sys/module/qla2xxx/parameters/ql2xmaxqdepth`

   The current value of the queue depth is displayed.

4. Modify the QLogic HBA queue depth by updating the firmware parameter `Execution Throttle` from the QLogic HBA BIOS.
   a. Log in to the QLogic HBA management CLI:

      `/opt/QLogic_Corporation/QConvergeConsoleCLI/qaucli`

   b. From the main menu, select the `Adapter Configuration` option.

```
[root@localhost ~]#
/opt/QLogic_Corporation/QConvergeConsoleCLI/qaucli
Using config file:
/opt/QLogic_Corporation/QConvergeConsoleCLI/qaucli.cfg
Installation directory: /opt/QLogic_Corporation/QConvergeConsoleCLI
Working dir: /root

QConvergeConsole

        CLI - Version 2.2.0 (Build 15)

    Main Menu

    1:  Adapter Information
    **2:  Adapter Configuration**
    3:  Adapter Updates
    4:  Adapter Diagnostics
    5:  Monitoring
    6:  FabricCache CLI
    7:  Refresh
    8:  Help
    9:  Exit


        Please Enter Selection: 2
```

c. From the list of adapter configuration parameters, select the `HBA Parameters` option.

```
1:  Adapter Alias
    2:  Adapter Port Alias
    **3:  HBA Parameters**
    4:  Persistent Names (udev)
    5:  Boot Devices Configuration
    6:  Virtual Ports (NPIV)
    7:  Target Link Speed (iiDMA)
    8:  Export (Save) Configuration
    9:  Generate Reports
   10:  Personality
   11:  FEC
(p or 0: Previous Menu; m or 98: Main Menu; ex or 99: Quit)
        Please Enter Selection: 3
```

d. From the list of HBA ports, select the required HBA port.

```
Fibre Channel Adapter Configuration

    HBA Model QLE2562 SN: BFD1524C78510
      1: Port   1: WWPN: 21-00-00-24-FF-8D-98-E0 Online
      2: Port   2: WWPN: 21-00-00-24-FF-8D-98-E1 Online
    HBA Model QLE2672 SN: RFE1241G81915
      3: Port   1: WWPN: 21-00-00-0E-1E-09-B7-62 Online
      4: Port   2: WWPN: 21-00-00-0E-1E-09-B7-63 Online


        (p or 0: Previous Menu; m or 98: Main Menu; ex or 99: Quit)
        Please Enter Selection: 1
```

The details of the HBA port are displayed.

e. From the HBA Parameters menu, select the `Display HBA Parameters` option to view the current value of the `Execution Throttle` option.

The default value of the `Execution Throttle` option is 65535.

```
HBA Parameters Menu

========================================================
HBA           : 2 Port: 1
SN            : BFD1524C78510
HBA Model     : QLE2562
HBA Desc.     : QLE2562 PCI Express to 8Gb FC Dual Channel
FW Version    : 8.01.02
WWPN          : 21-00-00-24-FF-8D-98-E0
WWNN          : 20-00-00-24-FF-8D-98-E0
Link          : Online
========================================================

    1:  Display HBA Parameters
    2:  Configure HBA Parameters
    3:  Restore Defaults


        (p or 0: Previous Menu; m or 98: Main Menu; x or 99: Quit)
        Please Enter Selection: 1
------------------------------------------------------------------------
-----------
HBA Instance 2: QLE2562 Port 1 WWPN 21-00-00-24-FF-8D-98-E0 PortID
03-07-00
Link: Online
```

```
--------------------------------------------------------------------
-----------
Connection Options              : 2 - Loop Preferred, Otherwise Point-
to-Point
Data Rate                       : Auto
Frame Size                      : 2048
Hard Loop ID                    : 0
Loop Reset Delay (seconds)      : 5
Enable Host HBA BIOS            : Enabled
Enable Hard Loop ID             : Disabled
Enable FC Tape Support          : Enabled
Operation Mode                  : 0 - Interrupt for every I/O
completion
Interrupt Delay Timer (100us)   : 0
**Execution Throttle              : 65535**
Login Retry Count               : 8
Port Down Retry Count           : 30
Enable LIP Full Login           : Enabled
Link Down Timeout (seconds)     : 30
Enable Target Reset             : Enabled
LUNs Per Target                 : 128
Out Of Order Frame Assembly     : Disabled
Enable LR Ext. Credits          : Disabled
Enable Fabric Assigned WWN      : N/A

Press <Enter> to continue:
```

f. Press **Enter** to continue.

g. From the HBA Parameters menu, select the `Configure HBA Parameters` option to modify the HBA parameters.

h. From the Configure Parameters menu, select the `Execute Throttle` option and update the value of this parameter.

```
Configure Parameters Menu

=====================================================
HBA           : 2 Port: 1
SN            : BFD1524C78510
HBA Model     : QLE2562
HBA Desc.     : QLE2562 PCI Express to 8Gb FC Dual Channel
FW Version    : 8.01.02
WWPN          : 21-00-00-24-FF-8D-98-E0
WWNN          : 20-00-00-24-FF-8D-98-E0
Link          : Online
=====================================================

    1:  Connection Options
    2:  Data Rate
    3:  Frame Size
    4:  Enable HBA Hard Loop ID
    5:  Hard Loop ID
    6:  Loop Reset Delay (seconds)
    7:  Enable BIOS
    8:  Enable Fibre Channel Tape Support
    9:  Operation Mode
   10:  Interrupt Delay Timer (100 microseconds)
   11:  Execution Throttle
   12:  Login Retry Count
   13:  Port Down Retry Count
   14:  Enable LIP Full Login
   15:  Link Down Timeout (seconds)
   16:  Enable Target Reset
   17:  LUNs per Target
   18:  Enable Receive Out Of Order Frame
   19:  Enable LR Ext. Credits
   20:  Commit Changes
   21:  Abort Changes


        (p or 0: Previous Menu; m or 98: Main Menu; x or 99: Quit)
        Please Enter Selection: 11
Enter Execution Throttle [1-65535] [65535]: 65500
```

i. Press **Enter** to continue.

j. From the Configure Parameters menu, select the `Commit Changes` option to save the changes.

k. Exit the menu.

Considerations for SAN configurations in a MetroCluster environment

- MetroCluster configurations do not support front-end FC fabric “routed” vSAN configurations.
- Beginning with ONTAP 9.12.1, four-node MetroCluster IP configurations are supported on NVMe/FC. MetroCluster configurations are not supported for NVMe prior to ONTAP 9.12.1.
- Other SAN protocols such as iSCSI, FC, and FCoE are supported on MetroCluster configurations.
- When using SAN client configurations, you must check whether any special considerations for MetroCluster configurations are included in the notes that are provided in the NetApp Interoperability Matrix Tool (IMT).
- Operating systems and applications must provide an I/O resiliency of 120 seconds to support MetroCluster automatic unplanned switchover and Tiebreaker or Mediator-initiated switchover.
- The MetroCluster is using the same WWPNs on both sides of the front-end SAN.

Related information

Understanding MetroCluster data protection and disaster recovery

For further MetroCluster-specific host information, refer to the following NetApp Knowledge Base articles:

What are AIX Host support considerations in a MetroCluster configuration?

Solaris host support considerations in a MetroCluster configuration

Copyright information

Trademark information