

FUJITSU Storage ETERNUS AX/HX Series

MetroCluster® IP Installation and Configuration Guide

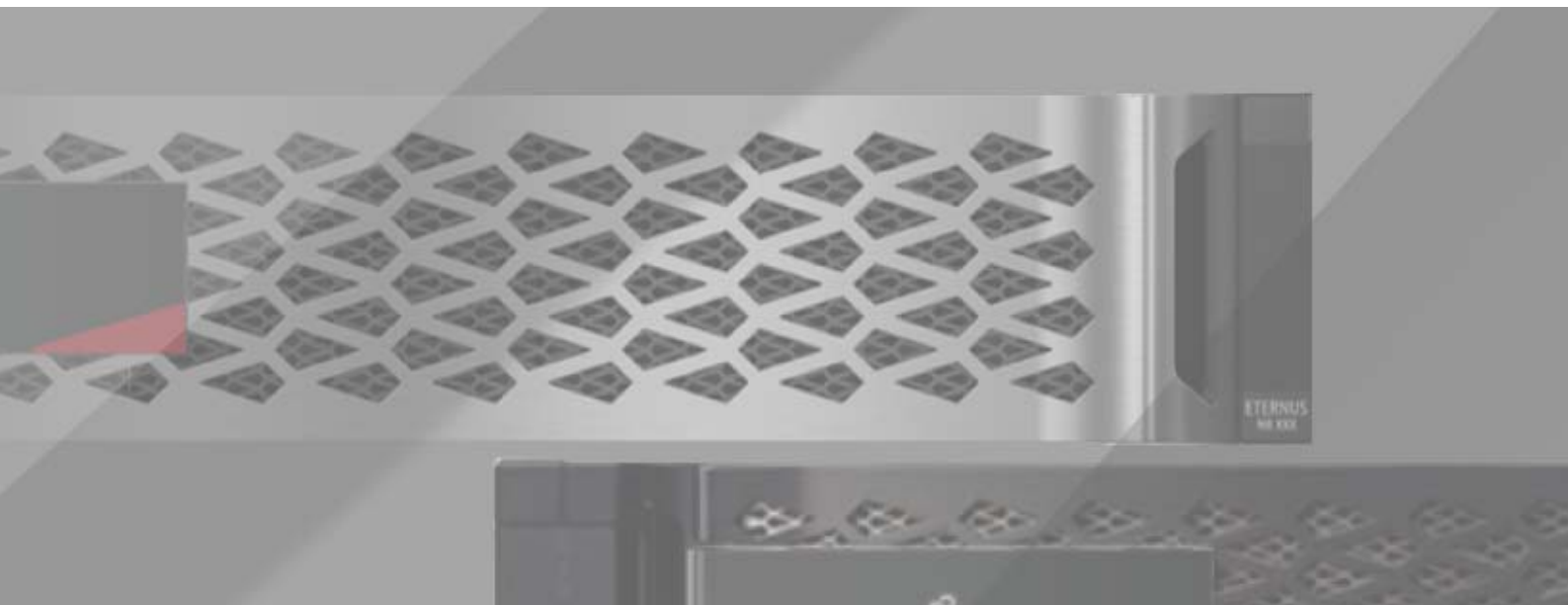


Table of Contents

1. Deciding whether to use the MetroCluster IP Installation and Configuration Guide	6
2. Preparing for the MetroCluster installation	7
Differences among the ONTAP MetroCluster configurations	7
Considerations for using All SAN Array systems in MetroCluster configurations	7
Access to remote storage in MetroCluster IP configurations	8
Considerations for using ONTAP Mediator or MetroCluster Tiebreaker	8
Interoperability of ONTAP Mediator with other applications and appliances	8
How the ONTAP Mediator supports automatic unplanned switchover	8
Considerations for MetroCluster IP configuration	9
Considerations for automatic drive assignment and ADP systems in ONTAP 9.7 and later	10
ADP and disk assignment differences by system in MetroCluster IP configurations	11
Considerations for configuring cluster peering	14
Prerequisites for cluster peering	14
Considerations when using dedicated ports	15
Considerations when sharing data ports	16
Considerations for sharing private layer 2 networks	16
ISL requirements in shared layer 2 networks	16
Considerations for ISLs	18
Required settings on intermediate switches	19
Examples of MetroCluster network topologies	20
Considerations for layer 3 wide-area networks	23
Considerations for using MetroCluster compliant switches	24
Considerations for using TDM/xWDM and encryption equipment with MetroCluster IP configurations	29
Considerations when using unmirrored aggregates	30
Considerations for firewall usage at MetroCluster sites	31
Preconfigured settings for new MetroCluster systems from the factory	31
Hardware setup checklist	31
3. Configuring the MetroCluster hardware components	33
Parts of a MetroCluster IP configuration	33
Illustration of the local HA pairs in a MetroCluster configuration	34

Illustration of the MetroCluster IP and cluster interconnect network.....	35
Illustration of the cluster peering network.....	36
Required MetroCluster IP components and naming conventions.....	36
Installing and cabling MetroCluster components	38
Racking the hardware components.....	38
Cabling the IP switches	39
Cabling the cluster peering connections	44
Cabling the management and data connections	45
Configuring the IP switches.....	45
4. Configuring the MetroCluster software in ONTAP.....	55
Gathering required information	55
IP network information worksheet for site A.....	55
IP network information worksheet for site B.....	57
Similarities and differences between standard cluster and MetroCluster configurations	59
Restoring system defaults on a previously used controller module	60
Verifying the ha-config state of components.....	61
Manually assigning drives to pool 0	61
Manually assigning drives for pool 0 (ONTAP 9.7 and later)	62
Setting up ONTAP	63
Configuring the clusters into a MetroCluster configuration.....	68
Verifying drive assignment of pool 0 drives	68
Peering the clusters	70
Creating the DR group	77
Configuring and connecting the MetroCluster IP interfaces	78
Verifying or manually performing pool 1 drives assignment.....	79
Mirroring the root aggregates	85
Creating a mirrored data aggregate on each node	86
Implementing the MetroCluster configuration.....	87
Creating unmirrored data aggregates.....	89
Checking the MetroCluster configuration	90
Completing ONTAP configuration.....	93
Verifying switchover, healing, and switchback	93
Configuring the MetroCluster Tiebreaker or ONTAP Mediator software	93
Protecting configuration backup files	94

5. Configuring the ONTAP Mediator service for unplanned automatic switchover	95
Installing and configuring the ONTAP Mediator service	95
Network requirements for using Mediator in a MetroCluster configuration.....	95
Guidelines for upgrading the ONTAP Mediator in a MetroCluster configuration	95
Installing or upgrading the ONTAP Mediator service	98
Configuring the ONTAP Mediator service from a MetroCluster IP configuration.....	112
Connecting a MetroCluster configuration to a different ONTAP Mediator instance	113
Changing the ONTAP Mediator password	113
Changing the ONTAP Mediator user name.....	114
Uninstalling the ONTAP Mediator service	115
6. Testing the MetroCluster configuration.....	116
Verifying negotiated switchover	116
Verifying healing and manual switchback.....	117
Verifying operation after power line disruption	119
Verifying operation after loss of a single storage shelf	120
7. Considerations when removing MetroCluster configurations	127
8. Considerations when using ONTAP in a MetroCluster configuration.....	128
MetroCluster operations in ONTAP System Manager.....	128
FlexCache support in a MetroCluster configuration.....	128
FabricPool support in MetroCluster configurations	129
Job schedules in a MetroCluster configuration.....	129
Cluster peering from the MetroCluster site to a third cluster	130
LDAP client configuration replication in a MetroCluster configuration	130
Networking and LIF creation guidelines for MetroCluster configurations.....	130
IPspace object replication and subnet configuration requirements	130
Requirements for LIF creation in a MetroCluster configuration	131
LIF replication and placement requirements and issues	131
Volume creation on a root aggregate	134
SVM disaster recovery in a MetroCluster configuration	134
SVM resynchronization at a disaster recovery site.....	135
Output for the storage aggregate plex show command is indeterminate after a MetroCluster switchover	136

Modifying volumes to set the NVFAIL flag in case of switchover	136
Monitoring and protecting the file system consistency using NVFAIL	136
How NVFAIL impacts access to NFS volumes or LUNs	137
Commands for monitoring data loss events.....	137
Accessing volumes in NVFAIL state after a switchover	138
Recovering LUNs in NVFAIL states after switchover.....	138
9. Where to find additional information	140
Copyright and trademark	141
Copyright.....	141
Trademark.....	141
How to send comments about documentation and receive update notifications	141

1. Deciding whether to use the MetroCluster IP Installation and Configuration Guide

This guide describes how to install and configure the MetroCluster IP hardware and software components. You should use this guide for planning, installing, and configuring a MetroCluster IP configuration under the following circumstances:

- You want to understand the architecture of a MetroCluster IP configuration.
- You want to understand the requirements and best practices for configuring a MetroCluster IP configuration.
- You want to use the command-line interface (CLI), not an automated scripting tool.

General information about ONTAP and MetroCluster configurations is also available.

[Fujitsu Manual Site](#)

2. Preparing for the MetroCluster installation

As you prepare for the MetroCluster installation, you should understand the MetroCluster hardware architecture and required components.

Differences among the ONTAP MetroCluster configurations

The various MetroCluster configurations have key differences in the required components.

In all configurations, each of the two MetroCluster sites is configured as an ONTAP cluster. In a two-node MetroCluster configuration, each node is configured as a single-node cluster.

Feature	IP configurations	Fabric-attached configurations		Stretch configurations	
		Four- or eight-node	Two-node	Two-node bridge-attached	Two-node direct-attached
Number of controllers	Four	Four or eight	Two	Two	Two
Uses an FC switch storage fabric	No	Yes	Yes	No	No
Uses an IP switch storage fabric	Yes	No	No	No	No
Uses FC-to-SAS bridges	No	Yes	Yes	Yes	No
Uses direct-attached SAS storage	Yes (local attached only)	No	No	No	Yes
Supports local HA	Yes	Yes	No	No	No
Supports automatic switchover	No	Yes	Yes	Yes	Yes
Supports unmirrored aggregates	Yes (starting in ONTAP 9.8)	Yes	Yes	Yes	Yes
Supports array LUNs	No	Yes	Yes	Yes	Yes

Caution

Note the following considerations for eight-node MetroCluster IP configurations:

- Eight-node configurations are supported starting in ONTAP 9.9.1.
- Only NetApp-validated MetroCluster switches (ordered from NetApp) are supported.
- Configurations using IP-routed (layer 3) backend connections are not supported.
- Configurations using shared private layer 2 networks are not supported.
- Configurations using a Cisco 9336C-FX2 shared switch are not supported.

Considerations for using All SAN Array systems in MetroCluster configurations

Some All SAN Arrays (AX Series) are supported in MetroCluster configurations. In the MetroCluster documentation, the information for ETERNUS AX models applies to the corresponding AX system. For example, all cabling and other information for the AX4100 system also applies to the AX4100 ASA system.

Access to remote storage in MetroCluster IP configurations

In MetroCluster IP configurations, the only way the local controllers can reach the remote storage pools is via the remote controllers. The IP switches are connected to the Ethernet ports on the controllers; they do not have direct connections to the disk shelves. If the remote controller is down, the local controllers cannot reach their remote storage pools.

This is different than MetroCluster FC configurations, in which the remote storage pools are connected to the local controllers via the FC fabric or the SAS connections. The local controllers still have access to the remote storage even if the remote controllers are down.

Considerations for using ONTAP Mediator or MetroCluster Tiebreaker

Starting with ONTAP 9.7, you can use either the ONTAP Mediator-assisted automatic unplanned switchover (MAUSO) in the MetroCluster IP configuration or you can use the MetroCluster Tiebreaker software. Only one of the two services can be used with the MetroCluster IP configuration.

The different MetroCluster configurations perform automatic switchover under different circumstances:

MetroCluster FC configurations using the AUSO capability (not present in MetroCluster IP configurations)

In these configurations, AUSO is initiated if controllers fail but the storage (and bridges, if present) remain operational.

MetroCluster IP configurations using the ONTAP Mediator service (ONTAP 9.7 and later)

In these configurations, MAUSO is initiated in the same circumstances as AUSO, as described above, and also after a complete site failure (controllers, storage, and switches).

Caution

MAUSO is initiated only if nonvolatile cache mirroring (*DR mirroring*) and SyncMirror plex mirroring is in sync at the time of the failure.

MetroCluster IP or FC configurations using the Tiebreaker software in active mode

In these configurations, the Tiebreaker initiates unplanned switchover after a complete site failure. Before using the Tiebreaker software, review the *Tiebreaker Software Installation and Configuration Guide*.

[Tiebreaker Software 1.21 Installation and Configuration Guide](#)

Interoperability of ONTAP Mediator with other applications and appliances

You cannot use any third-party applications or appliances that can trigger a switchover in combination with ONTAP Mediator. In addition, monitoring a MetroCluster configuration with MetroCluster Tiebreaker software is not supported when using ONTAP Mediator.

How the ONTAP Mediator supports automatic unplanned switchover

The ONTAP Mediator stores state information about the MetroCluster nodes in mailboxes located on the Mediator host. The MetroCluster nodes can use this information to monitor the state of their DR partners and implement a Mediator-assisted automatic unplanned switchover (MAUSO) in the case of a disaster.

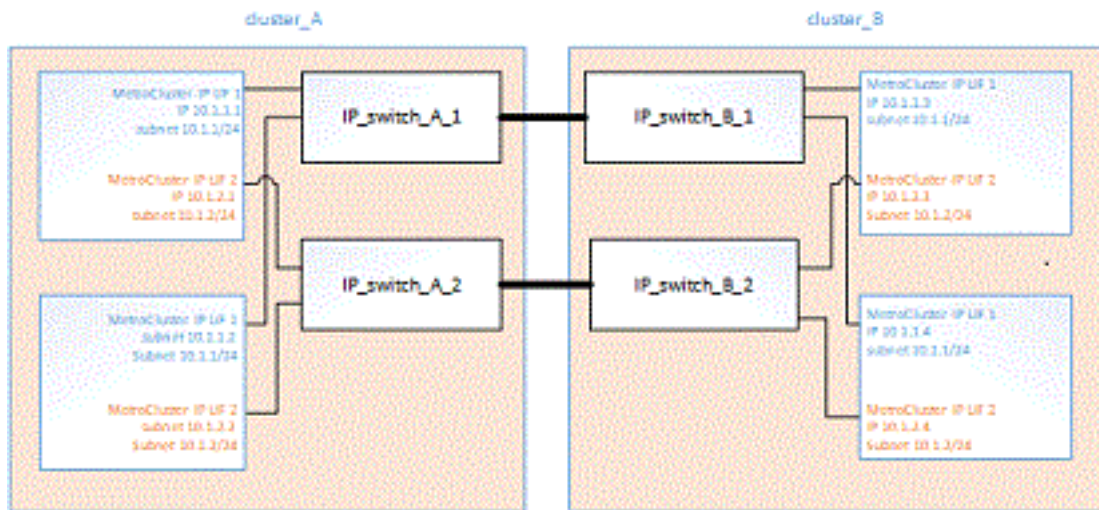
When a node detects a site failure requiring a switchover, it takes steps to confirm that the switchover is appropriate and, if so, performs the switchover.

MAUSO is only initiated if both SyncMirror mirroring and DR mirroring of each node's nonvolatile cache is operating and the caches and mirrors are synchronized at the time of the failure.

Considerations for MetroCluster IP configuration

You should be aware of how the MetroCluster IP addresses and interfaces are implemented in a MetroCluster IP configuration, as well as the associated requirements.

In a MetroCluster IP configuration, replication of storage and nonvolatile cache between the HA pairs and the DR partners is performed over high-bandwidth dedicated links in the MetroCluster IP fabric. iSCSI connections are used for storage replication. The IP switches are also used for all intra-cluster traffic within the local clusters. The MetroCluster traffic is kept separate from the intra-cluster traffic by using separate IP subnets and VLANs. The MetroCluster IP fabric is distinct and different from the cluster peering network.



The MetroCluster IP configuration requires two IP addresses on each node that are reserved for the back-end MetroCluster IP fabric. The reserved IP addresses are assigned to MetroCluster IP logical interfaces (LIFs) during initial configuration, and have the following requirements:

Caution

You must choose the MetroCluster IP addresses carefully because you cannot change them after initial configuration.

- They must fall in a unique IP range.
They must not overlap with any IP space in the environment.
- They must reside in one of two IP subnets that separate them from all other traffic.

For example, the nodes might be configured with the following IP addresses:

Node	Interface	IP address	Subnet
node_A_1	MetroCluster IP interface 1	10.1.1.1	10.1.1/24
	MetroCluster IP interface 2	10.1.2.1	10.1.2/24
node_A_2	MetroCluster IP interface 1	10.1.1.2	10.1.1/24
	MetroCluster IP interface 2	10.1.2.2	10.1.2/24
node_B_1	MetroCluster IP interface 1	10.1.1.3	10.1.1/24
	MetroCluster IP interface 2	10.1.2.3	10.1.2/24
node_B_2	MetroCluster IP interface 1	10.1.1.4	10.1.1/24
	MetroCluster IP interface 2	10.1.2.4	10.1.2/24

- Characteristics of MetroCluster IP interfaces

The MetroCluster IP interfaces are specific to MetroCluster IP configurations. They have different characteristics from other ONTAP interface types:

- They are created by the `metrocluster configuration-settings interface create` command as part of the initial MetroCluster configuration.

Note

Starting with ONTAP 9.9.1, if you are using a layer 3 configuration, you must also specify the `-gateway` parameter when creating MetroCluster IP interfaces.

They are not created or modified by the `network interface` commands.

- They do not appear in the output of the `network interface show` command.
- They do not fail over, but remain associated with the port on which they were created.
- MetroCluster IP configurations use specific Ethernet ports (depending on the platform) for the MetroCluster IP interfaces.

Considerations for automatic drive assignment and ADP systems in ONTAP 9.7 and later

In most configurations, partitioning and disk assignment is performed automatically during the initial configuration of the MetroCluster sites.

ONTAP 9.7 and later releases include the following changes for ADP support:

- Pool 0 disk assignments are done at the factory.
- The unmirrored root is created at the factory.
- Data partition assignment is done at the customer site during the setup procedure.
- In most cases, drive assignment and partitioning is done automatically during the setup procedures.

- Automatic partitioning

ADP is performed automatically during initial configuration of the platform.

- How shelf-by-shelf automatic assignment works

If there are four external shelves per site, each shelf is assigned to a different node and different pool, as shown in the following example:

- All of the disks on site_A-shelf_1 are automatically assigned to pool 0 of node_A_1
- All of the disks on site_A-shelf_3 are automatically assigned to pool 0 of node_A_2
- All of the disks on site_B-shelf_1 are automatically assigned to pool 0 of node_B_1
- All of the disks on site_B-shelf_3 are automatically assigned to pool 0 of node_B_2
- All of the disks on site_B-shelf_2 are automatically assigned to pool 1 of node_A_1
- All of the disks on site_B-shelf_4 are automatically assigned to pool 1 of node_A_2
- All of the disks on site_A-shelf_2 are automatically assigned to pool 1 of node_B_1
- All of the disks on site_A-shelf_4 are automatically assigned to pool 1 of node_B_2

- Adding shelves to an existing configuration.

Automatic drive assignment supports the symmetrical addition of shelves to an existing configuration.

2. Preparing for the MetroCluster installation

Considerations for automatic drive assignment and ADP systems in ONTAP 9.7 and later

When new shelves are added, the system applies the same assignment policy to newly added shelves. For example, with a single shelf per site, if an additional shelf is added, the systems applies the quarter-shelf assignment rules to the new shelf.

Related information

[Required MetroCluster IP components and naming conventions \(page 36\)](#)

[Disks and Aggregates Power Guide](#)

ADP and disk assignment differences by system in MetroCluster IP configurations

The operation of Advanced Drive Partitioning (ADP) and automatic disk assignment in MetroCluster IP configurations varies depending on the system model.

Caution

In systems using ADP, aggregates are created using partitions in which each drive is partitioned in to P1, P2 and P3 partitions. The root aggregate is created using P3 partitions.

You must meet the MetroCluster limits for the maximum number of supported drives and other guidelines.

● ADP and disk assignment on AX2100 systems

Guideline	Shelves per site	Drive assignment rules	ADP layout for root partition
Minimum recommended shelves (per site)	Internal drives only	<p>The internal drives are divided into four equal groups. Each group is automatically assigned to a separate pool and each pool is assigned to a separate controller in the configuration.</p> <p>Caution</p> <p>Half of the internal drives remain unassigned before MetroCluster is configured.</p>	<p>Two quarters are used by the local HA pair. The other two quarters are used by the remote HA pair.</p> <p>The root aggregate includes the following partitions in each plex:</p> <ul style="list-style-type: none">• Three partitions for data• Two parity partitions• One spare partition

2. Preparing for the MetroCluster installation

Considerations for automatic drive assignment and ADP systems in ONTAP 9.7 and later

Guideline	Shelves per site	Drive assignment rules	ADP layout for root partition
Minimum supported shelves (per site)	16 internal drives	<p>The drives are divided into four equal groups. Each quarter-shelf is automatically assigned to a separate pool.</p> <p>Two quarters on a shelf can have the same pool. The pool is chosen based on the node that owns the quarter:</p> <ul style="list-style-type: none"> • If owned by the local node, pool0 is used. • If owned by the remote node, pool1 is used. <p>For example: a shelf with quarters Q1 through Q4 can have following assignments:</p> <ul style="list-style-type: none"> • Q1: node_A_1 pool0 • Q2: node_A_2 pool0 • Q3: node_B_1 pool1 • Q4: node_B_2 pool1 <div style="background-color: #fff9c4; padding: 5px; margin-top: 10px;"> <p>Caution</p> <p>Half of the internal drives remain unassigned before MetroCluster is configured.</p> </div>	<p>Each of the two plexes in the root aggregate includes the following partitions:</p> <ul style="list-style-type: none"> • One partition for data • Two parity partitions • One spare partition

● ADP and disk assignment on AX2200 systems

Guideline	Shelves per site	Drive assignment rules	ADP layout for root partition
Minimum recommended shelves (per site)	Two shelves	<p>The drives on each external shelf are divided into two equal groups (halves). Each half-shelf is automatically assigned to a separate pool.</p>	<p>One shelf is used by the local HA pair. The second shelf is used by the remote HA pair. Partitions on each shelf are used to create the root aggregate. The root aggregate includes the following partitions in each plex:</p> <ul style="list-style-type: none"> • Eight partitions for data • Two parity partitions • Two spare partitions

2. Preparing for the MetroCluster installation

Considerations for automatic drive assignment and ADP systems in ONTAP 9.7 and later

Guideline	Shelves per site	Drive assignment rules	ADP layout for root partition
Minimum supported shelves (per site)	24 internal drives only	The drives are divided into four equal groups. Each quarter-shelf is automatically assigned to a separate pool.	Each of the two plexes in the root aggregate includes the following partitions: <ul style="list-style-type: none"> • Three partitions for data • Two parity partitions • One spare partition

● ADP and disk assignment on AX3000 systems

Guideline	Shelves per site	Drive assignment rules	ADP layout for root partition
Minimum recommended shelves (per site)	Two shelves	The drives on each external shelf are divided into two equal groups (halves). Each half-shelf is automatically assigned to a separate pool.	One shelf is used by the local HA pair. The second shelf is used by the remote HA pair. Partitions on each shelf are used to create the root aggregate. The root aggregate includes the following partitions in each plex: <ul style="list-style-type: none"> • Eight partitions for data • Two parity partitions • Two spare partitions
Minimum supported shelves (per site)	One shelf	The drives are divided into four equal groups. Each quarter-shelf is automatically assigned to a separate pool.	Each of the two plexes in the root aggregate includes the following partitions: <ul style="list-style-type: none"> • Three partitions for data • Two parity partitions • One spare partition

● Disk assignment on HX2200 systems

Guideline	Shelves per site	Drive assignment rules	ADP layout for root partition
Minimum recommended shelves (per site)	One internal and one external shelf	The internal and external shelves are divided into two equal halves. Each half is automatically assigned to different pool	Not applicable.
Minimum supported shelves (per site) (active/passive HA configuration)	Internal drives only	Manual assignment required.	

● Disk assignment on HX6100 systems

Guideline	Shelves per site	Drive assignment rules	ADP layout for root partition
Minimum supported shelves (per site)	Two shelves	The drives on the external shelves are divided into two equal groups (halves). Each half-shelf is automatically assigned to a separate pool.	Not applicable.
Minimum supported shelves (per site) (active/passive HA configuration)	One shelf	Manual assignment required.	

● Disk assignment on FAS500f systems

Guideline	Shelves per site	Drive assignment rules	ADP layout for root partition
Minimum recommended shelves (per site)	Four shelves	Drives are automatically assigned on a shelf-by-shelf basis.	Not applicable.
Minimum supported shelves (per site)	One shelf	The drives are divided into four equal groups. Each quarter-shelf is automatically assigned to a separate pool.	

Considerations for configuring cluster peering

Each MetroCluster site is configured as a peer to its partner site. You should be familiar with the prerequisites and guidelines for configuring the peering relationships and when deciding whether to use shared or dedicated ports for those relationships.

Related information

[Cluster and SVM Peering Power Guide](#)

Prerequisites for cluster peering

Before you set up cluster peering, you should confirm that the connectivity, port, IP address, subnet, firewall, and cluster-naming requirements are met.

● Connectivity requirements

Every intercluster LIF on the local cluster must be able to communicate with every intercluster LIF on the remote cluster.

Although it is not required, it is typically simpler to configure the IP addresses used for intercluster LIFs in the same subnet. The IP addresses can reside in the same subnet as data LIFs, or in a different subnet. The subnet used in each cluster must meet the following requirements:

- The subnet must have enough IP addresses available to allocate to one intercluster LIF per node. For example, in a six-node cluster, the subnet used for intercluster communication must have six available IP addresses.

Each node must have an intercluster LIF with an IP address on the intercluster network.

Intercluster LIFs can have an IPv4 address or an IPv6 address.

- **Port requirements**

You can use dedicated ports for intercluster communication, or share ports used by the data network. Ports must meet the following requirements:

- All ports that are used to communicate with a given remote cluster must be in the same IPspace. You can use multiple IPspaces to peer with multiple clusters. Pair-wise full-mesh connectivity is required only within an IPspace.
- The broadcast domain that is used for intercluster communication must include at least two ports per node so that intercluster communication can fail over from one port to another port. Ports added to a broadcast domain can be physical network ports, VLANs, or interface groups (ifgrps).
- All ports must be cabled.
- All ports must be in a healthy state.
- The MTU settings of the ports must be consistent.

- **Firewall requirements**

Firewalls and the intercluster firewall policy must allow the following protocols:

- ICMP service
- TCP to the IP addresses of all the intercluster LIFs over the ports 10000, 11104, and 11105
- Bidirectional HTTPS between the intercluster LIFs

The default `intercluster` firewall policy allows access through the HTTPS protocol and from all IP addresses (0.0.0.0/0). You can modify or replace the policy if necessary.

Related information

[Considerations when using dedicated ports \(page 15\)](#)

[Considerations when sharing data ports \(page 16\)](#)

Considerations when using dedicated ports

When determining whether using a dedicated port for intercluster replication is the correct intercluster network solution, you should consider configurations and requirements such as LAN type, available WAN bandwidth, replication interval, change rate, and number of ports.

Consider the following aspects of your network to determine whether using a dedicated port is the best intercluster network solution:

- If the amount of available WAN bandwidth is similar to that of the LAN ports and the replication interval is such that replication occurs while regular client activity exists, then you should dedicate Ethernet ports for intercluster replication to avoid contention between replication and the data protocols.
- If the network utilization generated by the data protocols (CIFS, NFS, and iSCSI) is such that the network utilization is above 50 percent, then you should dedicate ports for replication to allow for nondegraded performance if a node failover occurs.
- When physical 10 GbE or faster ports are used for data and replication, you can create VLAN ports for replication and dedicate the logical ports for intercluster replication. The bandwidth of the port is shared between all VLANs and the base port.
- Consider the data change rate and replication interval and whether the amount of data that must be replicated on each interval requires enough bandwidth that it might cause contention with data protocols if sharing data ports.

Related information

[Prerequisites for cluster peering \(page 14\)](#)

[Considerations when sharing data ports \(page 16\)](#)

Considerations when sharing data ports

When determining whether sharing a data port for intercluster replication is the correct intercluster network solution, you should consider configurations and requirements such as LAN type, available WAN bandwidth, replication interval, change rate, and number of ports.

Consider the following aspects of your network to determine whether sharing data ports is the best intercluster connectivity solution:

- For a high-speed network, such as a 40-Gigabit Ethernet (40-GbE) network, a sufficient amount of local LAN bandwidth might be available to perform replication on the same 40-GbE ports that are used for data access. In many cases, the available WAN bandwidth is far less than 10 GbE LAN bandwidth.
- All nodes in the cluster might have to replicate data and share the available WAN bandwidth, making data port sharing more acceptable.
- Sharing ports for data and replication eliminates the extra port counts required to dedicate ports for replication.
- The maximum transmission unit (MTU) size of the replication network will be the same size as that used on the data network.
- Consider the data change rate and replication interval and whether the amount of data that must be replicated on each interval requires enough bandwidth that it might cause contention with data protocols if sharing data ports.
- When data ports for intercluster replication are shared, the intercluster LIFs can be migrated to any other intercluster-capable port on the same node to control the specific data port that is used for replication.

Related information

[Prerequisites for cluster peering \(page 14\)](#)

[Considerations when using dedicated ports \(page 15\)](#)

Considerations for sharing private layer 2 networks

MetroCluster IP switches are dedicated to the MetroCluster configuration and cannot be shared. Therefore, a set of MetroCluster IP switches can only connect one MetroCluster configuration. Only the MetroCluster ISL ports on the MetroCluster IP switches can connect to the shared switches.



If using a shared network, the customer is responsible for meeting the MetroCluster network requirements in the shared network.

ISL requirements in shared layer 2 networks

When sharing ISL traffic in a shared network, you must ensure that you have adequate capacity and size the ISLs appropriately. Low latency is critical for replication of data between the MetroCluster sites. Latency issues on these connections can impact client I/O.

You should review these sections to correctly calculate the required end-to-end capacity of the ISLs. Continuous nonvolatile cache and storage replication with low latency is critical for MetroCluster configurations. The latency in the back-end network impacts the latency and throughput seen by client IO.

- Latency and packet loss limits in the ISLs

The following requirements must be met for round-trip traffic between the MetroCluster IP switches at site_A and site_B, with the MetroCluster configuration in steady state operation:

- Round trip latency must be less than or equal to 7 ms.

The maximum distance is 700 km, so the distance between the sites is limited by the latency or the maximum distance, whichever is reached first.

As the distance between two MetroCluster sites increases, latency increases, usually in the range of 1 ms round-trip delay time per 100 km (62 miles). This latency also depends on the network service level agreement (SLA) in terms of the bandwidth of the ISL links, packet drop rate, and jitter on the network. Low bandwidth, high jitter, and random packet drops lead to different recovery mechanisms by the switches or the TCP engine on the controller modules for successful packet delivery. These recovery mechanisms can increase overall latency.

Any device that contributes to latency must be accounted for.

- Packet loss must be less than or equal to 0.01%.

Packet loss includes physical loss or loss due to congestion or over-subscription.

Packet drops can cause retransmissions and a reduced congestion window.

- The supported jitter value is 3 ms for round trip (or 1.5 ms for one way).

- The network should allocate and maintain the SLA for the bandwidth required for MetroCluster traffic, accounting for microbursts and spikes in the traffic.

Low bandwidth can cause queuing delays and tail drops on switches. If you are using ONTAP 9.7 or later, the network intermediate between the two sites must provide a minimum bandwidth of 4.5 Gbps for the MetroCluster configuration.

- MetroCluster traffic should not consume the complete bandwidth and have negative impact on non-MetroCluster traffic.

- The shared network should have network monitoring configured to monitor the ISLs for utilization, errors (drops, link flaps, corruption, etc.) and failures.

- Connection limits and trunking in the customer switches

The intermediate customer-provided switches must meet the following requirements:

- The number of intermediate switches is not limited, and more than two switches between the MetroCluster IP switches is supported.

The MetroCluster IP switches should be located as close as possible to the intermediate switches providing the long-haul link. All of the ISL connections along the route must meet all of the requirements for MetroCluster ISL.

- The ISLs in the customer network (the ISLs between the customer switches) must be configured in such way that sufficient bandwidth is provided and order of delivery is preserved.

This can be done with trunking a sufficient number of links and enforcing load balancing policies to preserve order.

- Other network requirements

The intermediate customer-provided switches must meet the following requirements:

- The MetroCluster traffic uses fixed VLAN IDs that are set in the provided RCF files.

Layer 2 VLANs with IDs that match the MetroCluster VLAN IDs must span the shared network. AX2100 and HX2200 systems require VLAN 10 and VLAN 20 unless they are changed during interface creation. Other systems are not restricted to specific VLAN IDs.

- The MTU size must be set to 9216 on all devices in the end-to-end network.

- No other traffic can be configured with a higher priority than class of service (COS) five.

- ECN (explicit congestion notification) must be configured on all end-to-end paths.

Considerations for ISLs

When using shared ISLs in a MetroCluster IP configuration, you must be aware of the requirements for the end-to-end MetroCluster ISL running from controller ports on site A to controller ports on site B.

- **Basic MetroCluster ISL requirements**

The following requirements must be met:

- A native-speed ISL switch port must connect to a native-speed ISL switch port.
 For example, a 40 Gbps port connects to a 40 Gbps port.
- A 10 Gbps port that is in native mode (i.e., not using a breakout cable) can connect to a 10 Gbps port that is in native mode.
- The ISLs between the MetroCluster IP switches and the customer network, as well as the ISLs between the intermediate switches, follow the same rules in terms of speed.
- The number of ISLs that are between the MetroCluster switches and the customer network switches, and the number of ISLs that are between the customer network switches, do not need to match.
 For example, the MetroCluster switches can connect using two ISLs to the intermediate switches, and the intermediate switches can connect to each other using 10 ISLs.
- The speed of ISLs that are between the MetroCluster switches and the customer network switches, and the speed of ISLs that are between the customer network switches, do not need to match.
 For example, the MetroCluster switches can connect using a 40-Gbps ISL to the intermediate switches, and the intermediate switches can connect to each other using 100-Gbps ISLs.
- The number of and speed of ISLs connecting each MetroCluster switch to the intermediate switch must be the same on both MetroCluster sites.

- **Number of ISLs and breakout cables in the shared network**

The number of ISLs connecting the MetroCluster IP switches to the shared network varies depending on the switch model and port type.

MetroCluster IP switch model	Port type	Number of ISLs
Cisco 3132Q-V	Native ports	6 ISLs using 40 Gbps ports
	Breakout cables	16 x 10 Gbps ISLs
Cisco 3232C	Native ports	6 ISLs using 40 or 100 Gbps ports
	Breakout cables	16 x 10 Gbps ISLs

- The use of breakout cables (one physical port is used as 4 x 10 Gbps ports) is supported on Cisco switches.
- The RCF files for the IP switches have ports in native and breakout mode configured.
 A mix of ISL ports in native port speed mode and breakout mode is not supported. All ISLs from the MetroCluster IP switches to the intermediate switches in one network must be of same speed and length.
- The use of external encryption devices (for example, external link encryption or encryption provided via WDM devices) are supported as long as the round-trip latency remains within the above requirements.

For optimum performance, you should use at least a 1 x 40 Gbps or multiple 10 Gbps ISLs per network.

The maximum theoretical throughput of shared ISLs (for example, 240 Gbps with six 40 Gbps ISLs) is a best-case scenario. When using multiple ISLs, statistical load balancing can impact the maximum throughput. Uneven balancing can occur and reduce throughput to that of a single ISL.

If the configuration uses L2 VLANs, they must natively span the sites. VLAN overlay such as Virtual Extensible LAN (VXLAN) is not supported.

ISLs carrying MetroCluster traffic must be native links between the switches. Link sharing services such as Multiprotocol Label Switching (MPLS) links are not supported.

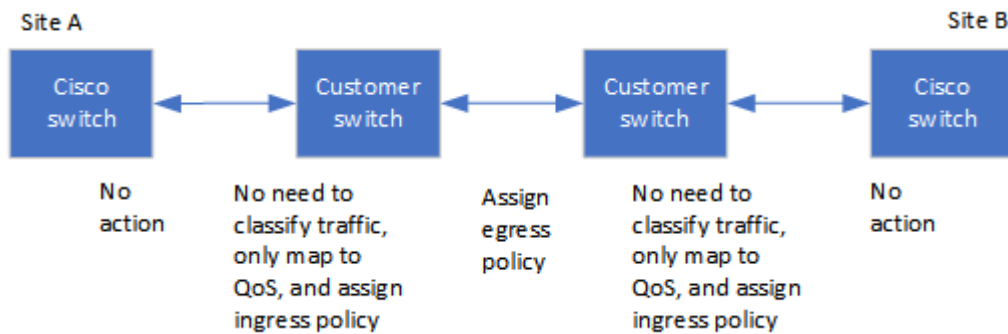
Required settings on intermediate switches

When sharing ISL traffic in a shared network, the configuration of the intermediate switches provided by the customer must ensure that the MetroCluster traffic (RDMA and storage) meets the required service levels across the entire path between the MetroCluster sites.

The following examples are for Cisco Nexus 3000 switches. Depending on your switch vendor and models, you must ensure that your intermediate switches have an equivalent configuration.

- Cisco Nexus switches

The following diagram gives an overview of the required settings for a shared network when the external switches are Cisco switches.



In this example, the following policies and maps are created for MetroCluster traffic:

- A MetroClusterIP_Ingress policy is applied to ports on the intermediate switch that connect to the MetroCluster IP switches.
The MetroClusterIP_Ingress policy maps the incoming tagged traffic to the appropriate queue on the intermediate switch. Tagging happens on the node-port, not on the ISL. Non-MetroCluster traffic that is using the same ports on the ISL remains in the default queue.
- A MetroClusterIP_Egress policy is applied to ports on the intermediate switch that connect to ISLs between intermediate switches

You must configure the intermediate switches with matching QoS access-maps, class-maps, and policy-maps along the path between the MetroCluster IP switches. The intermediate switches map RDMA traffic to COS5 and storage traffic to COS4.

The following example shows the configuration for a customer-provided Cisco Nexus 3000 switch. If you have Cisco switches, you can use the example to configure the switch along the path without much difficulty. If you do not have Cisco switches, you must determine and apply the equivalent configuration to your intermediate switches.

The following example shows the class map definitions:

Caution

This example is for configurations using Cisco MetroCluster IP switches. You can follow this example regardless of the switch types of the switches carrying MetroCluster traffic that do not connect to a MetroCluster IP switch.

```
class-map type qos match-all rdma
  match cos 5
class-map type qos match-all storage
  match cos 4
```

The following example shows the policy map definitions:

```
policy-map type qos MetroClusterIP_Ingress
  class rdma
    set dscp 40
    set cos 5
    set qos-group 5
  class storage
    set dscp 32
    set cos 4
    set qos-group 4
policy-map type queuing MetroClusterIP_Egress
  class type queuing c-out-8q-q7
    priority level 1
  class type queuing c-out-8q-q6
    priority level 2
  class type queuing c-out-8q-q5
    priority level 3
    random-detect threshold burst-optimized ecn
  class type queuing c-out-8q-q4
    priority level 4
    random-detect threshold burst-optimized ecn
  class type queuing c-out-8q-q3
    priority level 5
  class type queuing c-out-8q-q2
    priority level 6
  class type queuing c-out-8q-q1
    priority level 7
  class type queuing c-out-8q-q-default
    bandwidth remaining percent 100
    random-detect threshold burst-optimized ecn
```

- Intermediate customer switches
 - For intermediate customer switches, you must assign the egress policy to the ISL switch ports.
 - For all other interior switches along the path that carry MetroCluster traffic, follow the class map and policy map examples in the section *Cisco Nexus 3000 switches*.

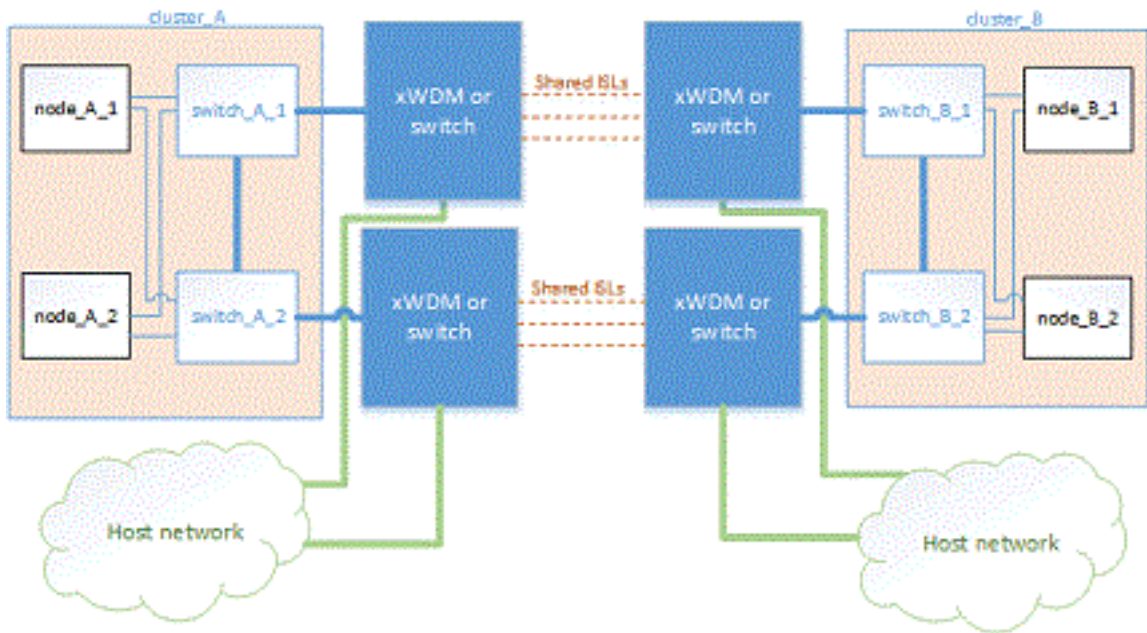
Examples of MetroCluster network topologies

- Shared network configuration with direct links

In this topology, two distinct sites are connected by direct links. These links can be between Wavelength Division Multiplexing equipment (xWDM) or switches. The capacity of the ISLs is not dedicated to the MetroCluster traffic but is shared with other traffic.

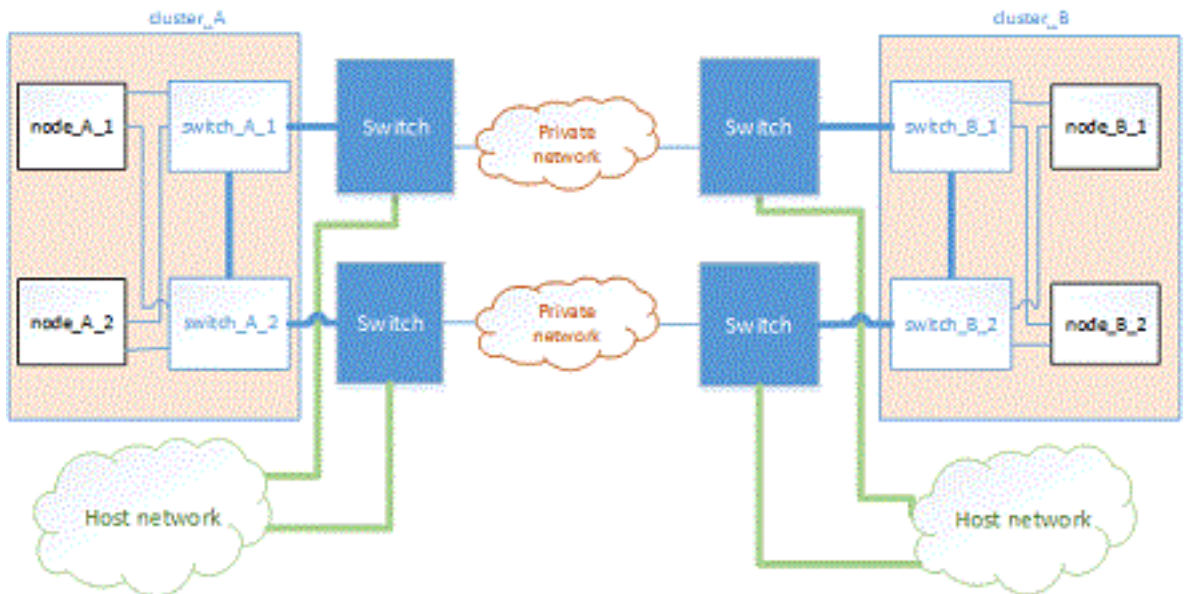
The ISL capacity must meet the minimum requirements. Depending on whether you use xWDM devices or switches a different combination of network configurations might apply.

2. Preparing for the MetroCluster installation
 Considerations for sharing private layer 2 networks



- Shared infrastructure with intermediate networks
 In this topology, the MetroCluster IP core switch traffic and the host traffic travel through a network that is not provided by Fujitsu. The network infrastructure and the links (including leased direct links) are outside of the MetroCluster configuration. The network can consist of a series of xWDM and switches but unlike the shared configuration with direct ISLs, the links are not direct between the sites. Depending on the infrastructure between the sites, any combination of network configurations is possible. The intermediate infrastructure is represented as a "cloud" (multiple devices can exist between the sites), but it is still under the control of the customer. Capacity through this intermediate infrastructure is not dedicated to the MetroCluster traffic but is shared with other traffic.

The VLAN and network xWDM or switch configuration must meet the minimum requirements.

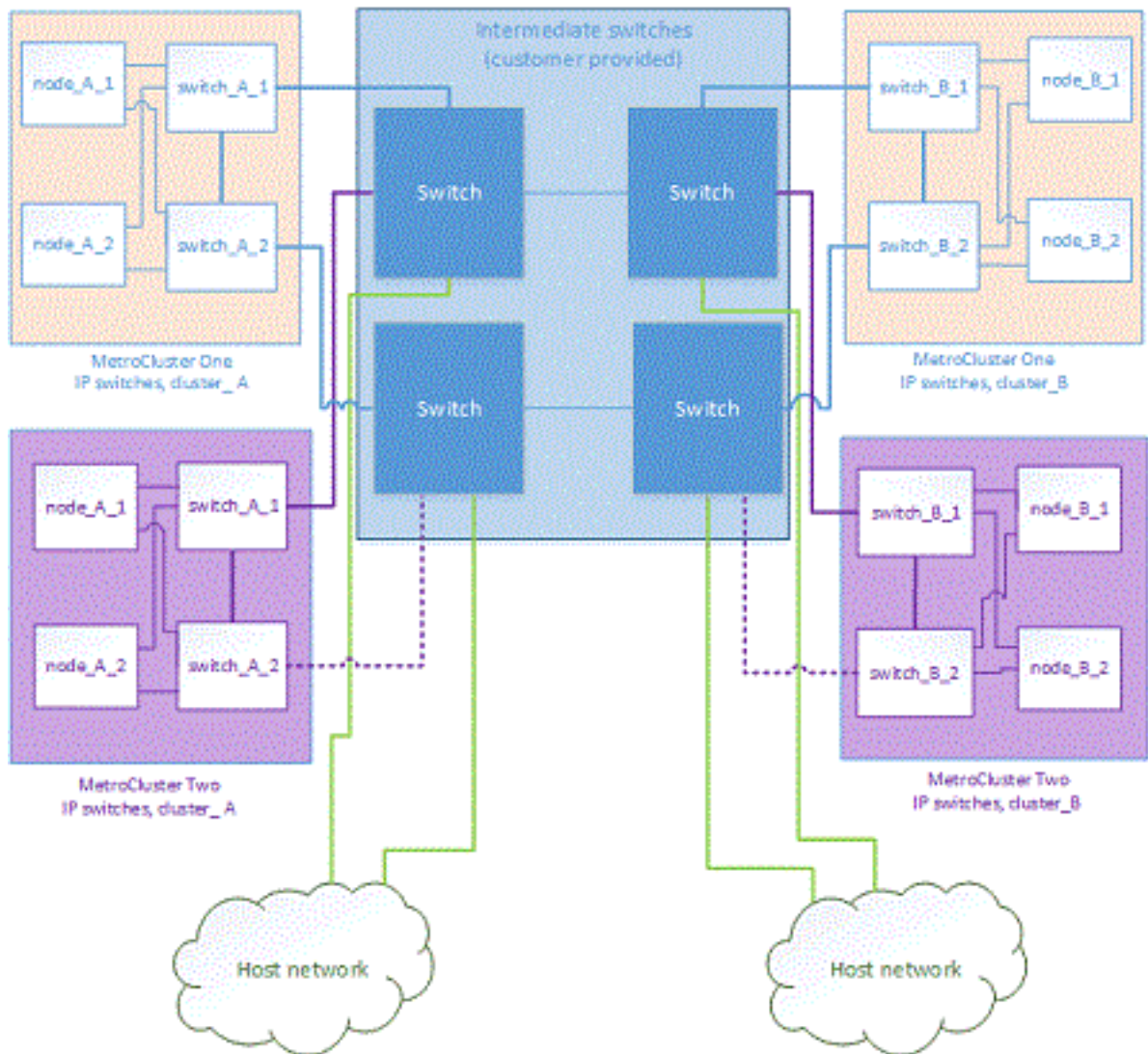


- Two MetroCluster configurations sharing an intermediate network
 In this topology, two separate MetroCluster configurations are sharing the same intermediate network. In the example, MetroCluster one switch_A_1 and MetroCluster two switch_A_1 both connect to the same intermediate switch.

The example is simplified for illustration purposes only:

2. Preparing for the MetroCluster installation

Considerations for sharing private layer 2 networks



- Two MetroCluster configurations with one connecting directly to the intermediate network
This topology is supported beginning with ONTAP 9.7. Two separate MetroCluster configurations share the same intermediate network and one MetroCluster configuration's nodes is directly connected to the intermediate switch.
MetroCluster Two is a MetroCluster configuration using Fujitsu compliant switches and ONTAP 9.7.

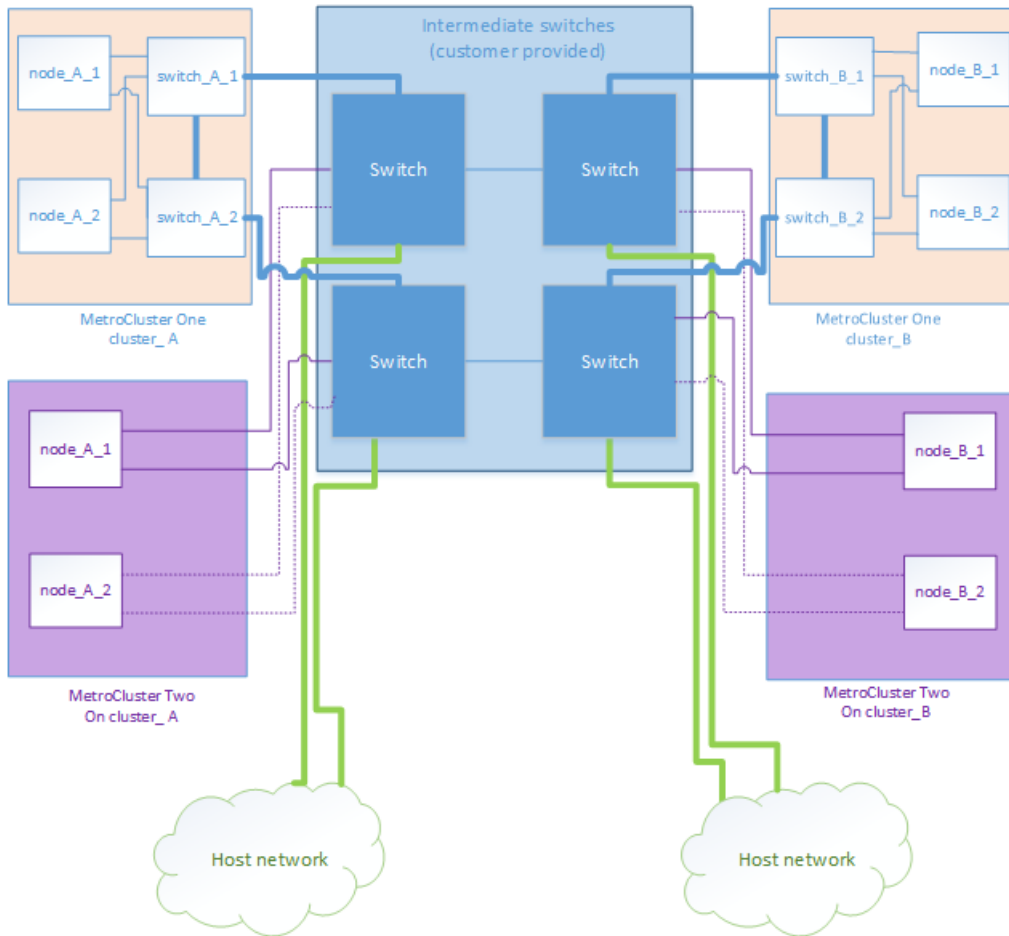
Caution

The intermediate switches must be compliant with Fujitsu specifications.

["Considerations for using MetroCluster compliant switches" \(page 24\)](#)

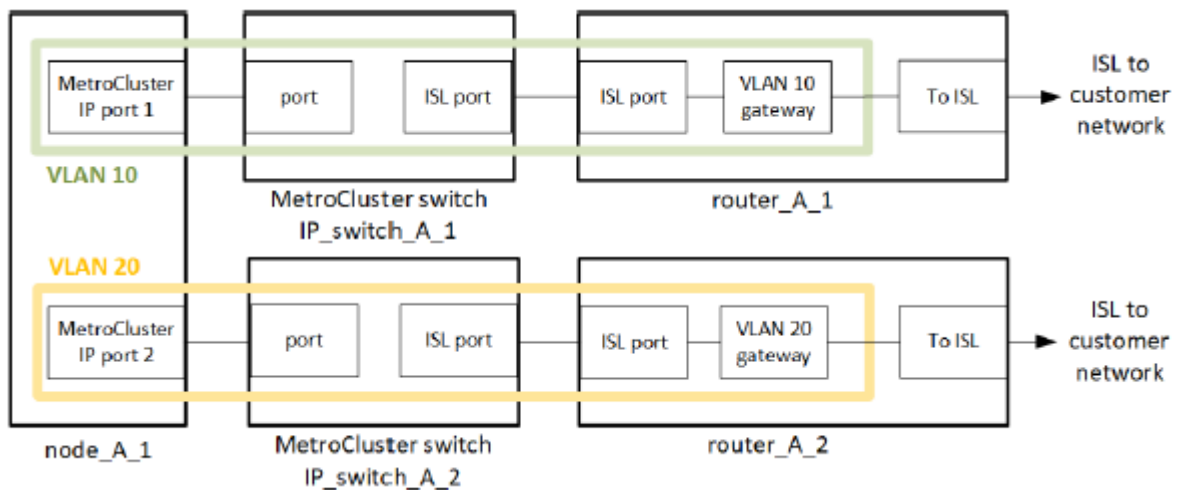
The example is simplified for illustration purposes only:

2. Preparing for the MetroCluster installation
 Considerations for layer 3 wide-area networks



Considerations for layer 3 wide-area networks

Starting with ONTAP 9.9.1, MetroCluster IP configurations can be implemented with IP-routed (layer 3) backend connections. The MetroCluster backend switches are connected to the routed IP network, either directly to routers (as shown in the following simplified example) or through other intervening switches.



Note

Fujitsu supports only Fujitsu -validated switches. These switches are tested and sold by Fujitsu. They are listed in "[Cabling the IP switches](#)" (page 39).

The MetroCluster environment is configured and cabled as a standard MetroCluster IP configuration as described in ["3. Configuring the MetroCluster hardware components" \(page 33\)](#). When you perform the installation and cabling procedure, you must perform the steps specific to the layer 3 configuration:

- The MetroCluster switches can be connected directly to the router or to one or more intervening switches. The VLAN must be extended to the gateway device.
- You use the `-gateway` parameter to configure the MetroCluster IP (MCC-IP) interface address with an IP gateway address.

When you configure routers and gateway IP addresses, ensure the following requirements are met:

- On each node, two interfaces cannot have the same gateway IP address.
- The corresponding interfaces on the HA pairs on each site must have the same gateway IP address.
- The corresponding interfaces on a node and its DR and AUX partners cannot have the same gateway IP address.
- The corresponding interfaces on a node and its DR and AUX partners must have the same VLAN ID.

The MetroCluster VLANs must extend from the edge MetroCluster switch to the gateway router so that MetroCluster traffic reaches the gateway (refer to the diagram shown above). The VLAN IDs for the MetroCluster VLANs must be the same at each site. However, the subnets can be different.

You use the RCF files that are created by the RcfFileGenerator tool. The network between the MetroCluster nodes and the gateway router must provide the same VLAN IDs as set in the RCF file.

- IP-routed network requirements

The IP-routed network must meet the following requirements:

- ["Basic MetroCluster ISL requirements" \(page 18\)](#)
- ["Considerations for sharing private layer 2 networks" \(page 16\)](#)
- ["Required settings on intermediate switches" \(page 19\)](#)
- Dynamic routing is not supported for the MetroCluster traffic.
- Only four-node MetroCluster configurations are supported (two nodes at each site).
- Two subnets are required on each MetroCluster site—one in each network.
- Auto-IP assignment is not supported.

Considerations for using MetroCluster compliant switches

MetroCluster IP switches provided by Fujitsu are Fujitsu validated. Beginning with ONTAP 9.7, MetroCluster IP configurations can support switches that are not Fujitsu validated provided that they are compliant with Fujitsu specifications.

- General requirements

The requirements show how to configure MetroCluster compliant switches without using reference configuration (RCF) files.

- Only platforms that provide dedicated ports for switchless cluster interconnects are supported. Platforms such as HX2200 and AX2100 are not supported because MetroCluster traffic and MetroCluster interconnect traffic share the same network ports.
Connecting local cluster connections to a MetroCluster compliant switch is not supported.
- The MetroCluster IP interface can be connected to any switch port that can be configured to meet the requirements.
- The speed of the switch ports must be 25 Gbps for HX6100 and AX3000 platforms, and at least 40 Gbps for all other platforms (40 Gbps or 100 Gbps).

2. Preparing for the MetroCluster installation

Considerations for using MetroCluster compliant switches

- The ISLs must be 10 Gbps or higher and must be sized appropriately for the load on the MetroCluster configuration.
 - The MetroCluster configuration must be connected to two networks. Connecting both the MetroCluster interfaces to the same network or switch is not supported. Each MetroCluster node must be connected to two network switches.
 - The network must meet the requirements as outlined in the sections *MetroCluster ISL requirements in shared networks*, *ISL cabling requirements*, and *Required settings on intermediate switches*.
 - In MetroCluster IP configurations using open networks, reverting to ONTAP 9.7 is not supported.
 - The MTU of 9216 must be configured on all switches that carry MetroCluster IP traffic.
- **Switch and cabling requirements**
 - The switches must support QoS/traffic classification.
 - The switches must support explicit congestion notification (ECN).
 - The switches must support L4 port-vlan load-balancing policies to preserve order along the path.
 - The switches must support L2 Flow Control (L2FC).
 - The cables connecting the nodes to the switches must be purchased from Fujitsu. The cables we provide must be supported by the switch vendor.
 - **Platform-specific network speeds for MetroCluster compliant switches**

The following table provides platform-specific network speeds for MetroCluster compliant switches.

Caution

Missing values indicate that the platform is not supported.

Platform	Network Speed (Gbps)
AX3000	25
AX2100	-
HX6100	25
HX2200	-

- **Assumptions for the examples**

The examples provided are valid for Cisco NX31xx and NX32xx switches. If other switches are used, these commands can be used as guidance, but the commands might be different. If a feature shown in the examples is not available on the switch, this means that the switch does not meet the minimum requirements and cannot be used to deploy a MetroCluster configuration. This is true for any switch that is connecting a MetroCluster configuration and for all switches on the path between those switches.

 - The ISL ports are 15 and 16 and operate at a speed of 40 Gbps.
 - The VLAN in network 1 is 10 and the VLAN in network 2 is 20. Examples might be shown for one network only.
 - The MetroCluster interface is connected to port 9 on each switch and operates at a speed of 100 Gbps.
 - The full context of the examples is not set or shown. You might need to enter further configuration information such as the profile, VLAN, or interface, to execute the commands.
- **Generic switch configuration**

A VLAN in each network must be configured. The example shows how to configure a VLAN in network 10.

Example:

```
# vlan 10
```

2. Preparing for the MetroCluster installation

Considerations for using MetroCluster compliant switches

The load balancing policy should be set so that order is preserved.

Example:

```
# port-channel load-balance src-dst ip-l4port-vlan
```

You must configure the access and class maps, which map the RDMA and iSCSI traffic to the appropriate classes.

All TCP traffic to and from the port 65200 is mapped to the storage (iSCSI) class. All TCP traffic to and from the port 10006 is mapped to the RDMA class.

Example:

```
ip access-list storage
 10 permit tcp any eq 65200 any
 20 permit tcp any any eq 65200
ip access-list rdma
 10 permit tcp any eq 10006 any
 20 permit tcp any any eq 10006

class-map type qos match-all storage
 match access-group name storage
class-map type qos match-all rdma
 match access-group name rdma
```

You must configure the ingress policy. The ingress policy maps the traffic as classified to the different COS groups. In this example, the RDMA traffic is mapped to COS group 5 and iSCSI traffic is mapped to COS group 4.

Example:

```
policy-map type qos MetroClusterIP_Ingress
class rdma
 set dscp 40
 set cos 5
 set qos-group 5
class storage
 set dscp 32
 set cos 4
 set qos-group 4
```

You must configure the egress policy on the switch. The egress policy maps the traffic to the egress queues. In this example, RDMA traffic is mapped to queue 5 and iSCSI traffic is mapped to queue 4.

2. Preparing for the MetroCluster installation

Considerations for using MetroCluster compliant switches

Example:

```
policy-map type queuing MetroClusterIP_Egress
class type queuing c-out-8q-q7
  priority level 1
class type queuing c-out-8q-q6
  priority level 2
class type queuing c-out-8q-q5
  priority level 3
  random-detect threshold burst-optimized ecn
class type queuing c-out-8q-q4
  priority level 4
  random-detect threshold burst-optimized ecn
class type queuing c-out-8q-q3
  priority level 5
class type queuing c-out-8q-q2
  priority level 6
class type queuing c-out-8q-q1
  priority level 7
class type queuing c-out-8q-q-default
  bandwidth remaining percent 100
  random-detect threshold burst-optimized ecn
```

You need to configure a switch that has MetroCluster traffic on an ISL but does not connect to any MetroCluster interfaces. In this case, the traffic is already classified and only needs to be mapped to the appropriate queue. In the following example, all of the COS5 traffic is mapped to the class RDMA, and all of the COS4 traffic is mapped to the class iSCSI. Note that this will affect all of the COS5 and COS4 traffic, not only the MetroCluster traffic. If you only want to map the MetroCluster traffic, then you must use the above class maps to identify the traffic using the access groups.

Example:

```
class-map type qos match-all rdma
  match cos 5
class-map type qos match-all storage
  match cos 4
```

● Configuring the ISLs

You can configure a 'trunk' mode port when setting an allowed VLAN.

There are two commands, one to **set** the allowed VLAN list, and one to **add** to the existing allowed VLAN list. You can set the allowed VLANs as shown in the example.

Example:

```
switchport trunk allowed vlan 10
```

You can **add** a VLAN to the allowed list as shown in the example.

Example:

```
switchport trunk allowed vlan add 10
```

In the example, port-channel 10 is configured for VLAN 10.

2. Preparing for the MetroCluster installation

Considerations for using MetroCluster compliant switches

Example:

```
interface port-channel10
switchport mode trunk
switchport trunk allowed vlan 10
mtu 9216
service-policy type queuing output MetroClusterIP_Egress
```

The ISL ports should be configured as part of a port-channel and be assigned the egress queues as shown in the example.

Example:

```
interface eth1/15-16
switchport mode trunk
switchport trunk allowed vlan 10
no lldp transmit
no lldp receive
mtu 9216
channel-group 10 mode active
service-policy type queuing output MetroClusterIP_Egress
no shutdown
```

- **Configuring the node ports**

You might need to configure the node port in breakout mode. In this example, ports 25 and 26 are configured in 4 x 25 Gbps breakout mode.

Example:

```
interface breakout module 1 port 25-26 map 25g-4x
```

You might need to configure the MetroCluster interface port speed. The example shows how to configure the speed to "auto".

Example:

```
speed auto
```

The following example shows how to fix the speed at 40 Gbps.

Example:

```
speed 40000
```

You might need to configure the interface. In the following example, the interface speed is set to "auto". The port is in access mode in VLAN 10, MTU is set to 9216 and the MetroCluster ingress policy is assigned.

Example:

```
interface eth1/9
description MetroCluster-IP Node Port
speed auto
switchport access vlan 10
spanning-tree port type edge
spanning-tree bpduguard enable
mtu 9216
flowcontrol receive on
flowcontrol send on
service-policy type qos input MetroClusterIP_Ingress
no shutdown
```

On 25-Gbps ports, the FEC setting might need to be set to "off" as shown in the example.

Example:

```
fec off
```

Caution

You must always run this command **after** the interface is configured. A transceiver module might need to be inserted for the command to work.

Considerations for using TDM/xWDM and encryption equipment with MetroCluster IP configurations

You should be aware of certain considerations for using multiplexing equipment in the MetroCluster IP configuration.

These considerations apply only to direct, dedicated MetroCluster back-end links and switches, not links shared with non-MetroCluster traffic.

● Using encryption on WDM or external encryption devices

When using encryption on WDM devices in the MetroCluster IP configuration, your environment must meet the following requirements:

- The external encryption devices or DWDM equipment must have been certified by the vendor with the switch in question.
The certification should cover the operating mode (such as trunking and encryption).
- The overall end-to-end latency and jitter, including the encryption, cannot be above the maximum stated in the Server Support Matrix or in this document.

● SFP considerations

Any SFPs or QSFPs supported by the equipment vendor are supported for the MetroCluster ISLs. SFPs and QSFPs can be acquired from Fujitsu or the equipment vendor.

● Considerations for ISLs

The ISLs on one fabric should all be the same speed and length.

The ISLs on one fabric should all have the same topology. For example, they should all be direct links, or if the configuration uses WDM, then they should all use WDM.

If you are sharing ISLs with a non-MetroCluster network, you must follow the guidelines in the section "[Considerations for sharing private layer 2 networks](#)" (page 16).

Considerations when using unmirrored aggregates

If your configuration includes unmirrored aggregates, you must be aware of potential access issues after switchover operations.

- **Considerations for unmirrored aggregates when doing maintenance requiring power shutdown**
If you are performing negotiated switchover for maintenance reasons requiring site-wide power shutdown, you should first manually take offline any unmirrored aggregates owned by the disaster site.
If you do not, nodes at the surviving site might go down due to multi-disk panics. This could occur if switched-over unmirrored aggregates go offline or are missing because of the loss of connectivity to storage at the disaster site due to the power shutdown or a loss of ISLs.
- **Considerations for unmirrored aggregates and hierarchical namespaces**
If you are using hierarchical namespaces, you should configure the junction path so that all of the volumes in that path are either on mirrored aggregates only or on unmirrored aggregates only. Configuring a mix of unmirrored and mirrored aggregates in the junction path might prevent access to the unmirrored aggregates after the switchover operation.
- **Considerations for unmirrored aggregates and CRS metadata volume and data SVM root volumes**
The configuration replication service (CRS) metadata volume and data SVM root volumes must be on a mirrored aggregate. You cannot move these volumes to unmirrored aggregate. If they are on unmirrored aggregate, negotiated switchover and switchback operations are vetoed. The `metrocluster check` command provides a warning if this is the case.
- **Considerations for unmirrored aggregates and SVMs**
SVMs should be configured on mirrored aggregates only or on unmirrored aggregates only. Configuring a mix of unmirrored and mirrored aggregates can result in a switchover operation that exceeds 120 seconds and result in a data outage if the unmirrored aggregates do not come online.
- **Considerations for unmirrored aggregates and SAN**
In ONTAP versions prior to 9.9.1, a LUN should not be located on an unmirrored aggregate. Configuring a LUN on an unmirrored aggregate can result in a switchover operation that exceeds 120 seconds and a data outage.
- **Considerations for adding storage shelves for unmirrored aggregates**

Caution

If you are adding shelves that will be used for unmirrored aggregates in a MetroCluster IP configuration, you must do the following:

- 1 Before starting the procedure to add the shelves, issue the following command: `metrocluster modify -enable-unmirrored-aggr-deployment true`
- 2 Verify that automatic disk assignment is off: `disk option show`
- 3 Follow the steps of the procedure to add the shelf.
- 4 Manually assign all disks from new shelf to the node that will own the unmirrored aggregate or aggregates.
- 5 Create the aggregates using the `storage aggregate create` command. .
- 6 After completing the procedure, issue the following command: `metrocluster modify -enable-unmirrored-aggr-deployment false`
- 7 Verify that automatic disk assignment is enabled: `disk option show`

Considerations for firewall usage at MetroCluster sites

If you are using a firewall at a MetroCluster site, you must ensure access for certain required ports. The following table shows TCP/UDP port usage in an external firewall positioned between two MetroCluster sites.

Traffic type	Port/services
Cluster peering	11104 / TCP 11105 / TCP
ONTAP System Manager	443 / TCP
MetroCluster IP intercluster LIFs	65200 / TCP 10006 / TCP and UDP
Hardware assist	4444 / TCP

Preconfigured settings for new MetroCluster systems from the factory

New MetroCluster nodes are preconfigured with a root aggregate. Additional hardware and software settings are configured using the detailed procedures provided in this guide.


- Hardware racking and cabling
Depending on the configuration you ordered, you might need to rack the systems and complete the cabling.
- Software configuration of the MetroCluster configuration
Nodes received with the new MetroCluster configuration are preconfigured with a single root aggregate. Additional configuration must be performed using the detailed procedures provided in this guide.

Hardware setup checklist

You need to know which hardware setup steps were completed at the factory and which steps you need to complete at each MetroCluster site.

Step	Completed at factory	Completed by you
Mount components in one or more cabinets.	Yes	No
Position cabinets in the desired location.	No	Yes Position them in the original order so that the supplied cables are long enough.
Connect multiple cabinets to each other, if applicable.	No	Yes Use the cabinet interconnect kit if it is included in the order. The kit box is labeled.
Secure the cabinets to the floor, if applicable.	No	Yes Use the universal bolt-down kit if it is included in the order. The kit box is labeled.

2. Preparing for the MetroCluster installation
 Preconfigured settings for new MetroCluster systems from the factory

Step	Completed at factory	Completed by you
Cable the components within the cabinet.	Yes Cables 5 meters and longer are removed for shipping and placed in the accessories box.	No
Connect the cables between cabinets, if applicable.	No	Yes Cables are in the accessories box.
Connect management cables to the customer's network.	No	Yes Connect them directly or through the CN1601 management switches, if present. <div style="border: 1px solid black; padding: 5px; display: inline-block;">  CAUTION </div> To avoid address conflicts, do not connect management ports to the customer's network until after you change the default IP addresses to the customer's values.
Connect console ports to the customer's terminal server, if applicable.	No	Yes
Connect the customer's data cables to the cluster.	No	Yes
Connect the cabinets to power and power on the components.	No	Yes Power them on in the following order: 1 PDUs 2 Disk shelves 3 Nodes
Assign IP addresses to the management ports of the cluster switches and to the management ports of the management switches, if present.	No	Yes Connect to the serial console port of each switch and log in with user name "admin" with no password. Suggested management addresses are 10.10.10.81, 10.10.10.82, 10.10.10.83, and 10.10.10.84.
Verify cabling by running the Config Advisor tool.	No	Yes

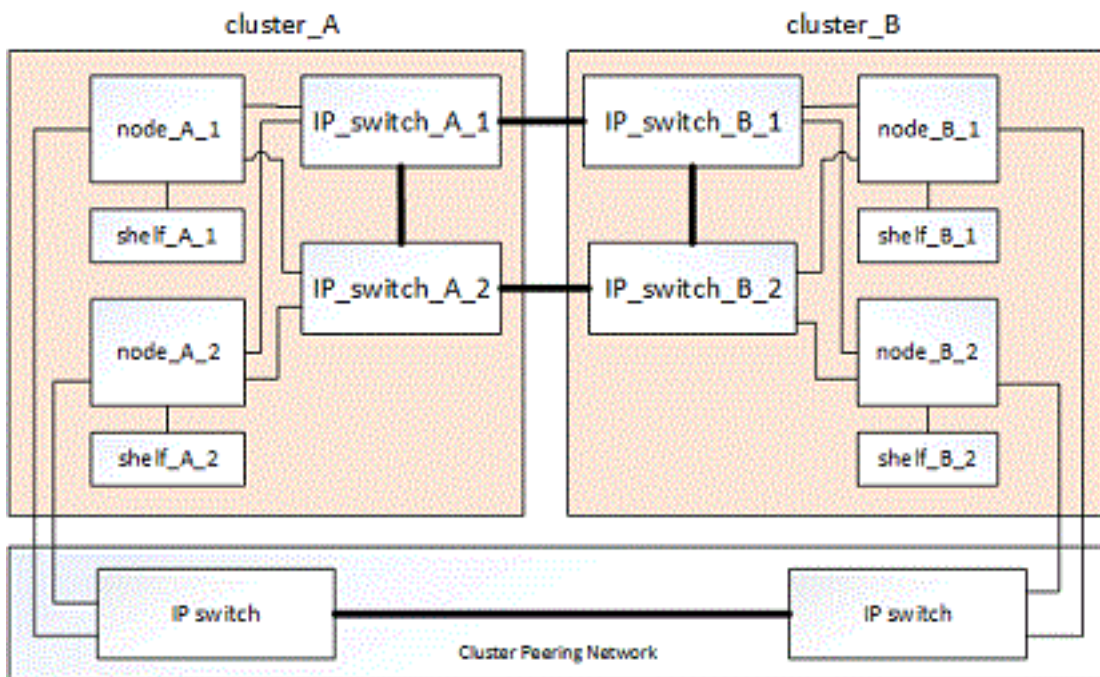
3. Configuring the MetroCluster hardware components

The MetroCluster components must be physically installed, cabled, and configured at both geographic sites.

Parts of a MetroCluster IP configuration

As you plan your MetroCluster IP configuration, you should understand the hardware components and how they interconnect.

- Key hardware elements
A MetroCluster IP configuration includes the following key hardware elements:
 - Storage controllers
The storage controllers are configured as two two-node clusters.
 - IP network
This back-end IP network provides connectivity for two distinct uses:
 - Standard cluster connectivity for intra-cluster communications.
This is the same cluster switch functionality used in non-MetroCluster switched ONTAP clusters.
 - MetroCluster back-end connectivity for replication of storage data and non-volatile cache.
 - Cluster peering network
The cluster peering network provides connectivity for mirroring of the cluster configuration, which includes storage virtual machine (SVM) configuration. The configuration of all of the SVMs on one cluster is mirrored to the partner cluster.



- Disaster Recovery (DR) groups
A MetroCluster IP configuration consists of one DR group of four nodes.
The following illustration shows the organization of nodes in a four-node MetroCluster configuration:

3. Configuring the MetroCluster hardware components
Parts of a MetroCluster IP configuration

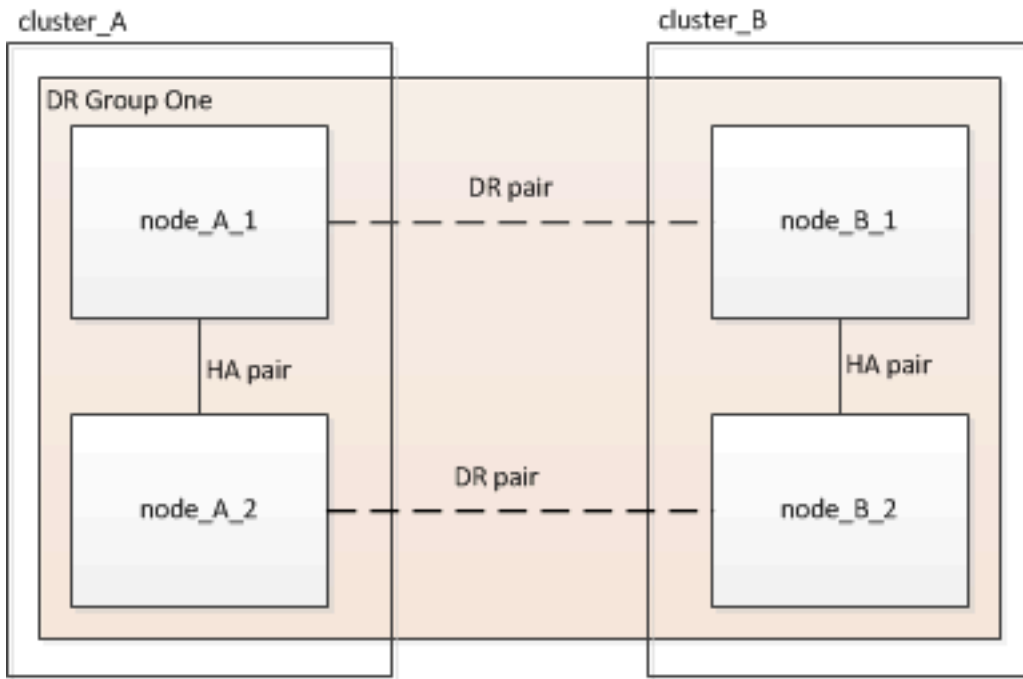
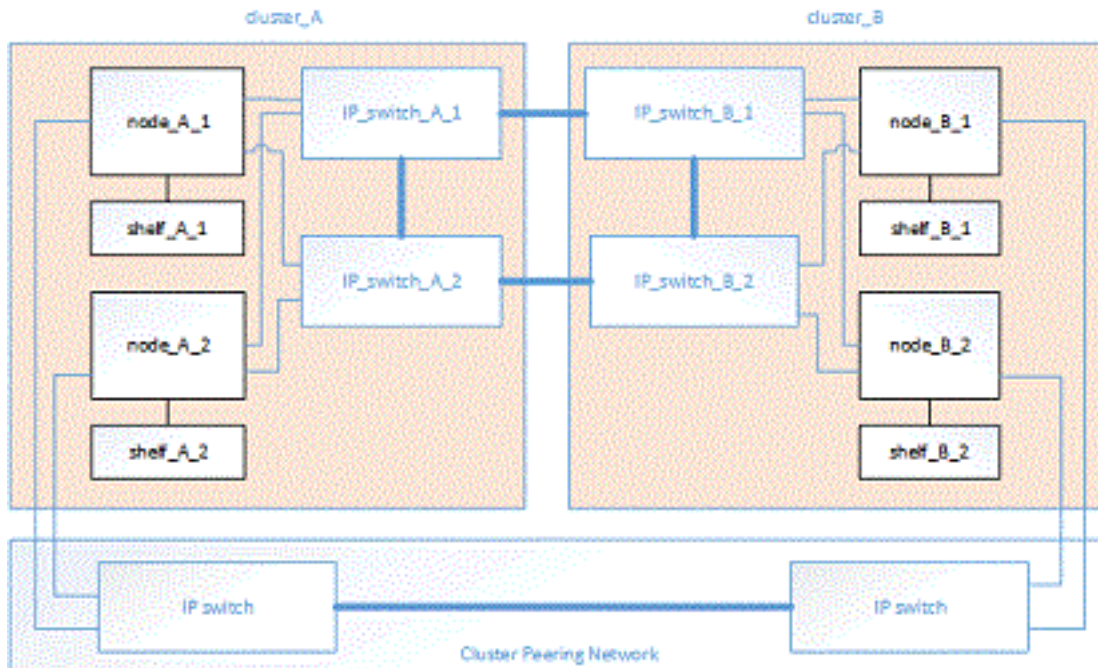


Illustration of the local HA pairs in a MetroCluster configuration

Each MetroCluster site consists of storage controllers configured as an HA pair. This allows local redundancy so that if one storage controller fails, its local HA partner can take over. Such failures can be handled without a MetroCluster switchover operation.

Local HA failover and giveback operations are performed with the `storage failover` commands, in the same manner as a non-MetroCluster configuration.

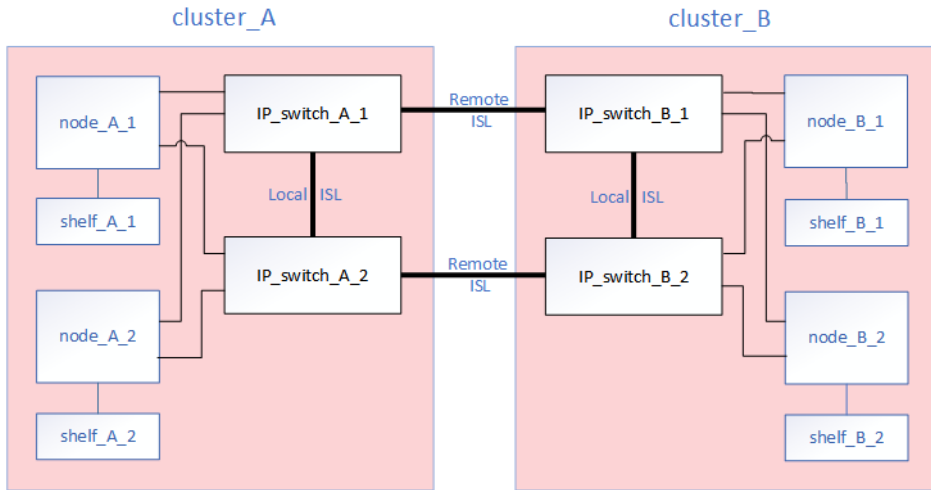


Related information

[Concepts Guide](#)

Illustration of the MetroCluster IP and cluster interconnect network

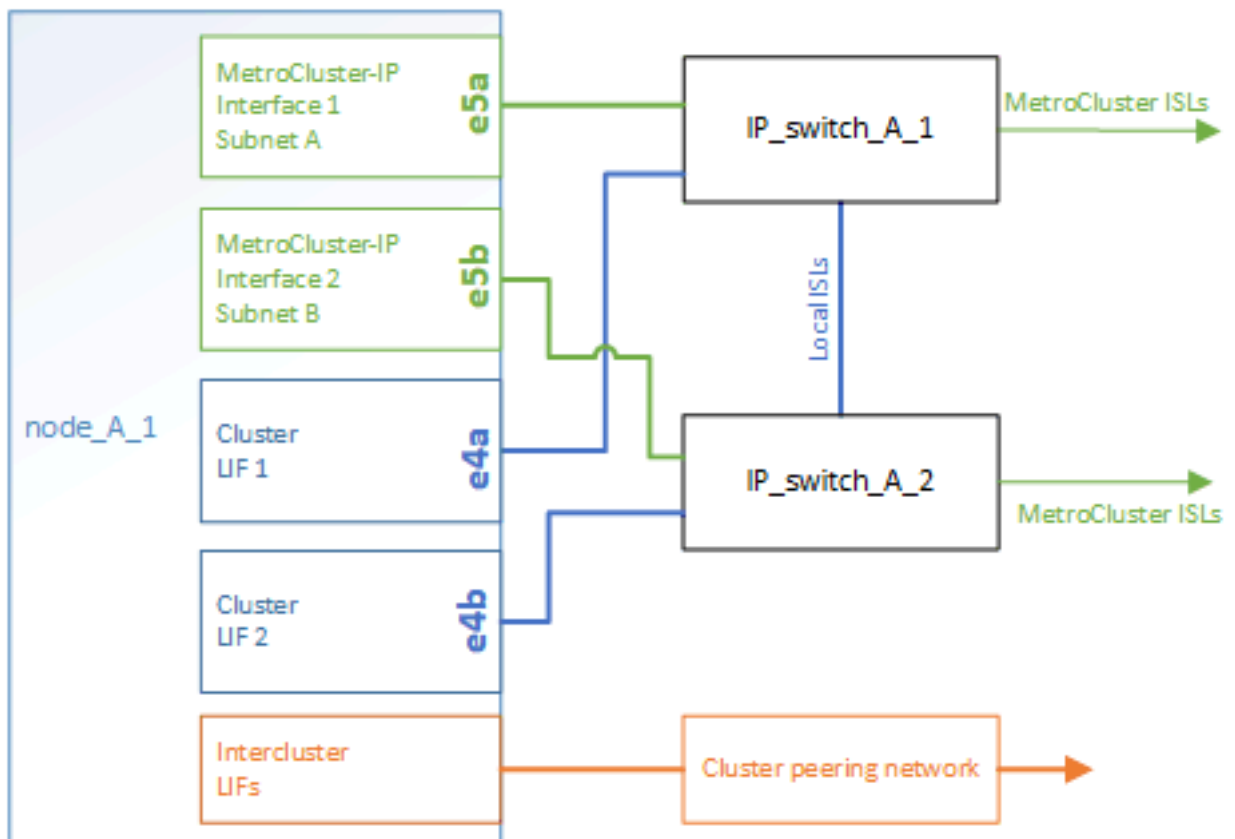
ONTAP clusters typically include a cluster interconnect network for traffic between the nodes in the cluster. In MetroCluster IP configurations, this network is also used for carrying data replication traffic between the MetroCluster sites.



Each node in the MetroCluster IP configuration has specialized LIFs for connection to the back-end IP network:

- Two MetroCluster IP interfaces
- One intercluster LIF

The following illustration shows these interfaces.



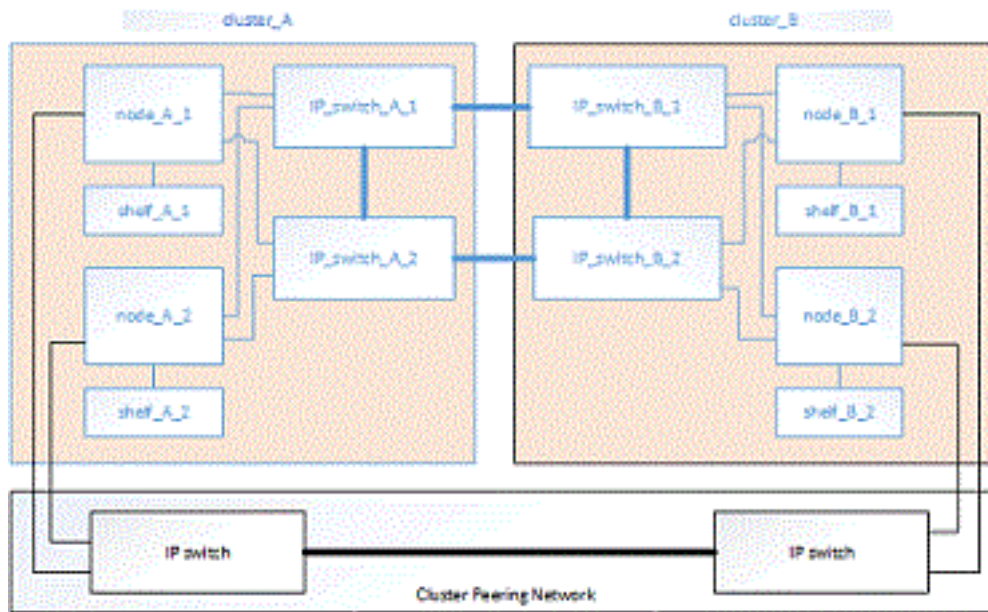
Related information

[Considerations for MetroCluster IP configuration \(page 9\)](#)

Illustration of the cluster peering network

The two clusters in the MetroCluster configuration are peered through a customer-provided cluster peering network. Cluster peering supports the synchronous mirroring of storage virtual machines (SVMs, formerly known as Vservers) between the sites.

Intercluster LIFs must be configured on each node in the MetroCluster configuration, and the clusters must be configured for peering. The ports with the intercluster LIFs are connected to the customer-provided cluster peering network. Replication of the SVM configuration is carried out over this network through the Configuration Replication Service.



Related information

- [Cluster and SVM Peering Power Guide](#)
- [Considerations for configuring cluster peering \(page 14\)](#)
- [Cabling the cluster peering connections \(page 44\)](#)
- [Peering the clusters \(page 70\)](#)

Required MetroCluster IP components and naming conventions

When planning your MetroCluster IP configuration, you must understand the required and supported hardware and software components. For convenience and clarity, you should also understand the naming conventions used for components in examples throughout the documentation.

- **Supported software and hardware**
The hardware and software must be supported for the MetroCluster IP configuration.
[NetApp Hardware Universe](#)
When using ETERNUS AX systems, all controller modules in the MetroCluster configuration must be configured as ETERNUS AX systems.
- **Hardware redundancy requirements in a MetroCluster IP configuration**
Because of the hardware redundancy in the MetroCluster IP configuration, there are two of each component at each site. The sites are arbitrarily assigned the letters A and B, and the individual components are arbitrarily assigned the numbers 1 and 2.

- **ONTAP cluster requirements in a MetroCluster IP configuration**
 MetroCluster IP configurations require two ONTAP clusters, one at each MetroCluster site.
 Example names:
 - Site A: cluster_A
 - Site B: cluster_B

- **IP switch requirements in a MetroCluster IP configuration**
 MetroCluster IP configurations require four IP switches. The four switches form two switch storage fabrics that provide the ISL between each of the clusters in the MetroCluster IP configuration.
 The IP switches also provide intracluster communication among the controller modules in each cluster.
 Example names:
 - Site A: cluster_A
 - IP_switch_A_1
 - IP_switch_A_2
 - Site B: cluster_B
 - IP_switch_B_1
 - IP_switch_B_2

- **Controller module requirements in a MetroCluster IP configuration**
 MetroCluster IP configurations require four controller modules.
 The controller modules at each site form an HA pair. Each controller module has a DR partner at the other site.

- **Gigabit Ethernet adapter requirements in a MetroCluster IP configuration**
 MetroCluster IP configurations use a 40/100 Gbps or 10/25 Gbps Ethernet adapter for the IP interfaces to the IP switches used for the MetroCluster IP fabric.

Platform model	Required Gigabit Ethernet adapter	Required slot for adapter	Ports
AX3000 and HX6100	X1116A	Slot 1	e1a, e1b
AX2100 and HX2200	Onboard ports	Slot 0	e0a, e0b

- **Pool and drive requirements (minimum supported)**
 Eight SAS disk shelves are recommended (four shelves at each site) to allow disk ownership on a per-shelf basis.
 A four-node MetroCluster IP configuration requires the minimum configuration at each site:
 - Each node has at least one local pool and one remote pool at the site.
 - At least seven drives in each pool.
 In a four-node MetroCluster configuration with a single mirrored data aggregate per node, the minimum configuration requires 28 disks at the site.
 In a minimum supported configuration, each pool has the following drive layout:
 - Three root drives
 - Three data drives
 - One spare drive
 In a minimum supported configuration, at least one shelf is needed per site.
 MetroCluster configurations support RAID-DP and RAID4.

- Drive location considerations for partially populated shelves
For correct auto-assignment of drives when using shelves that are half populated (12 drives in a 24-drive shelf), drives should be located in slots 0-5 and 18-23.
In a configuration with a partially populated shelf, the drives must be evenly distributed in the four quadrants of the shelf.
- Mixing IOM12 and IOM 6 modules in a stack
Your version of ONTAP must support shelf mixing. Refer to the Server Support Matrix to see if your version of ONTAP supports shelf mixing. Please contact Fujitsu support personnel for Server Support Matrix.
For further details on shelf mixing, see: [Hot-adding shelves with IOM12 modules to a stack of shelves with IOM6 modules](#)

Related information

[Considerations for automatic drive assignment and ADP systems in ONTAP 9.7 and later \(page 10\)](#)

Installing and cabling MetroCluster components

The storage controllers must be cabled to the IP switches and the ISLs must be cabled to link the MetroCluster sites. The storage controllers must also be cabled to the storage, to each other, and to the data and management networks.

Racking the hardware components

If you have not received the equipment already installed in cabinets, you must rack the components.

About this task

This task must be performed on both MetroCluster sites.

Procedure ▶▶▶

- 1 **Plan out the positioning of the MetroCluster components.**
The rack space depends on the platform model of the controller modules, the switch types, and the number of disk shelf stacks in your configuration.
- 2 **Properly ground yourself.**
- 3 **Install the controller modules in the rack or cabinet.**
- 4 **Install the IP switches in the rack or cabinet.**
- 5 **Install the disk shelves, power them on, and then set the shelf IDs.**
[ETERNUS AX/HX Series Manual Page](#)
 - You must power-cycle each disk shelf.
 - Shelf IDs must be unique for each SAS disk shelf within each MetroCluster DR group (including both sites).

Caution

Do not cable disk shelves intended to contain unmirrored aggregates at this time. You must wait to deploy shelves intended for unmirrored aggregates until after the MetroCluster configuration is complete and only deploy them after using the `metrocluster modify -enable-unmirrored-aggr-deployment true` command.

Cabling the IP switches

You must cable each IP switch to the local controller modules and to the ISLs.

About this task

- This task must be repeated for each switch in the MetroCluster configuration.
- The controller module Ethernet port usage depends on the model of the controller module.

Procedure ▶▶▶

- 1 Cable the switch and node ports, using the table for your switch model and platform.
-

Using the port tables with the RcfFileGenerator tool or multiple MetroCluster configurations

You must understand how to use the information in the port tables to correctly generate your RCF files.

Review these considerations before using the tables:

- The following tables show the port usage for site A. The same cabling is used for site B.
 - The switches cannot be configured with ports of different speeds (for example, a mix of 100 Gbps ports and 40 Gbps ports).
 - Keep track of the MetroCluster port group (MetroCluster 1, MetroCluster 2, etc.). You will need it when using the RcfFileGenerator tool as described later in this configuration procedure.
 - The RcfFileGenerator for MetroCluster IP also provides a per-port cabling overview for each switch. Use this cabling overview to verify your cabling.
- **Cabling eight-node MetroCluster configurations**

For MetroCluster configuration running ONTAP 9.8 and earlier, some procedures that are performed to transition an upgrade require the addition of a second four-node DR group to the configuration to create a temporary eight-node configuration. Starting with ONTAP 9.9.1, permanent 8-node MetroCluster configurations are supported.

For such configurations, you use the same method as described above. Instead of a second MetroCluster, you are cabling an additional four-node DR group

For example, your configuration includes the following:

 - Cisco 3132Q-V switches
 - MetroCluster 1: HX2200 platforms

For MetroCluster 1, cable the Cisco 3132Q-V switches using the table for the HX2200 platform and the rows for MetroCluster 1 interfaces.

Platform port assignments for Cisco 3132Q-V switches

The port usage in a MetroCluster IP configuration depends on the switch model and platform type.

3. Configuring the MetroCluster hardware components
 Installing and cabling MetroCluster components

- Port usage for HX2200 or AX2100 systems and a Cisco 3132Q-V switch

Cabling an AX2100 or HX2200 to a Cisco 3132Q-V switch			
Port use	HX2200, AX2100		Switch Port
	IP_switch_x_1	IP_switch_x_2	
Unused	-		1
			2
			3
			4
			5
			6
ISL, Local Cluster native speed / 40G / 100G	ISL, Local Cluster		7
			8
MetroCluster 1, Shared Cluster and MetroCluster interface	e0a	e0b	9/1
	disabled		9/2-4
	e0a	e0b	10/1
	disabled		10/2-4
MetroCluster 2, Shared Cluster and MetroCluster interface	e0a	e0b	11/1
	disabled		11/2-4
	e0a	e0b	12/1
	disabled		12/2-4
MetroCluster 3, Shared Cluster and MetroCluster interface	e0a	e0b	13/1
	disabled		13/2-4
	e0a	e0b	14/1
	disabled		14/2-4
ISL, MetroCluster native speed 40G	ISL, MetroCluster		15 - 20
ISL, MetroCluster breakout mode 10G	ISL, MetroCluster		21/1-4
			22/1-4
			23/1-4
			24/1-4
Unused	-		25 - 32

Platform port assignments for Cisco 3232C or Cisco 9336C switches

The port usage in a MetroCluster IP configuration depends on the switch model and platform type.

Review these considerations before using the tables:

- The following tables show the port usage for site A. The same cabling is used for site B.
- The switches cannot be configured with ports of different speeds (for example, a mix of 100 Gbps ports and 40 Gbps ports).
- If you are configuring a single MetroCluster with the switches, use the **MetroCluster 1** port group. Keep track of the MetroCluster port group (MetroCluster 1, MetroCluster 2, or MetroCluster 3). You will need it when using the RcfFileGenerator tool as described later in this configuration procedure.

3. Configuring the MetroCluster hardware components
 Installing and cabling MetroCluster components

- The RcfFileGenerator for MetroCluster IP also provides a per-port cabling overview for each switch. Use this cabling overview to verify your cabling.
- Cabling two MetroCluster configurations to the switches
 When cabling more than one MetroCluster configuration to a Cisco 3132Q-V switch, then cable each MetroCluster according to the appropriate table. For example, if cabling a HX2200 and an A700 to the same Cisco 3132Q-V switch. Then you cable the HX2200 as per 'MetroCluster 1' in Table 1, and the A700 as per 'MetroCluster 2' or 'MetroCluster 3' in Table 2. You cannot physically cable both the HX2200 and A700 as 'MetroCluster 1'.
- Port usage for HX2200, AX2100, systems

Cabling an AX2100 or HX2200 to a Cisco 3232C or Cisco 9336C switch			
Port use	HX2200, AX2100		Switch port
	IP_switch_x_1	IP_switch_x_2	
Unused	-		1 - 6
ISL, Local Cluster native speed / 100G	ISL, Local Cluster		7
			8
MetroCluster 1, Shared Cluster and MetroCluster interface	e0a	e0b	9/1
	disabled		9/2-4
	e0a	e0b	10/1
	disabled		10/2-4
MetroCluster 2, Shared Cluster and MetroCluster interface	e0a	e0b	11/1
	disabled		11/2-4
	e0a	e0b	12/1
	disabled		12/2-4
MetroCluster 3, Shared Cluster and MetroCluster interface	e0a	e0b	13/1
	disabled		13/2-4
	e0a	e0b	14/1
	disabled		14/2-4
ISL, MetroCluster native speed 40G / 100G			15
			16
	ISL, MetroCluster		17
			18
			19
			20
ISL, MetroCluster breakout mode 10G			21/1-4
			22/1-4
			23/1-4
			24/1-4
Unused	-		25 - 32

3. Configuring the MetroCluster hardware components
 Installing and cabling MetroCluster components

- Cabling a AX3000 or HX6100 to a Cisco 3232C or Cisco 9336C switch

Cabling a AX3000 or HX6100 to a Cisco 3232C or Cisco 9336C switch			
Port use	HX6100, AX3000		Switch port
	IP_switch_x_1	IP_switch_x_2	
MetroCluster 1 Local Cluster interface	See Hardware Universe for available ports		1/1
			1/2 - 4
			2/1
			2/2 - 4
MetroCluster 2 Local Cluster interface			3/1
			3/2 - 4
			4/1
			4/2 - 4
MetroCluster 3 Local Cluster interface			5/1
			5/2 - 4
			6/1
			6/2 - 4
ISL, Local Cluster native speed / 100G	ISL, Local Cluster		7
			8
MetroCluster 1 MetroCluster interface	e1a	e1b	9/1
	disabled		9/2-4
	e1a	e1b	10/1
	disabled		10/2-4
MetroCluster 2 MetroCluster interface	e1a	e1b	11/1
	disabled		11/2-4
	e1a	e1b	12/1
	disabled		12/2-4
MetroCluster 3 MetroCluster interface	e1a	e1b	13/1
	disabled		13/2-4
	e1a	e1b	14/1
	disabled		14/2-4
ISL, MetroCluster	ISL, MetroCluster		15 - 20
ISL, MetroCluster breakout mode 10G	ISL, MetroCluster		21/1-4
			22/1-4
			23/1-4
			24/1-4
MetroCluster 4 MetroCluster interface	e1a	e1b	25/1
	disabled		25/2-4
	e1a	e1b	26/1
	disabled		26/2-4
Unused	-		27 - 28
MetroCluster 4 Local Cluster interface	See Hardware Universe		29/1
	disabled		29/2-4
	See Hardware Universe		30/1
	disabled		30/2-4
Unused	-		31 - 32

3. Configuring the MetroCluster hardware components
 Installing and cabling MetroCluster components

- Cabling a AX2200 or FAS500f to a Cisco 3232C or Cisco 9336C switch

Cabling an AX2200 or FAS500f to a Cisco 3232C or Cisco 9336C switch			
Port use	FAS500f, AX2200		Switch port
	IP_switch_x_1	IP_switch_x_2	
Unused	-		1 - 6
ISL, Local Cluster native speed / 100G	ISL, Local Cluster		7
			8
MetroCluster 1, Shared Cluster and MetroCluster interface	e0c	e0d	9/1
	disabled		9/2-4
	e0c	e0d	10/1
	disabled		10/2-4
MetroCluster 2, Shared Cluster and MetroCluster interface	e0c	e0d	11/1
	disabled		11/2-4
	e0c	e0d	12/1
	disabled		12/2-4
MetroCluster 3, Shared Cluster and MetroCluster interface	e0c	e0d	13/1
	disabled		13/2-4
	e0c	e0d	14/1
	disabled		14/2-4
ISL, MetroCluster native speed 40G / 100G	ISL, MetroCluster		15
			16
			17
			18
			19
ISL, MetroCluster breakout mode 10G	ISL, MetroCluster		20
			21/1-4
			22/1-4
			23/1-4
Unused	-		24/1-4
			25 - 32

3. Configuring the MetroCluster hardware components
 Installing and cabling MetroCluster components

- Cabling an AX4100 or HX6100 to a Cisco 3232C or Cisco 9336C switch

Cabling a AX4100 or HX6100 to a Cisco 3232C or Cisco 9336C switch			
Port use	HX6100, AX4100		Switch port
	IP_switch_x_1	IP_switch_x_2	
MetroCluster 1, Local Cluster interface	See Hardware Universe for available ports		1
			2
			3
4			
MetroCluster 2, Local Cluster interface			5
MetroCluster 3, Local Cluster interface			6
ISL, Local Cluster native speed / 100G	ISL, Local Cluster		7
			8
MetroCluster 1, MetroCluster interface	e1a	e1b	9
	e1a	e1b	10
MetroCluster 2, MetroCluster interface	e1a	e1b	11
	e1a	e1b	12
MetroCluster 3, MetroCluster interface	e1a	e1b	13
	e1a	e1b	14
ISL, MetroCluster native speed 40G / 100G	ISL, MetroCluster		15
			16
			17
			18
			19
ISL, MetroCluster breakout mode 10G	ISL, MetroCluster		20
			21/1-4
			22/1-4
			23/1-4
Unused	-		24/1-4
			25
			26
			27
			28
			29
			30
			31
32			

Cabling the cluster peering connections

You must cable the controller module ports used for cluster peering so that they have connectivity with the cluster on the partner site.

About this task

This task must be performed on each controller module in the MetroCluster configuration.

At least two ports on each controller module should be used for cluster peering.

The recommended minimum bandwidth for the ports and network connectivity is 1 GbE.

Procedure ▶▶▶

- 1 Identify and cable at least two ports for cluster peering and verify they have network connectivity with the partner cluster.
Cluster peering can be done on dedicated ports or on data ports. Using dedicated ports provides higher throughput for the cluster peering traffic.
[Cluster and SVM Peering Power Guide](#)

Related information

- [Cluster and SVM Peering Power Guide](#)
- [Considerations for configuring cluster peering \(page 14\)](#)

Cabling the management and data connections

You must cable the management and data ports on each storage controller to the site networks.

About this task

This task must be repeated for each new controller at both MetroCluster sites.

You can connect the controller and cluster switch management ports to existing switches in your network or to new dedicated network switches such as Fujitsu CN1601 cluster management switches.

Procedure ▶▶▶

- 1 Cable the controller's management and data ports to the management and data networks at the local site.

Configuring the IP switches

You must configure the IP switches for use as the cluster interconnect and for back-end MetroCluster IP connectivity. The procedure you use depends on the switch model.

Configuring Cisco IP switches

You must configure the Cisco IP switches for use as the cluster interconnect and for backend MetroCluster IP connectivity.

Resetting the Cisco IP switch to factory defaults

Before installing a new software version and RCFs, you must erase the Cisco switch configuration and perform basic configuration.

About this task

You must repeat these steps on each of the IP switches in the MetroCluster IP configuration.

Procedure ▶▶▶

- 1 Reset the switch to factory defaults:
 - (1) Erase the existing configuration: `write erase`
 - (2) Reload the switch software: `reload`

3. Configuring the MetroCluster hardware components

Installing and cabling MetroCluster components

The system reboots and enters the configuration wizard. During the boot, if you receive the prompt `Abort Auto Provisioning and continue with normal setup?(yes/no) [n]`, you should respond `yes` to proceed.

(3) In the configuration wizard, enter the basic switch settings:

- Admin password
- Switch name
- Out-of-band management configuration
- Default gateway
- SSH service (RSA)

After completing the configuration wizard, the switch reboots.

(4) When prompted, enter the user name and password to log in to the switch.

The following example shows the prompts and system responses when configuring the switch. The angle brackets (`<<<`) show where you enter the information.

```
---- System Admin Account Setup ----
Do you want to enforce secure password standard (yes/no) [y]:y <<<

Enter the password for "admin": password <<<
Confirm the password for "admin": password <<<
---- Basic System Configuration Dialog VDC: 1 ----

This setup utility will guide you through the basic configuration of
the system. Setup configures only enough connectivity for management
of the system.

Please register Cisco Nexus3000 Family devices promptly with your
supplier. Failure to register may affect response times for initial
service calls. Nexus3000 devices must be registered to receive
entitled support services.

Press Enter at anytime to skip a dialog. Use ctrl-c at anytime
to skip the remaining dialogs.
```

3. Configuring the MetroCluster hardware components

Installing and cabling MetroCluster components

You enter basic information in the next set of prompts, including the switch name, management address, and gateway, and select SSH with RSA.

```
Would you like to enter the basic configuration dialog (yes/no): yes
Create another login account (yes/no) [n]:
Configure read-only SNMP community string (yes/no) [n]:
Configure read-write SNMP community string (yes/no) [n]:
Enter the switch name : switch-name <<<
Continue with Out-of-band (mgmt0) management configuration? (yes/no) [y]:
  Mgmt0 IPv4 address : management-IP-address <<<
  Mgmt0 IPv4 netmask : management-IP-netmask <<<
Configure the default gateway? (yes/no) [y]: y <<<
  IPv4 address of the default gateway : gateway-IP-address <<<
Configure advanced IP options? (yes/no) [n]:
Enable the telnet service? (yes/no) [n]:
Enable the ssh service? (yes/no) [y]: y <<<
  Type of ssh key you would like to generate (dsa/rsa) [rsa]: rsa <<<
  Number of rsa key bits <1024-2048> [1024]:
Configure the ntp server? (yes/no) [n]:
Configure default interface layer (L3/L2) [L2]:
Configure default switchport interface state (shut/noshut) [noshut]: shut <<<
Configure CoPP system profile (strict/moderate/lenient/dense) [strict]:
```

3. Configuring the MetroCluster hardware components Installing and cabling MetroCluster components

The final set of prompts completes the configuration:

```
The following configuration will be applied:
  password strength-check
  switchname IP_switch_A_1
vrf context management
ip route 0.0.0.0/0 10.10.99.1
exit
  no feature telnet
  ssh key rsa 1024 force
  feature ssh
  system default switchport
  system default switchport shutdown
  copp profile strict
interface mgmt0
ip address 10.10.99.10 255.255.255.0
no shutdown

Would you like to edit the configuration? (yes/no) [n]:

Use this configuration and save it? (yes/no) [y]:
2017 Jun 13 21:24:43 A1 %$ VDC-1 %$ %COPP-2-COPP_POLICY: Control-Plane is protected with policy copp-system-p-
policy-strict.

[#####] 100%
Copy complete.

User Access Verification
IP_switch_A_1 login: admin
Password:
Cisco Nexus Operating System (NX-OS) Software
.
.
.
IP_switch_A_1#
```

2 Save the configuration:

```
IP_switch-A-1# copy running-config startup-config
```

3 Reboot the switch and wait for the switch to reload:

```
IP_switch-A-1# reload
```

4 Repeat the previous steps on the other three switches in the MetroCluster IP configuration.

Downloading and installing the Cisco switch NX-OS software

You must download the switch operating system file and RCF file to each switch in the MetroCluster IP configuration.

Before you begin

This task requires file transfer software, such as FTP, TFTP, SFTP, or SCP, to copy the files to the switches.

About this task

3. Configuring the MetroCluster hardware components Installing and cabling MetroCluster components

These steps must be repeated on each of the IP switches in the MetroCluster IP configuration. You must use the supported switch software version.

Procedure ►►►

- 1 **Download the supported NX-OS software file.**
- 2 **Copy the switch software to the switch:** `copy sftp://root@server-ip-address/tftpboot/NX-OS-file-name bootflash: vrf management`

In this example, the `nxos.7.0.3.I4.6.bin` file is copied from SFTP server 10.10.99.99 to the local bootflash:

```
IP_switch_A_1# copy sftp://root@10.10.99.99/tftpboot/nxos.7.0.3.I4.6.bin bootflash: vrf management
root@10.10.99.99's password: password
sftp> progress
Progress meter enabled
sftp> get /tftpboot/nxos.7.0.3.I4.6.bin /bootflash/nxos.7.0.3.I4.6.bin
Fetching /tftpboot/nxos.7.0.3.I4.6.bin to /bootflash/nxos.7.0.3.I4.6.bin
/tftpboot/nxos.7.0.3.I4.6.bin          100% 666MB 7.2MB/s 01:32
sftp> exit
Copy complete, now saving to disk (please wait)...
```

- 3 **Verify on each switch that the switch NX-OS files are present in each switch's bootflash directory:** `dir bootflash:`

The following example shows that the files are present on IP_switch_A_1:

```
IP_switch_A_1# dir bootflash:
.
.
.
698629632   Jun 13 21:37:44 2017  nxos.7.0.3.I4.6.bin
.
.
.

Usage for bootflash://sup-local
 1779363840 bytes used
13238841344 bytes free
15018205184 bytes total
IP_switch_A_1#
```

- 4 **Install the switch software:** `install all nxos bootflash:nxos.version-number.bin`
The switch will reload (reboot) automatically after the switch software has been installed.

3. Configuring the MetroCluster hardware components Installing and cabling MetroCluster components

The following example shows the software installation on IP_switch_A_1:

```
IP_switch_A_1# install all nxos bootflash:nxos.7.0.3.I4.6.bin
Installer will perform compatibility check first. Please wait.
Installer is forced disruptive

Verifying image bootflash:/nxos.7.0.3.I4.6.bin for boot variable "nxos".
[#####] 100% -- SUCCESS

Verifying image type.
[#####] 100% -- SUCCESS

Preparing "nxos" version info using image bootflash:/nxos.7.0.3.I4.6.bin.
[#####] 100% -- SUCCESS

Preparing "bios" version info using image bootflash:/nxos.7.0.3.I4.6.bin.
[#####] 100% -- SUCCESS      [#####] 100%      -- SUCCESS

Performing module support checks.      [#####] 100%      -- SUCCESS

Notifying services about system upgrade.      [#####] 100%      -- SUCCESS

Compatibility check is done:
Module  bootable          Impact  Install-type  Reason
-----  -----  -----  -----  -----
      1      yes      disruptive      reset  default upgrade is not hitless

Images will be upgraded according to following table:
Module      Image      Running-Version(pri:alt)      New-Version      Upg-Required
-----  -----  -----  -----  -----
      1      nxos      7.0(3)I4(1)      7.0(3)I4(6)      yes
      1      bios      v04.24(04/21/2016)      v04.24(04/21/2016)      no

Switch will be reloaded for disruptive upgrade.
Do you want to continue with the installation (y/n)? [n] y
```

3. Configuring the MetroCluster hardware components

Installing and cabling MetroCluster components

```
Install is in progress, please wait.

Performing runtime checks.          [#####] 100%  -- SUCCESS

Setting boot variables.
[#####] 100% -- SUCCESS

Performing configuration copy.
[#####] 100% -- SUCCESS

Module 1: Refreshing compact flash and upgrading bios/loader/bootrom.
Warning: please do not remove or power off the module at this time.
[#####] 100% -- SUCCESS

Finishing the upgrade, switch will reboot in 10 seconds.
IP_switch_A_1#
```

- 5** Wait for the switch to reload and then log in to the switch.
After the switch has rebooted the login prompt is displayed:

```
User Access Verification
IP_switch_A_1 login: admin
Password:
Cisco Nexus Operating System (NX-OS) Software
TAC support: http://www.cisco.com/tac
Copyright (C) 2002-2017, Cisco and/or its affiliates.
All rights reserved.
.
.
.
MDP database restore in progress.
IP_switch_A_1#

The switch software is now installed.
```

- 6** Verify that the switch software has been installed: `show version`

3. Configuring the MetroCluster hardware components

Installing and cabling MetroCluster components

The following example shows the output:

```
IP_switch_A_1# show version
Cisco Nexus Operating System (NX-OS) Software
TAC support: http://www.cisco.com/tac
Copyright (C) 2002-2017, Cisco and/or its affiliates.
All rights reserved.
.
.
.

Software
  BIOS: version 04.24
  NXOS: version 7.0(3)I4(6)   <<< switch software version
  BIOS compile time: 04/21/2016
  NXOS image file is: bootflash:///nxos.7.0.3.I4.6.bin
  NXOS compile time: 3/9/2017 22:00:00 [03/10/2017 07:05:18]

Hardware
  cisco Nexus 3132QV Chassis
  Intel(R) Core(TM) i3- CPU @ 2.50GHz with 16401416 kB of memory.
  Processor Board ID FOC20123GPS

  Device name: A1
  bootflash: 14900224 kB
  usb1: 0 kB (expansion flash)

Kernel uptime is 0 day(s), 0 hour(s), 1 minute(s), 49 second(s)

Last reset at 403451 usecs after Mon Jun 10 21:43:52 2017

Reason: Reset due to upgrade
System version: 7.0(3)I4(1)
Service:

plugin
  Core Plugin, Ethernet Plugin
IP_switch_A_1#
```

7 Repeat these steps on the remaining three IP switches in the MetroCluster IP configuration.

Downloading and installing the Cisco IP RCF files

You must download the RCF file to each switch in the MetroCluster IP configuration.

Before you begin

This task requires file transfer software, such as FTP, TFTP, SFTP, or SCP, to copy the files to the switches.

About this task

These steps must be repeated on each of the IP switches in the MetroCluster IP configuration.

You must use the supported switch software version.

There are four RCF files, one for each of the four switches in the MetroCluster IP configuration. You must use the correct RCF files for the switch model you are using.

3. Configuring the MetroCluster hardware components

Installing and cabling MetroCluster components

Switch	RCF file
IP_switch_A_1	NX3232_v1.8T-X1_Switch-A1.txt
IP_switch_A_2	NX3232_v1.8T-X1_Switch-A2.txt
IP_switch_B_1	NX3232_v1.8T-X1_Switch-B1.txt
IP_switch_B_2	NX3232_v1.8T-X1_Switch-B2.txt

Procedure ▶▶▶

1 Download the MetroCluster IP RCF files.

2 Copy the RCF files to the switches:

(1) **Copy the RCF files to the first switch:** `copy sftp://root@FTP-server-IP-address/tftpboot/switch-specific-RCF bootflash: vrf management`

In this example, the `NX3232_v1.8T-X1_Switch-A1.txt` RCF file is copied from the SFTP server at `10.10.99.99` to the local bootflash. You must use the IP address of your TFTP/SFTP server and the file name of the RCF file that you need to install.

```
IP_switch_A_1# copy sftp://root@10.10.99.99/tftpboot/NX3232_v1.8T-X1_Switch-A1.txt bootflash: vrf
management
root@10.10.99.99's password: password
sftp> progress
Progress meter enabled
sftp> get /tftpboot/NX3232_v1.8T-X1_Switch-A1.txt /bootflash/NX3232_v1.8T-X1_Switch-A1.txt
Fetching /tftpboot/NX3232_v1.8T-X1_Switch-A1.txt to /bootflash/NX3232_v1.8T-X1_Switch-A1.txt
/tftpboot/NX3232_v1.8T-X1_Switch-A1.txt      100% 5141      5.0KB/s   00:00
sftp> exit
Copy complete, now saving to disk (please wait)...
IP_switch_A_1#
```

(2) Repeat the previous substep for each of the other three switches, being sure to copy the matching RCF file to the corresponding switch.

3 Verify on each switch that the RCF file is present in each switch's `bootflash` directory: `dir bootflash:`

The following example shows that the files are present on `IP_switch_A_1`:

```
IP_switch_A_1# dir bootflash:
.
.
.
5514   Jun 13 22:09:05 2017  NX3232_v1.8T-X1_Switch-A1.txt
.
.
.

Usage for bootflash://sup-local
 1779363840 bytes used
13238841344 bytes free
15018205184 bytes total
IP_switch_A_1#
```

4 Copy the matching RCF file from the local bootflash to the running configuration on each switch: `copy bootflash:switch-specific-RCF.txt running-config`

- 5 **Copy the RCF files from the running configuration to the startup configuration on each switch:** `copy running-config startup-config`

You should see output similar to the following:

```
IP_switch_A_1# copy bootflash:NX3232_v1.8T-X1_Switch-A1.txt running-config
IP_switch-A-1# copy running-config startup-config
```

- 6 **Reload the switch:** `reload`

```
IP_switch_A_1# reload
```

- 7 **Repeat the previous steps on the other three switches in the MetroCluster IP configuration.**



Setting Forward Error Correction for AX3000 and HX6100 systems using 25-Gbps connectivity

If your AX3000 or HX6100 system is configured using 25-Gbps connectivity, you need to set the Forward Error Correction (`fec`) parameter manually to `off` after applying the RCF file. The RCF file does not apply this setting.

Before you begin

The 25-Gbps ports must be cabled prior to performing this procedure.

["Platform port assignments for Cisco 3232C or Cisco 9336C switches" \(page 40\)](#)

About this task

This task only applies to AX3000 and HX6100 platforms using 25-Gbps connectivity.

This task must be performed on all four switches in the MetroCluster IP configuration.

Procedure ▶▶▶

- 1 **Set the `fec` parameter to `off` on each 25-Gbps port that is connected to a controller module, and then copy the running configuration to the startup configuration:**

- (1) **Enter configuration mode:** `conf t`
- (2) **Specify the 25-Gbps interface to configure:** `interface interface-ID`
- (3) **Set `fec` to `off`:** `fec off`
- (4) **Repeat the previous steps for each 25-Gbps port on the switch.**
- (5) **Exit configuration mode:** `exit`

The following example shows the commands for interface Ethernet1/25/1 on switch IP_switch_A_1:

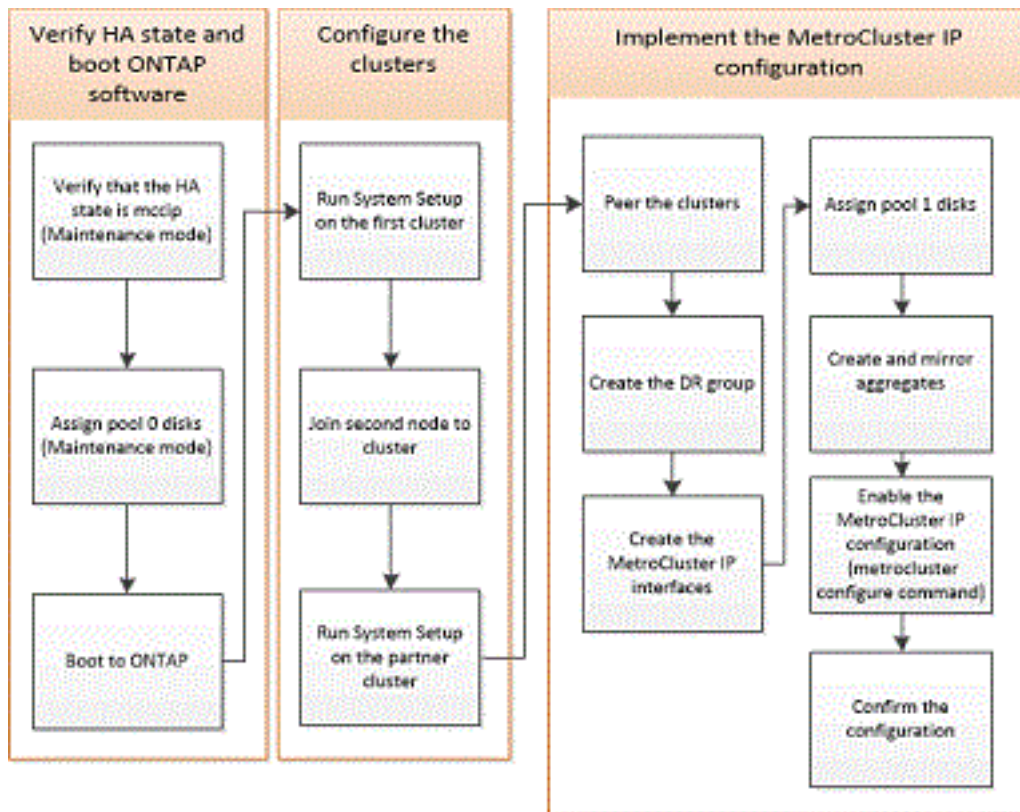
```
IP_switch_A_1# conf t
IP_switch_A_1(config)# interface Ethernet1/25/1
IP_switch_A_1(config-if)# fec off
IP_switch_A_1(config-if)# exit
IP_switch_A_1(config-if)# end
IP_switch_A_1# copy running-config startup-config
```

- 2 **Repeat the previous step on the other three switches in the MetroCluster IP configuration.**



4. Configuring the MetroCluster software in ONTAP

You must set up each node in the MetroCluster configuration in ONTAP, including the node-level configurations and the configuration of the nodes into two sites. You must also implement the MetroCluster relationship between the two sites.



Gathering required information

You need to gather the required IP addresses for the controller modules before you begin the configuration process.

IP network information worksheet for site A

You must obtain IP addresses and other network information for the first MetroCluster site (site A) from your network administrator before you configure the system.

- Site A switch information

When you cable the system, you need a host name and management IP address for each cluster switch.

Cluster switch	Host name	IP address	Network mask	Default gateway
Interconnect 1				
Interconnect 2				
Management 1				
Management 2				

- Site A cluster creation information

When you first create the cluster, you need the following information:

4. Configuring the MetroCluster software in ONTAP
Gathering required information

Type of information	Your values
Cluster name Example used in this guide: site_A	
DNS domain	
DNS name servers	
Location	
Administrator password	

● Site A node information

For each node in the cluster, you need a management IP address, a network mask, and a default gateway.

Node	Port	IP address	Network mask	Default gateway
Node 1 Example used in this guide: controller_A_1				
Node 2 Example used in this guide: controller_A_2				

● Site A LIFs and ports for MetroCluster IP back-end connectivity

For each node in the cluster, you need the IP addresses of two MetroCluster IP LIFs, including a network mask and a default gateway. The MetroCluster IP LIFs are used for MetroCluster IP back-end connectivity. ["Considerations for MetroCluster IP configuration" \(page 9\)](#)

Node	Port	IP address of MetroCluster IP LIF	Network mask	Default gateway
Node 1 MetroCluster IP LIF 1	e5a			
Node 1 MetroCluster IP LIF 2	e5b			
Node 2 MetroCluster IP LIF 1	e5a			
Node 2 MetroCluster IP LIF 2	e5b			

● Site A LIFs and ports for cluster peering

For each node in the cluster, you need the IP addresses of two intercluster LIFs, including a network mask and a default gateway. The intercluster LIFs are used to peer the clusters.

Node	Port	IP address of intercluster LIF	Network mask	Default gateway
Node 1 IC LIF 1				
Node 1 IC LIF 2				
Node 2 IC LIF 1				
Node 2 IC LIF 2				

- Site A time server information

You must synchronize the time, which requires one or more NTP time servers.

Node	Host name	IP address	Network mask	Default gateway
NTP server 1				
NTP server 2				

- Site A AutoSupport information

You must configure AutoSupport on each node, which requires the following information:

Type of information	Your values	
From email address		
Mail hosts	IP addresses or names	
Transport protocol	HTTP, HTTPS, or SMTP	
	Proxy server	
Recipient email addresses or distribution lists	Full-length messages	
	Concise messages	
	Partners	

- Site A SP information

You must enable access to the Service Processor (SP) of each node for troubleshooting and maintenance, which requires the following network information for each node:

Node	IP address	Network mask	Default gateway
Node 1			

IP network information worksheet for site B

You must obtain IP addresses and other network information for the second MetroCluster site (site B) from your network administrator before you configure the system.

- Site B switch information

When you cable the system, you need a host name and management IP address for each cluster switch.

Cluster switch	Host name	IP address	Network mask	Default gateway
Interconnect 1				
Interconnect 2				
Management 1				
Management 2				

- Site B cluster creation information

When you first create the cluster, you need the following information:

Type of information	Your values
Cluster name Example used in this guide: site_B	
DNS domain	

4. Configuring the MetroCluster software in ONTAP
Gathering required information

Type of information	Your values
DNS name servers	
Location	
Administrator password	

● Site B node information

For each node in the cluster, you need a management IP address, a network mask, and a default gateway.

Node	Port	IP address	Network mask	Default gateway
Node 1 Example used in this guide: controller_B_1				
Node 2 Example used in this guide: controller_B_2				

● Site B LIFs and ports for MetroCluster IP back-end connectivity

For each node in the cluster, you need the IP addresses of two MetroCluster IP LIFs, including a network mask and a default gateway. The MetroCluster IP LIFs are used for MetroCluster IP back-end connectivity. ["Considerations for MetroCluster IP configuration" \(page 9\)](#)

Node	Port	IP address of MetroCluster IP LIF	Network mask	Default gateway
Node 1 MetroCluster IP LIF 1	e5a			
Node 1 MetroCluster IP LIF 2	e5b			
Node 2 MetroCluster IP LIF 1	e5a			
Node 2 MetroCluster IP LIF 2	e5b			

● Site B LIFs and ports for cluster peering

For each node in the cluster, you need the IP addresses of two intercluster LIFs, including a network mask and a default gateway. The intercluster LIFs are used to peer the clusters.

Node	Port	IP address of intercluster LIF	Network mask	Default gateway
Node 1 IC LIF 1				
Node 1 IC LIF 2				
Node 2 IC LIF 1				
Node 2 IC LIF 2				

● Site B time server information

You must synchronize the time, which requires one or more NTP time servers.

4. Configuring the MetroCluster software in ONTAP
 Similarities and differences between standard cluster and MetroCluster configurations

Node	Host name	IP address	Network mask	Default gateway
NTP server 1				
NTP server 2				

● Site B AutoSupport information

You must configure AutoSupport on each node, which requires the following information:

Type of information	Your values	
From email address		
Mail hosts	IP addresses or names	
Transport protocol	HTTP, HTTPS, or SMTP	
	Proxy server	
Recipient email addresses or distribution lists	Full-length messages	
	Concise messages	
	Partners	

● Site B SP information

You must enable access to the Service Processor (SP) of each node for troubleshooting and maintenance, which requires the following network information for each node:

Node	IP address	Network mask	Default gateway
Node 1 (controller_B_1)			

Similarities and differences between standard cluster and MetroCluster configurations

The configuration of the nodes in each cluster in a MetroCluster configuration is similar to that of nodes in a standard cluster.

The MetroCluster configuration is built on two standard clusters. Physically, the configuration must be symmetrical, with each node having the same hardware configuration, and all of the MetroCluster components must be cabled and configured. However, the basic software configuration for nodes in a MetroCluster configuration is the same as that for nodes in a standard cluster.

Configuration step	Standard cluster configuration	MetroCluster configuration
Configure management, cluster, and data LIFs on each node.	Same in both types of clusters	
Configure the root aggregate.	Same in both types of clusters	
Set up the cluster on one node in the cluster.	Same in both types of clusters	
Join the other node to the cluster.	Same in both types of clusters	
Create a mirrored root aggregate.	Optional	Required
Peer the clusters.	Optional	Required
Enable the MetroCluster configuration.	Does not apply	Required

Restoring system defaults on a previously used controller module

If your controller modules have been used previously, you must reset them for a successful MetroCluster configuration.

About this task

IMPORTANT

This task is required only on controller modules that have been previously configured. You do not need to perform this task if you received the controller modules from the factory.

Procedure ▶▶▶

- 1 At the **LOADER** prompt, return the environmental variables to their default setting: `set-defaults`
- 2 Boot the node to the boot menu: `boot_ontap menu`
After you run the command, wait until the boot menu is shown.
- 3 Clear the node configuration:
 - If you are using systems configured for ADP, select option 9a from the boot menu, and respond `yes` when prompted.

Caution

This process is disruptive.

The following screen shows the boot menu prompt:

```
Please choose one of the following:

(1) Normal Boot.
(2) Boot without /etc/rc.
(3) Change password.
(4) Clean configuration and initialize all disks.
(5) Maintenance mode boot.
(6) Update flash from backup config.
(7) Install new software first.
(8) Reboot node.
(9) Configure Advanced Drive Partitioning.
Selection (1-9)? 9a
##### WARNING #####

This is a disruptive operation and will result in the
loss of all filesystem data. Before proceeding further,
make sure that:

1) This option (9a) has been executed or will be executed
on the HA partner node, prior to reinitializing either
system in the HA-pair.
2) The HA partner node is currently in a halted state or
at the LOADER prompt.

Do you still want to continue (yes/no)? yes
```

- If your system is not configured for ADP, type `wipeconfig` at the boot menu prompt, and then press Enter.

The following screen shows the boot menu prompt:

```
Please choose one of the following:

(1) Normal Boot.
(2) Boot without /etc/rc.
(3) Change password.
(4) Clean configuration and initialize all disks.
(5) Maintenance mode boot.
(6) Update flash from backup config.
(7) Install new software first.
(8) Reboot node.
(9) Configure Advanced Drive Partitioning.
Selection (1-9)? wipeconfig

This option deletes critical system configuration, including cluster membership.
Warning: do not run this option on a HA node that has been taken over.
Are you sure you want to continue?: yes
Rebooting to finish wipeconfig request.
```



Verifying the ha-config state of components

In a MetroCluster IP configuration that is not preconfigured at the factory, you must verify that the ha-config state of the controller and chassis components is set to `mccip` so that they boot up properly. For systems received from the factory, this value is preconfigured and you do not need to verify it.

Before you begin

The system must be in Maintenance mode.

Procedure ▶▶▶

- 1 Display the HA state of the controller module and chassis:** `ha-config show`
The controller module and chassis should show the value `mccip`.
- 2 If the displayed system state of the controller is not `mccip`, set the HA state for the controller:** `ha-config modify controller mccip`
- 3 If the displayed system state of the chassis is not `mccip`, set the HA state for the chassis:** `ha-config modify chassis mccip`
- 4 Repeat these steps on each node in the MetroCluster configuration.**



Manually assigning drives to pool 0

If you did not receive the systems pre-configured from the factory, you might have to manually assign the pool 0 drives. Depending on the platform model and whether the system is using ADP, you must manually assign drives to pool 0 for each node in the MetroCluster IP configuration. The procedure you use depends on the version of ONTAP you are using.

Manually assigning drives for pool 0 (ONTAP 9.7 and later)

If the system has not been pre-configured at the factory and does not meet the requirements for automatic drive assignment, you must manually assign the pool 0 drives.

About this task

This procedure applies to configurations running ONTAP 9.7 or later.

To determine if your system requires manual disk assignment, you should review ["Considerations for automatic drive assignment and ADP systems in ONTAP 9.7 and later"](#) (page 10).

You perform these steps in Maintenance mode. The procedure must be performed on each node in the configuration.

Examples in this section are based on the following assumptions:

- node_A_1 and node_A_2 own drives on:
 - site_A-shelf_1 (local)
 - site_B-shelf_2 (remote)
- node_B_1 and node_B_2 own drives on:
 - site_B-shelf_1 (local)
 - site_A-shelf_2 (remote)

Procedure ▶▶▶

1 Display the boot menu: `boot_ontap` menu

2 Select option 9a.

The following screen shows the boot menu prompt:

```
Please choose one of the following:

(1) Normal Boot.
(2) Boot without /etc/rc.
(3) Change password.
(4) Clean configuration and initialize all disks.
(5) Maintenance mode boot.
(6) Update flash from backup config.
(7) Install new software first.
(8) Reboot node.
(9) Configure Advanced Drive Partitioning.
Selection (1-9)? 9a
##### WARNING #####

This is a disruptive operation and will result in the
loss of all filesystem data. Before proceeding further,
make sure that:

1) This option (9a) has been executed or will be executed
on the HA partner node (and DR/DR-AUX partner nodes if
applicable), prior to reinitializing any system in the
HA-pair (or MCC setup).

2) The HA partner node (and DR/DR-AUX partner nodes if
applicable) is currently waiting at the boot menu.

Do you still want to continue (yes/no)? yes
```

- 3 When the node restarts, press Ctrl-C when prompted to display the boot menu and then select the option for [Maintenance mode boot].
- 4 In Maintenance mode, manually assign drives for the local aggregates on the node: `disk assign disk-id -p 0 -s local-node-sysid`
The drives should be assigned symmetrically, so each node has an equal number of drives. The following steps are for a configuration with two storage shelves at each site.
 - (1) When configuring node_A_1, manually assign drives from slot 0 to 11 to pool0 of node A1 from site_A-shelf_1.
 - (2) When configuring node_A_2, manually assign drives from slot 12 to 23 to pool0 of node A2 from site_A-shelf_1.
 - (3) When configuring node_B_1, manually assign drives from slot 0 to 11 to pool0 of node B1 from site_B-shelf_1.
 - (4) When configuring node_B_2, manually assign drives from slot 12 to 23 to pool0 of node B2 from site_B-shelf_1.
- 5 Exit Maintenance mode: `halt`
- 6 Display the boot menu: `boot_ontap menu`
- 7 Select option 4 from the boot menu and let the system boot.
- 8 Repeat these steps on the other nodes in the MetroCluster IP configuration.
- 9 Proceed to ["Setting up ONTAP" \(page 63\)](#).



Setting up ONTAP

After you boot each node, you are prompted to perform basic node and cluster configuration. After configuring the cluster, you return to the ONTAP CLI to create aggregates and create the MetroCluster configuration.

Before you begin

- You must have cabled the MetroCluster configuration.
- You must not have configured the Service Processor.

If you need to netboot the new controllers, see [Netbooting the new controller modules](#) in the *MetroCluster Upgrade, Transition, and Expansion Guide*.

About this task

This task must be performed on both clusters in the MetroCluster configuration.

Procedure ►►►

- 1 Power up each node at the local site if you have not already done so and let them all boot completely. If the system is in Maintenance mode, you need to issue the `halt` command to exit Maintenance mode, and then issue the `boot_ontap` command to boot the system and get to cluster setup.
- 2 On the first node in each cluster, proceed through the prompts to configure the cluster
 - (1) Enable the AutoSupport tool by following the directions provided by the system.

The output should be similar to the following:

```
Welcome to the cluster setup wizard.

You can enter the following commands at any time:
"help" or "?" - if you want to have a question clarified,
"back" - if you want to change previously answered questions, and
"exit" or "quit" - if you want to quit the cluster setup wizard.
Any changes you made before quitting will be saved.

You can return to cluster setup at any time by typing "cluster setup".
To accept a default or omit a question, do not enter a value.

This system will send event messages and periodic reports to Fujitsu
Technical
Support. To disable this feature, enter
autosupport modify -support disable
within 24 hours.

Enabling AutoSupport can significantly speed problem determination and
resolution should a problem occur on your system.
For further information on AutoSupport, see:
http://support.netapp.com/autosupport/

Type yes to confirm and continue {yes}: yes

.
.
.
```

(2) Configure the node management interface by responding to the prompts.

The prompts are similar to the following:

```
Enter the node management interface port [e0M]:
Enter the node management interface IP address: 172.17.8.229
Enter the node management interface netmask: 255.255.254.0
Enter the node management interface default gateway: 172.17.8.1
A node management interface on port e0M with IP address 172.17.8.229 has
been created.
```

(3) Create the cluster by responding to the prompts.

The prompts are similar to the following:

```
Do you want to create a new cluster or join an existing cluster? {create,
join}:
create

Do you intend for this node to be used as a single node cluster? {yes,
no} [no]:
no

Existing cluster interface configuration found:

Port MTU IP Netmask
e0a 1500 169.254.18.124 255.255.0.0
e1a 1500 169.254.184.44 255.255.0.0

Do you want to use this configuration? {yes, no} [yes]: no

System Defaults:
Private cluster network ports [e0a,e1a].
Cluster port MTU values will be set to 9000.
Cluster interface IP addresses will be automatically generated.

Do you want to use these defaults? {yes, no} [yes]: no

Enter the cluster administrator's (username "admin") password:

Retype the password:

Step 1 of 5: Create a Cluster
You can type "back", "exit", or "help" at any question.

List the private cluster network ports [e0a,e1a]:
Enter the cluster ports' MTU size [9000]:
Enter the cluster network netmask [255.255.0.0]: 255.255.254.0
Enter the cluster interface IP address for port e0a: 172.17.10.228
Enter the cluster interface IP address for port e1a: 172.17.10.229
Enter the cluster name: cluster_A

Creating cluster cluster_A

Starting cluster support services ...

Cluster cluster_A has been created.
```

- (4) Add licenses, set up a Cluster Administration SVM, and enter DNS information by responding to the prompts.

The prompts are similar to the following:

```
Step 2 of 5: Add Feature License Keys
You can type "back", "exit", or "help" at any question.

Enter an additional license key []:

Step 3 of 5: Set Up a Vserver for Cluster Administration
You can type "back", "exit", or "help" at any question.

Enter the cluster management interface port [e3a]:
Enter the cluster management interface IP address: 172.17.12.153
Enter the cluster management interface netmask: 255.255.252.0
Enter the cluster management interface default gateway: 172.17.12.1

A cluster management interface on port e3a with IP address 172.17.12.153
has been created. You can use this address to connect to and manage the
cluster.

Enter the DNS domain names: lab.netapp.com
Enter the name server IP addresses: 172.19.2.30
DNS lookup for the admin Vserver will use the lab.netapp.com domain.

Step 4 of 5: Configure Storage Failover (SFO)
You can type "back", "exit", or "help" at any question.

SFO will be enabled when the partner joins the cluster.

Step 5 of 5: Set Up the Node
You can type "back", "exit", or "help" at any question.

Where is the controller located []: svl
```

(5) Enable storage failover and set up the node by responding to the prompts.

The prompts are similar to the following:

```
Step 4 of 5: Configure Storage Failover (SFO)
You can type "back", "exit", or "help" at any question.

SFO will be enabled when the partner joins the cluster.

Step 5 of 5: Set Up the Node
You can type "back", "exit", or "help" at any question.

Where is the controller located []: site_A
```

(6) Complete the configuration of the node, but do not create data aggregates.

You can use ONTAP System Manager, pointing your web browser to the cluster management IP address (<https://172.17.12.153>).

Cluster management using ONTAP System Manager

3 Boot the next controller and join it to the cluster, following the prompts.

4 Confirm that nodes are configured in high-availability mode: `storage failover show -fields mode`

If not, you must configure HA mode on each node, and then reboot the nodes: `storage failover modify -mode ha -node localhost`

This command configures high-availability mode but does not enable storage failover. Storage failover is automatically enabled when you configure the MetroCluster configuration later in the process.

5 Confirm that you have four ports configured as cluster interconnects: `network port show`

The MetroCluster IP interfaces are not configured at this time and do not appear in the command output.

The following example shows two cluster ports on node_A_1:

```
cluster_A:*> network port show -role cluster

Node: node_A_1

                                                    Ignore
                                                    Speed(Mbps) Health Health
Port      IPspace      Broadcast Domain Link MTU  Admin/Oper  Status  Status
-----
e4a      Cluster      Cluster      up   9000  auto/40000  healthy  false
e4e      Cluster      Cluster      up   9000  auto/40000  healthy  false

Node: node_A_2

                                                    Ignore
                                                    Speed(Mbps) Health Health
Port      IPspace      Broadcast Domain Link MTU  Admin/Oper  Status  Status
-----
e4a      Cluster      Cluster      up   9000  auto/40000  healthy  false
e4e      Cluster      Cluster      up   9000  auto/40000  healthy  false

4 entries were displayed.
```

6 Repeat these steps on the partner cluster.



What to do next

Return to the ONTAP command-line interface and complete the MetroCluster configuration by performing the tasks that follow.

Configuring the clusters into a MetroCluster configuration

You must peer the clusters, mirror the root aggregates, create a mirrored data aggregate, and then issue the command to implement the MetroCluster operations.

Verifying drive assignment of pool 0 drives

You must verify that the remote drives are visible to the nodes and have been assigned correctly.

About this task

Automatic assignment depends on the storage system platform model and drive shelf arrangement.

["Considerations for automatic drive assignment and ADP systems in ONTAP 9.7 and later" \(page 10\)](#)

Procedure ▶▶▶

- 1 Verify that pool 0 drives are assigned automatically: `disk show`

4. Configuring the MetroCluster software in ONTAP Configuring the clusters into a MetroCluster configuration

One quarter (8 drives) were automatically assigned to node_A_1 and one quarter were automatically assigned to node_A_2. The remaining drives will be remote (pool 1) drives for node_B_1 and node_B_2.

```
cluster_A:.*> disk show
```

Disk	Usable Size	Disk Shelf	Disk Bay	Container Type	Container Type	Container Name	Container Owner
node_A_1:0n.12	1.75TB	0	12	SSD-NVM shared	shared	aggr0	node_A_1
node_A_1:0n.13	1.75TB	0	13	SSD-NVM shared	shared	aggr0	node_A_1
node_A_1:0n.14	1.75TB	0	14	SSD-NVM shared	shared	aggr0	node_A_1
node_A_1:0n.15	1.75TB	0	15	SSD-NVM shared	shared	aggr0	node_A_1
node_A_1:0n.16	1.75TB	0	16	SSD-NVM shared	shared	aggr0	node_A_1
node_A_1:0n.17	1.75TB	0	17	SSD-NVM shared	shared	aggr0	node_A_1
node_A_1:0n.18	1.75TB	0	18	SSD-NVM shared	shared	aggr0	node_A_1
node_A_1:0n.19	1.75TB	0	19	SSD-NVM shared	shared	-	node_A_1
node_A_2:0n.0	1.75TB	0	0	SSD-NVM shared	shared	aggr0_node_A_2_0	node_A_2
node_A_2:0n.1	1.75TB	0	1	SSD-NVM shared	shared	aggr0_node_A_2_0	node_A_2
node_A_2:0n.2	1.75TB	0	2	SSD-NVM shared	shared	aggr0_node_A_2_0	node_A_2
node_A_2:0n.3	1.75TB	0	3	SSD-NVM shared	shared	aggr0_node_A_2_0	node_A_2
node_A_2:0n.4	1.75TB	0	4	SSD-NVM shared	shared	aggr0_node_A_2_0	node_A_2
node_A_2:0n.5	1.75TB	0	5	SSD-NVM shared	shared	aggr0_node_A_2_0	node_A_2
node_A_2:0n.6	1.75TB	0	6	SSD-NVM shared	shared	aggr0_node_A_2_0	node_A_2
node_A_2:0n.7	1.75TB	0	7	SSD-NVM shared	shared	-	node_A_2
node_A_2:0n.24	-	0	24	SSD-NVM unassigned	unassigned	-	-
node_A_2:0n.25	-	0	25	SSD-NVM unassigned	unassigned	-	-
node_A_2:0n.26	-	0	26	SSD-NVM unassigned	unassigned	-	-
node_A_2:0n.27	-	0	27	SSD-NVM unassigned	unassigned	-	-
node_A_2:0n.28	-	0	28	SSD-NVM unassigned	unassigned	-	-
node_A_2:0n.29	-	0	29	SSD-NVM unassigned	unassigned	-	-
node_A_2:0n.30	-	0	30	SSD-NVM unassigned	unassigned	-	-
node_A_2:0n.31	-	0	31	SSD-NVM unassigned	unassigned	-	-
node_A_2:0n.36	-	0	36	SSD-NVM unassigned	unassigned	-	-
node_A_2:0n.37	-	0	37	SSD-NVM unassigned	unassigned	-	-
node_A_2:0n.38	-	0	38	SSD-NVM unassigned	unassigned	-	-
node_A_2:0n.39	-	0	39	SSD-NVM unassigned	unassigned	-	-
node_A_2:0n.40	-	0	40	SSD-NVM unassigned	unassigned	-	-
node_A_2:0n.41	-	0	41	SSD-NVM unassigned	unassigned	-	-
node_A_2:0n.42	-	0	42	SSD-NVM unassigned	unassigned	-	-
node_A_2:0n.43	-	0	43	SSD-NVM unassigned	unassigned	-	-

32 entries were displayed.

4. Configuring the MetroCluster software in ONTAP Configuring the clusters into a MetroCluster configuration

The following example shows the cluster_B output:

```
cluster_B::> disk show
```

Disk	Usable Size	Disk Shelf	Disk Bay	Disk Type	Container Type	Container Name	Container Owner
Info: This cluster has partitioned disks. To get a complete list of spare disk capacity use "storage aggregate show-spare-disks".							
node_B_1:0n.12	1.75TB	0	12	SSD-NVM	shared	aggr0	node_B_1
node_B_1:0n.13	1.75TB	0	13	SSD-NVM	shared	aggr0	node_B_1
node_B_1:0n.14	1.75TB	0	14	SSD-NVM	shared	aggr0	node_B_1
node_B_1:0n.15	1.75TB	0	15	SSD-NVM	shared	aggr0	node_B_1
node_B_1:0n.16	1.75TB	0	16	SSD-NVM	shared	aggr0	node_B_1
node_B_1:0n.17	1.75TB	0	17	SSD-NVM	shared	aggr0	node_B_1
node_B_1:0n.18	1.75TB	0	18	SSD-NVM	shared	aggr0	node_B_1
node_B_1:0n.19	1.75TB	0	19	SSD-NVM	shared	-	node_B_1
node_B_2:0n.0	1.75TB	0	0	SSD-NVM	shared	aggr0_node_B_1_0	node_B_2
node_B_2:0n.1	1.75TB	0	1	SSD-NVM	shared	aggr0_node_B_1_0	node_B_2
node_B_2:0n.2	1.75TB	0	2	SSD-NVM	shared	aggr0_node_B_1_0	node_B_2
node_B_2:0n.3	1.75TB	0	3	SSD-NVM	shared	aggr0_node_B_1_0	node_B_2
node_B_2:0n.4	1.75TB	0	4	SSD-NVM	shared	aggr0_node_B_1_0	node_B_2
node_B_2:0n.5	1.75TB	0	5	SSD-NVM	shared	aggr0_node_B_1_0	node_B_2
node_B_2:0n.6	1.75TB	0	6	SSD-NVM	shared	aggr0_node_B_1_0	node_B_2
node_B_2:0n.7	1.75TB	0	7	SSD-NVM	shared	-	node_B_2
node_B_2:0n.24	-	0	24	SSD-NVM	unassigned	-	-
node_B_2:0n.25	-	0	25	SSD-NVM	unassigned	-	-
node_B_2:0n.26	-	0	26	SSD-NVM	unassigned	-	-
node_B_2:0n.27	-	0	27	SSD-NVM	unassigned	-	-
node_B_2:0n.28	-	0	28	SSD-NVM	unassigned	-	-
node_B_2:0n.29	-	0	29	SSD-NVM	unassigned	-	-
node_B_2:0n.30	-	0	30	SSD-NVM	unassigned	-	-
node_B_2:0n.31	-	0	31	SSD-NVM	unassigned	-	-
node_B_2:0n.36	-	0	36	SSD-NVM	unassigned	-	-
node_B_2:0n.37	-	0	37	SSD-NVM	unassigned	-	-
node_B_2:0n.38	-	0	38	SSD-NVM	unassigned	-	-
node_B_2:0n.39	-	0	39	SSD-NVM	unassigned	-	-
node_B_2:0n.40	-	0	40	SSD-NVM	unassigned	-	-
node_B_2:0n.41	-	0	41	SSD-NVM	unassigned	-	-
node_B_2:0n.42	-	0	42	SSD-NVM	unassigned	-	-
node_B_2:0n.43	-	0	43	SSD-NVM	unassigned	-	-

```
32 entries were displayed.  
  
cluster_B::>
```

Peering the clusters

The clusters in the MetroCluster configuration must be in a peer relationship so that they can communicate with each other and perform the data mirroring essential to MetroCluster disaster recovery.

Related information

[Cluster and SVM Peering Power Guide](#)

4. Configuring the MetroCluster software in ONTAP Configuring the clusters into a MetroCluster configuration

[Considerations when using dedicated ports \(page 15\)](#)

[Considerations when sharing data ports \(page 16\)](#)

Configuring intercluster LIFs

You must create intercluster LIFs on ports used for communication between the MetroCluster partner clusters. You can use dedicated ports or ports that also have data traffic.

Configuring intercluster LIFs on dedicated ports

You can configure intercluster LIFs on dedicated ports. Doing so typically increases the available bandwidth for replication traffic.

Procedure ▶▶▶

- 1 List the ports in the cluster:**`network port show`
For complete command syntax, see the man page.
The following example shows the network ports in `cluster01`:

```
cluster01::> network port show
```

Node	Port	IPspace	Broadcast Domain	Link	MTU	Speed (Mbps) Admin/Oper

cluster01-01						
	e0a	Cluster	Cluster	up	1500	auto/1000
	e0b	Cluster	Cluster	up	1500	auto/1000
	e0c	Default	Default	up	1500	auto/1000
	e0d	Default	Default	up	1500	auto/1000
	e0e	Default	Default	up	1500	auto/1000
	e0f	Default	Default	up	1500	auto/1000
cluster01-02						
	e0a	Cluster	Cluster	up	1500	auto/1000
	e0b	Cluster	Cluster	up	1500	auto/1000
	e0c	Default	Default	up	1500	auto/1000
	e0d	Default	Default	up	1500	auto/1000
	e0e	Default	Default	up	1500	auto/1000
	e0f	Default	Default	up	1500	auto/1000

- 2 Determine which ports are available to dedicate to intercluster communication:**`network interface show -fields home-port,curr-port`
For complete command syntax, see the man page.

4. Configuring the MetroCluster software in ONTAP
Configuring the clusters into a MetroCluster configuration

The following example shows that ports `e0e` and `e0f` have not been assigned LIFs:

```
cluster01::> network interface show -fields home-port,curr-port
vserver lif                home-port curr-port
-----
Cluster cluster01-01_clus1 e0a       e0a
Cluster cluster01-01_clus2 e0b       e0b
Cluster cluster01-02_clus1 e0a       e0a
Cluster cluster01-02_clus2 e0b       e0b
cluster01
  cluster_mgmt             e0c       e0c
cluster01
  cluster01-01_mgmt1      e0c       e0c
cluster01
  cluster01-02_mgmt1      e0c       e0c
```

- 3 Create a failover group for the dedicated ports:** `network interface failover-groups create -vserver system_SVM -failover-group failover_group -targets physical_or_logical_ports`
The following example assigns ports `e0e` and `e0f` to the failover group `intercluster01` on the system SVM `cluster01`:

```
cluster01::> network interface failover-groups create -vserver cluster01 -failover-group
intercluster01 -targets
cluster01-01:e0e,cluster01-01:e0f,cluster01-02:e0e,cluster01-02:e0f
```

- 4 Verify that the failover group was created:** `network interface failover-groups show`
For complete command syntax, see the man page.

```
cluster01::> network interface failover-groups show
Vserver      Group      Failover
              Targets
-----
Cluster
  Cluster
              cluster01-01:e0a, cluster01-01:e0b,
              cluster01-02:e0a, cluster01-02:e0b
cluster01
  Default
              cluster01-01:e0c, cluster01-01:e0d,
              cluster01-02:e0c, cluster01-02:e0d,
              cluster01-01:e0e, cluster01-01:e0f
              cluster01-02:e0e, cluster01-02:e0f
  intercluster01
              cluster01-01:e0e, cluster01-01:e0f
              cluster01-02:e0e, cluster01-02:e0f
```

- 5 Create intercluster LIFs on the system SVM and assign them to the failover group.**

Option	Description
In ONTAP 9.7 and later:	<code>network interface create -vserver <i>system_SVM</i> -lif <i>LIF_name</i> -service-policy default-intercluster -home-node <i>node</i> -home-port <i>port</i> -address <i>port_IP</i> -netmask <i>netmask</i> -failover-group <i>failover_group</i></code>

For complete command syntax, see the man page.

4. Configuring the MetroCluster software in ONTAP
 Configuring the clusters into a MetroCluster configuration

The following example creates intercluster LIFs `cluster01_icl01` and `cluster01_icl02` in the failover group `intercluster01`:

```
cluster01::> network interface create -vserver cluster01 -lif cluster01_icl01 -service-policy default-intercluster -home-node cluster01-01 -home-port e0e -address 192.168.1.201 -netmask 255.255.255.0 -failover-group intercluster01

cluster01::> network interface create -vserver cluster01 -lif cluster01_icl02 -service-policy default-intercluster -home-node cluster01-02 -home-port e0e -address 192.168.1.202 -netmask 255.255.255.0 -failover-group intercluster01
```

6 Verify that the intercluster LIFs were created:

Option	Description
In ONTAP 9.7 and later:	<code>network interface show -service-policy default-intercluster</code>

For complete command syntax, see the man page.

```
cluster01::> network interface show -service-policy default-intercluster

      Logical      Status      Network      Current      Current Is
Vserver  Interface  Admin/Oper  Address/Mask  Node         Port      Home
-----
cluster01
      cluster01_icl01
                up/up      192.168.1.201/24  cluster01-01  e0e      true
      cluster01_icl02
                up/up      192.168.1.202/24  cluster01-02  e0f      true
```

7 Verify that the intercluster LIFs are redundant:

Option	Description
In ONTAP 9.7 and later:	<code>network interface show -service-policy default-intercluster -failover</code>

For complete command syntax, see the man page.

The following example shows that the intercluster LIFs `cluster01_icl01` and `cluster01_icl02` on the SVM `e0e` port will fail over to the `e0f` port.

```
cluster01::> network interface show -service-policy default-intercluster -failover

      Logical      Home      Failover      Failover
Vserver  Interface  Node:Port  Policy      Group
-----
cluster01
      cluster01_icl01 cluster01-01:e0e  local-only  intercluster01
                Failover Targets: cluster01-01:e0e,
                cluster01-01:e0f
      cluster01_icl02 cluster01-02:e0e  local-only  intercluster01
                Failover Targets: cluster01-02:e0e,
                cluster01-02:e0f
```

Related information

[Considerations when using dedicated ports \(page 15\)](#)

Configuring intercluster LIFs on shared data ports

You can configure intercluster LIFs on ports shared with the data network. Doing so reduces the number of ports you need for intercluster networking.

Procedure ▶▶▶

- 1 List the ports in the cluster:** `network port show`
 For complete command syntax, see the man page.
 The following example shows the network ports in `cluster01`:

```
cluster01::> network port show
```

Node	Port	IPspace	Broadcast Domain	Link	MTU	Speed (Mbps)	Admin/Oper

cluster01-01							
	e0a	Cluster	Cluster	up	1500	auto/1000	
	e0b	Cluster	Cluster	up	1500	auto/1000	
	e0c	Default	Default	up	1500	auto/1000	
	e0d	Default	Default	up	1500	auto/1000	
cluster01-02							
	e0a	Cluster	Cluster	up	1500	auto/1000	
	e0b	Cluster	Cluster	up	1500	auto/1000	
	e0c	Default	Default	up	1500	auto/1000	
	e0d	Default	Default	up	1500	auto/1000	

- 2 Create intercluster LIFs on the system SVM:**

Option	Description
In ONTAP 9.7 and later:	<code>network interface create -vserver system_SVM -lif LIF_name -service-policy default-intercluster -home-node node -home-port port -address port_IP -netmask netmask</code>

For complete command syntax, see the man page.

The following example creates intercluster LIFs `cluster01_icl01` and `cluster01_icl02`:

```
cluster01::> network interface create -vserver cluster01 -lif cluster01_icl01 -service-policy default-intercluster -home-node cluster01-01 -home-port e0c -address 192.168.1.201 -netmask 255.255.255.0
```

```
cluster01::> network interface create -vserver cluster01 -lif cluster01_icl02 -service-policy default-intercluster -home-node cluster01-02 -home-port e0c -address 192.168.1.202 -netmask 255.255.255.0
```

- 3 Verify that the intercluster LIFs were created:**

Option	Description
In ONTAP 9.7 and later:	<code>network interface show -service-policy default-intercluster</code>

4. Configuring the MetroCluster software in ONTAP Configuring the clusters into a MetroCluster configuration

For complete command syntax, see the man page.

```
cluster01::> network interface show -service-policy default-intercluster
      Logical      Status      Network      Current      Current Is
Vserver  Interface  Admin/Oper  Address/Mask  Node        Port      Home
-----
cluster01
      cluster01_icl01
              up/up      192.168.1.201/24  cluster01-01  e0c      true
      cluster01_icl02
              up/up      192.168.1.202/24  cluster01-02  e0c      true
```

4 Verify that the intercluster LIFs are redundant:

Option	Description
In ONTAP 9.7 and later:	<code>network interface show -service-policy default-intercluster -failover</code>

For complete command syntax, see the man page.

The following example shows that the intercluster LIFs `cluster01_icl01` and `cluster01_icl02` on the `e0c` port will fail over to the `e0d` port.

```
cluster01::> network interface show -service-policy default-intercluster -failover
      Logical      Home      Failover      Failover
Vserver  Interface  Node:Port  Policy        Group
-----
cluster01
      cluster01_icl01 cluster01-01:e0c  local-only  192.168.1.201/24
              Failover Targets: cluster01-01:e0c,
              cluster01-01:e0d
      cluster01_icl02 cluster01-02:e0c  local-only  192.168.1.201/24
              Failover Targets: cluster01-02:e0c,
              cluster01-02:e0d
```

Related information

[Considerations when sharing data ports \(page 16\)](#)

Creating a cluster peer relationship

You can use the `cluster peer create` command to create a peer relationship between a local and remote cluster. After the peer relationship has been created, you can run `cluster peer create` on the remote cluster to authenticate it to the local cluster.

Before you begin

- You must have created intercluster LIFs on every node in the clusters that are being peered.

Procedure ▶▶▶

- 1 **On the destination cluster, create a peer relationship with the source cluster:** `cluster peer create -generate-passphrase -offer-expiration MM/DD/YYYY HH:MM:SS | 1...7days | 1...168hours -peer-addr peer_LIF_IPs -ip-space ip-space`
If you specify both `-generate-passphrase` and `-peer-addr`s, only the cluster whose intercluster LIFs are specified in `-peer-addr`s can use the generated password.

4. Configuring the MetroCluster software in ONTAP Configuring the clusters into a MetroCluster configuration

You can ignore the `-ip-space` option if you are not using a custom IPspace. For complete command syntax, see the man page.

The following example creates a cluster peer relationship on an unspecified remote cluster:

```
cluster02:> cluster peer create -generate-passphrase -offer-expiration 2days

          Passphrase: UCa+6lRVICXeL/gq1WrK7ShR
          Expiration Time: 6/7/2017 08:16:10 EST
Initial Allowed Vserver Peers: -
          Intercluster LIF IP: 192.140.112.101
          Peer Cluster Name: Clus_7ShR (temporary generated)

Warning: make a note of the passphrase - it cannot be displayed again.
```

- 2** On source cluster, authenticate the source cluster to the destination cluster: `cluster peer create -peer-addr peer_LIF_IPs -ip-space ip-space`

For complete command syntax, see the man page.

The following example authenticates the local cluster to the remote cluster at intercluster LIF IP addresses 192.140.112.101 and 192.140.112.102:

```
cluster01:> cluster peer create -peer-addr 192.140.112.101,192.140.112.102

Notice: Use a generated passphrase or choose a passphrase of 8 or more characters.
          To ensure the authenticity of the peering relationship, use a phrase or sequence of characters that would
          be hard to guess.

Enter the passphrase:
Confirm the passphrase:

Clusters cluster02 and cluster01 are peered.
```

Enter the passphrase for the peer relationship when prompted.

- 3** Verify that the cluster peer relationship was created: `cluster peer show -instance`

```
cluster01:> cluster peer show -instance

          Peer Cluster Name: cluster02
          Remote Intercluster Addresses: 192.140.112.101, 192.140.112.102
Availability of the Remote Cluster: Available
          Remote Cluster Name: cluster2
          Active IP Addresses: 192.140.112.101, 192.140.112.102
          Cluster Serial Number: 1-80-123456
          Address Family of Relationship: ipv4
Authentication Status Administrative: no-authentication
Authentication Status Operational: absent
          Last Update Time: 02/05 21:05:41
          IPspace for the Relationship: Default
```

4 Check the connectivity and status of the nodes in the peer relationship: `cluster peer health show`

```
cluster01::> cluster peer health show
Node      cluster-Name      Node-Name
      Ping-Status      RDB-Health Cluster-Health Avail...
-----
cluster01-01
      cluster02      cluster02-01
      Data: interface_reachable
      ICMP: interface_reachable true      true      true
      cluster02-02
      Data: interface_reachable
      ICMP: interface_reachable true      true      true
cluster01-02
      cluster02      cluster02-01
      Data: interface_reachable
      ICMP: interface_reachable true      true      true
      cluster02-02
      Data: interface_reachable
      ICMP: interface_reachable true      true      true
```



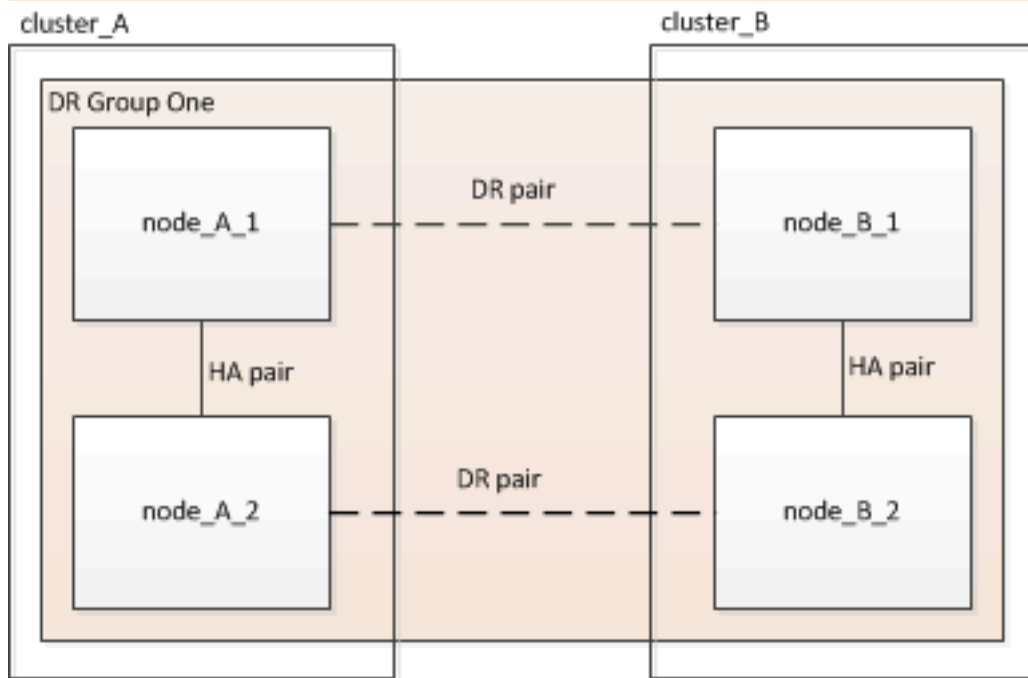
Creating the DR group

You must create the disaster recovery (DR) group relationships between the clusters.

About this task

You perform this procedure on one of the clusters in the MetroCluster configuration to create the DR relationships between the nodes in both clusters.

Caution
 The DR relationships cannot be changed after the DR groups are created.



Procedure ▶▶▶

- 1 **Verify that the nodes are ready for creation of the DR group by entering the following command on each:** `metrocluster configuration-settings show-status`

The command output should show that the nodes are ready:

```
cluster_A::> metrocluster configuration-settings show-status
Cluster                Node                Configuration Settings Status
-----
cluster_A              node_A_1           ready for DR group create
                      node_A_2           ready for DR group create
2 entries were displayed.
```

```
cluster_B::> metrocluster configuration-settings show-status
Cluster                Node                Configuration Settings Status
-----
cluster_B              node_B_1           ready for DR group create
                      node_B_2           ready for DR group create
2 entries were displayed.
```

- 2 **Create the DR group:** `metrocluster configuration-settings dr-group create -partner-cluster partner-cluster-name -local-node local-node-name -remote-node remote-node-name`

This command is issued only once. It does not need to be repeated on the partner cluster. In the command, you specify the name of the remote cluster and the name of one local node and one node on the partner cluster.

The two nodes you specify are configured as DR partners and the other two nodes (which are not specified in the command) are configured as the second DR pair in the DR group. These relationships cannot be changed after you enter this command.

The following command creates these DR pairs:

- node_A_1 and node_B_1
- node_A_2 and node_B_2

```
Cluster_A::> metrocluster configuration-settings dr-group create -partner-
cluster cluster_B -local-node node_A_1 -remote-node node_B_1
[Job 27] Job succeeded: DR Group Create is successful.
```



Configuring and connecting the MetroCluster IP interfaces

You must configure the MetroCluster IP (MCCIP) interfaces that are used for replication of each node's storage and nonvolatile cache. You then establish the connections using the MCCIP interfaces. This creates iSCSI connections for storage replication.

About this task

You must create two interfaces for each node. The interfaces must be associated with the VLANs defined in the MetroCluster RCF file.

You must create all MetroCluster IP interface 'A' ports in the same VLAN and all MetroCluster IP interface 'B' ports in the other VLAN.

4. Configuring the MetroCluster software in ONTAP
 Configuring the clusters into a MetroCluster configuration

Caution

You must choose the MetroCluster IP addresses carefully because you cannot change them after initial configuration.

["Considerations for MetroCluster IP configuration" \(page 9\)](#)

Caution

Starting with ONTAP 9.8, certain platforms use a VLAN for the MetroCluster IP interface. By default, each of the two ports uses a different VLAN: 10 and 20. You can also specify a different (non-default) VLAN higher than 100 (between 101–4095) using the `-vlan-id` parameter in the `metrocluster configuration-settings interface create` command.

The following platform models use VLANs.

ETERNUS AX platforms	ETERNUS HX platforms
<ul style="list-style-type: none"> • AX2100 • AX2200 • AX4100 	<ul style="list-style-type: none"> • HX2200 • FAS500f • HX6100

The following IP addresses and subnets are used in the examples:

Node	Interface	IP address	Subnet
node_A_1	MetroCluster IP interface 1	10.1.1.1	10.1.1/24
	MetroCluster IP interface 2	10.1.2.1	10.1.2/24
node_A_2	MetroCluster IP interface 1	10.1.1.2	10.1.1/24
	MetroCluster IP interface 2	10.1.2.2	10.1.2/24
node_B_1	MetroCluster IP interface 1	10.1.1.3	10.1.1/24
	MetroCluster IP interface 2	10.1.2.3	10.1.2/24
node_B_2	MetroCluster IP interface 1	10.1.1.4	10.1.1/24
	MetroCluster IP interface 2	10.1.2.4	10.1.2/24

The physical ports used by the MetroCluster IP interfaces depends on the platform model, as shown in the following table.

Platform model	MetroCluster IP port	VLAN ID	
AX3000 and HX6100	e1a	Not used	On these systems, these physical ports are also used as cluster interfaces.
	e1b		
AX2100 and HX2200	e0a	10	
	e0b	20	
AX2200 and FAS500f	e0c	10	
	e0d	20	

Verifying or manually performing pool 1 drives assignment

Depending on the storage configuration, you must either verify pool 1 drive assignment or manually assign drives to pool 1 for each node in the MetroCluster IP configuration. The procedure you use depends on the version of ONTAP you are using.

Verifying disk assignment for pool 1 disks

You must verify that the remote disks are visible to the nodes and have been assigned correctly.

Before you begin

You must wait at least ten minutes for disk auto-assignment to complete after the MetroCluster IP interfaces and connections were created with the `metrocluster configuration-settings connection connect` command.

About this task

Command output will show disk names in the form: `node-name:0m.i1.0L1`

["Considerations for automatic drive assignment and ADP systems in ONTAP 9.7 and later"](#) (page 10)

Procedure ▶▶▶

1 Verify pool 1 disks are auto-assigned: `disk show`

Drive autoassignment has assigned one quarter (8 drives) to `node_A_1` and one quarter to `node_A_2`. The remaining drives will be remote (pool1) disks for `node_B_1` and `node_B_2`.

```
cluster_B:> disk show -host-adapter 0m -owner node_B_2
```

Disk	Usable Size	Disk Shelf Bay	Type	Container Type	Container Name	Owner
node_B_2:0m.i0.2L4	894.0GB	0 29	SSD-NVM	shared	-	node_B_2
node_B_2:0m.i0.2L10	894.0GB	0 25	SSD-NVM	shared	-	node_B_2
node_B_2:0m.i0.3L3	894.0GB	0 28	SSD-NVM	shared	-	node_B_2
node_B_2:0m.i0.3L9	894.0GB	0 24	SSD-NVM	shared	-	node_B_2
node_B_2:0m.i0.3L11	894.0GB	0 26	SSD-NVM	shared	-	node_B_2
node_B_2:0m.i0.3L12	894.0GB	0 27	SSD-NVM	shared	-	node_B_2
node_B_2:0m.i0.3L15	894.0GB	0 30	SSD-NVM	shared	-	node_B_2
node_B_2:0m.i0.3L16	894.0GB	0 31	SSD-NVM	shared	-	node_B_2

8 entries were displayed.

```
cluster_B:> disk show -host-adapter 0m -owner node_B_1
```

Disk	Usable Size	Disk Shelf Bay	Type	Container Type	Container Name	Owner
node_B_1:0m.i2.3L19	1.75TB	0 42	SSD-NVM	shared	-	node_B_1
node_B_1:0m.i2.3L20	1.75TB	0 43	SSD-NVM	spare	Pool1	node_B_1
node_B_1:0m.i2.3L23	1.75TB	0 40	SSD-NVM	shared	-	node_B_1
node_B_1:0m.i2.3L24	1.75TB	0 41	SSD-NVM	spare	Pool1	node_B_1
node_B_1:0m.i2.3L29	1.75TB	0 36	SSD-NVM	shared	-	node_B_1
node_B_1:0m.i2.3L30	1.75TB	0 37	SSD-NVM	shared	-	node_B_1
node_B_1:0m.i2.3L31	1.75TB	0 38	SSD-NVM	shared	-	node_B_1
node_B_1:0m.i2.3L32	1.75TB	0 39	SSD-NVM	shared	-	node_B_1

8 entries were displayed.

4. Configuring the MetroCluster software in ONTAP Configuring the clusters into a MetroCluster configuration

```
cluster_B::> disk show
```

Disk	Usable Size	Disk Shelf Bay	Type	Container Type	Container Name	Container Owner
node_B_1:0m.i1.0L6	1.75TB	0 1	SSD-NVM shared	-	-	node_A_2
node_B_1:0m.i1.0L8	1.75TB	0 3	SSD-NVM shared	-	-	node_A_2
node_B_1:0m.i1.0L17	1.75TB	0 18	SSD-NVM shared	-	-	node_A_1
node_B_1:0m.i1.0L22	1.75TB	0 17	SSD-NVM shared	-	node_A_1	
node_B_1:0m.i1.0L25	1.75TB	0 12	SSD-NVM shared	-	node_A_1	
node_B_1:0m.i1.2L2	1.75TB	0 5	SSD-NVM shared	-	node_A_2	
node_B_1:0m.i1.2L7	1.75TB	0 2	SSD-NVM shared	-	node_A_2	
node_B_1:0m.i1.2L14	1.75TB	0 7	SSD-NVM shared	-	node_A_2	
node_B_1:0m.i1.2L21	1.75TB	0 16	SSD-NVM shared	-	node_A_1	
node_B_1:0m.i1.2L27	1.75TB	0 14	SSD-NVM shared	-	node_A_1	
node_B_1:0m.i1.2L28	1.75TB	0 15	SSD-NVM shared	-	node_A_1	
node_B_1:0m.i2.1L1	1.75TB	0 4	SSD-NVM shared	-	node_A_2	
node_B_1:0m.i2.1L5	1.75TB	0 0	SSD-NVM shared	-	node_A_2	
node_B_1:0m.i2.1L13	1.75TB	0 6	SSD-NVM shared	-	node_A_2	
node_B_1:0m.i2.1L18	1.75TB	0 19	SSD-NVM shared	-	node_A_1	
node_B_1:0m.i2.1L26	1.75TB	0 13	SSD-NVM shared	-	node_A_1	
node_B_1:0m.i2.3L19	1.75TB	0 42	SSD-NVM shared	-	node_B_1	
node_B_1:0m.i2.3L20	1.75TB	0 43	SSD-NVM shared	-	node_B_1	
node_B_1:0m.i2.3L23	1.75TB	0 40	SSD-NVM shared	-	node_B_1	
node_B_1:0m.i2.3L24	1.75TB	0 41	SSD-NVM shared	-	node_B_1	
node_B_1:0m.i2.3L29	1.75TB	0 36	SSD-NVM shared	-	node_B_1	
node_B_1:0m.i2.3L30	1.75TB	0 37	SSD-NVM shared	-	node_B_1	
node_B_1:0m.i2.3L31	1.75TB	0 38	SSD-NVM shared	-	node_B_1	
node_B_1:0m.i2.3L32	1.75TB	0 39	SSD-NVM shared	-	node_B_1	
node_B_1:0n.12	1.75TB	0 12	SSD-NVM shared	aggr0	node_B_1	
node_B_1:0n.13	1.75TB	0 13	SSD-NVM shared	aggr0	node_B_1	
node_B_1:0n.14	1.75TB	0 14	SSD-NVM shared	aggr0	node_B_1	
node_B_1:0n.15	1.75TB	0 15	SSD-NVM shared	aggr0	node_B_1	
node_B_1:0n.16	1.75TB	0 16	SSD-NVM shared	aggr0	node_B_1	
node_B_1:0n.17	1.75TB	0 17	SSD-NVM shared	aggr0	node_B_1	
node_B_1:0n.18	1.75TB	0 18	SSD-NVM shared	aggr0	node_B_1	
node_B_1:0n.19	1.75TB	0 19	SSD-NVM shared	-	node_B_1	
node_B_1:0n.24	894.0GB	0 24	SSD-NVM shared	-	node_A_2	
node_B_1:0n.25	894.0GB	0 25	SSD-NVM shared	-	node_A_2	
node_B_1:0n.26	894.0GB	0 26	SSD-NVM shared	-	node_A_2	
node_B_1:0n.27	894.0GB	0 27	SSD-NVM shared	-	node_A_2	
node_B_1:0n.28	894.0GB	0 28	SSD-NVM shared	-	node_A_2	
node_B_1:0n.29	894.0GB	0 29	SSD-NVM shared	-	node_A_2	
node_B_1:0n.30	894.0GB	0 30	SSD-NVM shared	-	node_A_2	
node_B_1:0n.31	894.0GB	0 31	SSD-NVM shared	-	node_A_2	
node_B_1:0n.36	1.75TB	0 36	SSD-NVM shared	-	node_A_1	
node_B_1:0n.37	1.75TB	0 37	SSD-NVM shared	-	node_A_1	
node_B_1:0n.38	1.75TB	0 38	SSD-NVM shared	-	node_A_1	
node_B_1:0n.39	1.75TB	0 39	SSD-NVM shared	-	node_A_1	
node_B_1:0n.40	1.75TB	0 40	SSD-NVM shared	-	node_A_1	
node_B_1:0n.41	1.75TB	0 41	SSD-NVM shared	-	node_A_1	
node_B_1:0n.42	1.75TB	0 42	SSD-NVM shared	-	node_A_1	
node_B_1:0n.43	1.75TB	0 43	SSD-NVM shared	-	node_A_1	

4. Configuring the MetroCluster software in ONTAP Configuring the clusters into a MetroCluster configuration

```
node_B_2:0m.i0.2L4 894.0GB 0 29 SSD-NVM shared - node_B_2
node_B_2:0m.i0.2L10 894.0GB 0 25 SSD-NVM shared - node_B_2
node_B_2:0m.i0.3L3 894.0GB 0 28 SSD-NVM shared - node_B_2
node_B_2:0m.i0.3L9 894.0GB 0 24 SSD-NVM shared - node_B_2
node_B_2:0m.i0.3L11 894.0GB 0 26 SSD-NVM shared - node_B_2
node_B_2:0m.i0.3L12 894.0GB 0 27 SSD-NVM shared - node_B_2
node_B_2:0m.i0.3L15 894.0GB 0 30 SSD-NVM shared - node_B_2
node_B_2:0m.i0.3L16 894.0GB 0 31 SSD-NVM shared - node_B_2
node_B_2:0n.0 1.75TB 0 0 SSD-NVM shared aggr0_rhal2_b1_cm_02_0 node_B_2
node_B_2:0n.1 1.75TB 0 1 SSD-NVM shared aggr0_rhal2_b1_cm_02_0 node_B_2
node_B_2:0n.2 1.75TB 0 2 SSD-NVM shared aggr0_rhal2_b1_cm_02_0 node_B_2
node_B_2:0n.3 1.75TB 0 3 SSD-NVM shared aggr0_rhal2_b1_cm_02_0 node_B_2
node_B_2:0n.4 1.75TB 0 4 SSD-NVM shared aggr0_rhal2_b1_cm_02_0 node_B_2
node_B_2:0n.5 1.75TB 0 5 SSD-NVM shared aggr0_rhal2_b1_cm_02_0 node_B_2
node_B_2:0n.6 1.75TB 0 6 SSD-NVM shared aggr0_rhal2_b1_cm_02_0 node_B_2
node_B_2:0n.7 1.75TB 0 7 SSD-NVM shared - node_B_2
64 entries were displayed.

cluster_B:>

cluster_A:> disk show
Usable Disk Container Container
Disk Size Shelf Bay Type Type Name Owner
-----
node_A_1:0m.i1.0L2 1.75TB 0 5 SSD-NVM shared - node_B_2
node_A_1:0m.i1.0L8 1.75TB 0 3 SSD-NVM shared - node_B_2
node_A_1:0m.i1.0L18 1.75TB 0 19 SSD-NVM shared - node_B_1
node_A_1:0m.i1.0L25 1.75TB 0 12 SSD-NVM shared - node_B_1
node_A_1:0m.i1.0L27 1.75TB 0 14 SSD-NVM shared - node_B_1
node_A_1:0m.i1.2L1 1.75TB 0 4 SSD-NVM shared - node_B_2
node_A_1:0m.i1.2L6 1.75TB 0 1 SSD-NVM shared - node_B_2
node_A_1:0m.i1.2L7 1.75TB 0 2 SSD-NVM shared - node_B_2
node_A_1:0m.i1.2L14 1.75TB 0 7 SSD-NVM shared - node_B_2
node_A_1:0m.i1.2L17 1.75TB 0 18 SSD-NVM shared - node_B_1
node_A_1:0m.i1.2L22 1.75TB 0 17 SSD-NVM shared - node_B_1
node_A_1:0m.i2.1L5 1.75TB 0 0 SSD-NVM shared - node_B_2
node_A_1:0m.i2.1L13 1.75TB 0 6 SSD-NVM shared - node_B_2
node_A_1:0m.i2.1L21 1.75TB 0 16 SSD-NVM shared - node_B_1
node_A_1:0m.i2.1L26 1.75TB 0 13 SSD-NVM shared - node_B_1
node_A_1:0m.i2.1L28 1.75TB 0 15 SSD-NVM shared - node_B_1
node_A_1:0m.i2.3L19 1.75TB 0 42 SSD-NVM shared - node_A_1
node_A_1:0m.i2.3L20 1.75TB 0 43 SSD-NVM shared - node_A_1
node_A_1:0m.i2.3L23 1.75TB 0 40 SSD-NVM shared - node_A_1
node_A_1:0m.i2.3L24 1.75TB 0 41 SSD-NVM shared - node_A_1
node_A_1:0m.i2.3L29 1.75TB 0 36 SSD-NVM shared - node_A_1
node_A_1:0m.i2.3L30 1.75TB 0 37 SSD-NVM shared - node_A_1
node_A_1:0m.i2.3L31 1.75TB 0 38 SSD-NVM shared - node_A_1
node_A_1:0m.i2.3L32 1.75TB 0 39 SSD-NVM shared - node_A_1
```

4. Configuring the MetroCluster software in ONTAP

Configuring the clusters into a MetroCluster configuration

```
node_A_1:0n.12 1.75TB 0 12 SSD-NVM shared aggr0 node_A_1
node_A_1:0n.13 1.75TB 0 13 SSD-NVM shared aggr0 node_A_1
node_A_1:0n.14 1.75TB 0 14 SSD-NVM shared aggr0 node_A_1
node_A_1:0n.15 1.75TB 0 15 SSD-NVM shared aggr0 node_A_1
node_A_1:0n.16 1.75TB 0 16 SSD-NVM shared aggr0 node_A_1
node_A_1:0n.17 1.75TB 0 17 SSD-NVM shared aggr0 node_A_1
node_A_1:0n.18 1.75TB 0 18 SSD-NVM shared aggr0 node_A_1
node_A_1:0n.19 1.75TB 0 19 SSD-NVM shared - node_A_1
node_A_1:0n.24 894.0GB 0 24 SSD-NVM shared - node_B_2
node_A_1:0n.25 894.0GB 0 25 SSD-NVM shared - node_B_2
node_A_1:0n.26 894.0GB 0 26 SSD-NVM shared - node_B_2
node_A_1:0n.27 894.0GB 0 27 SSD-NVM shared - node_B_2
node_A_1:0n.28 894.0GB 0 28 SSD-NVM shared - node_B_2
node_A_1:0n.29 894.0GB 0 29 SSD-NVM shared - node_B_2
node_A_1:0n.30 894.0GB 0 30 SSD-NVM shared - node_B_2
node_A_1:0n.31 894.0GB 0 31 SSD-NVM shared - node_B_2
node_A_1:0n.36 1.75TB 0 36 SSD-NVM shared - node_B_1
node_A_1:0n.37 1.75TB 0 37 SSD-NVM shared - node_B_1
node_A_1:0n.38 1.75TB 0 38 SSD-NVM shared - node_B_1
node_A_1:0n.39 1.75TB 0 39 SSD-NVM shared - node_B_1
node_A_1:0n.40 1.75TB 0 40 SSD-NVM shared - node_B_1
node_A_1:0n.41 1.75TB 0 41 SSD-NVM shared - node_B_1
node_A_1:0n.42 1.75TB 0 42 SSD-NVM shared - node_B_1
node_A_1:0n.43 1.75TB 0 43 SSD-NVM shared - node_B_1
node_A_2:0m.i2.3L3 894.0GB 0 28 SSD-NVM shared - node_A_2
node_A_2:0m.i2.3L4 894.0GB 0 29 SSD-NVM shared - node_A_2
node_A_2:0m.i2.3L9 894.0GB 0 24 SSD-NVM shared - node_A_2
node_A_2:0m.i2.3L10 894.0GB 0 25 SSD-NVM shared - node_A_2
node_A_2:0m.i2.3L11 894.0GB 0 26 SSD-NVM shared - node_A_2
node_A_2:0m.i2.3L12 894.0GB 0 27 SSD-NVM shared - node_A_2
node_A_2:0m.i2.3L15 894.0GB 0 30 SSD-NVM shared - node_A_2
node_A_2:0m.i2.3L16 894.0GB 0 31 SSD-NVM shared - node_A_2
node_A_2:0n.0 1.75TB 0 0 SSD-NVM shared aggr0_node_A_2_0 node_A_2
node_A_2:0n.1 1.75TB 0 1 SSD-NVM shared aggr0_node_A_2_0 node_A_2
node_A_2:0n.2 1.75TB 0 2 SSD-NVM shared aggr0_node_A_2_0 node_A_2
node_A_2:0n.3 1.75TB 0 3 SSD-NVM shared aggr0_node_A_2_0 node_A_2
node_A_2:0n.4 1.75TB 0 4 SSD-NVM shared aggr0_node_A_2_0 node_A_2
node_A_2:0n.5 1.75TB 0 5 SSD-NVM shared aggr0_node_A_2_0 node_A_2
node_A_2:0n.6 1.75TB 0 6 SSD-NVM shared aggr0_node_A_2_0 node_A_2
node_A_2:0n.7 1.75TB 0 7 SSD-NVM shared - node_A_2
64 entries were displayed.

cluster_A::>
```

Manually assigning drives for pool 1 (ONTAP 9.7 or later)

If the system was not preconfigured at the factory and does not meet the requirements for automatic drive assignment, you must manually assign the remote pool 1 drives.

About this task

This procedure applies to configurations running ONTAP 9.7 or later.

Details for determining whether your system requires manual disk assignment are included in ["Considerations for automatic drive assignment and ADP systems in ONTAP 9.7 and later"](#) (page 10).

If the configuration has fewer than four external shelves per site, you must check that automatic drive assignment on all nodes is disabled and manually assign the drives.

When the configuration includes only two external shelves per site, pool 1 drives for each site should be shared from the same shelf as shown in the following examples:

- node_A_1 is assigned drives in bays 0-11 on site_B-shelf_2 (remote)
- node_A_2 is assigned drives in bays 12-23 on site_B-shelf_2 (remote)

Procedure ▶▶▶

1 From each node in the MetroCluster IP configuration, assign remote drives to pool 1.

- (1) Display the list of unassigned drives: `disk show -host-adapter 0m -container-type unassigned`

```
cluster_A::> disk show -host-adapter 0m -container-type unassigned
          Usable      Disk      Container      Container
Disk      Size Shelf Bay Type      Type        Name      Owner
-----
6.23.0          -    23   0 SSD      unassigned  -        -
6.23.1          -    23   1 SSD      unassigned  -        -
.
.
.
node_A_2:0m.i1.2L51  -    21  14 SSD      unassigned  -        -
node_A_2:0m.i1.2L64  -    21  10 SSD      unassigned  -        -
.
.
.
48 entries were displayed.

cluster_A::>
```

- (2) Assign ownership of remote drives (0m) to pool 1 of the first node (for example, node_A_1):
`disk assign -disk disk-id -pool 1 -owner owner-node-name`
disk-id must identify a drive on a remote shelf of *owner-node-name*.
- (3) Confirm that the drives were assigned to pool 1: `disk show -host-adapter 0m -container-type unassigned`

Caution

The iSCSI connection used to access the remote drives appears as device 0m.

4. Configuring the MetroCluster software in ONTAP Configuring the clusters into a MetroCluster configuration

The following output shows that the drives on shelf 23 were assigned because they no longer appear in the list of unassigned drives:

```
cluster_A::> disk show -host-adapter 0m -container-type unassigned
              Usable      Disk      Container  Container
Disk          Size Shelf Bay Type      Type      Name      Owner
-----
node_A_2:0m.i1.2L51      -    21  14 SSD      unassigned -        -
node_A_2:0m.i1.2L64      -    21  10 SSD      unassigned -        -
.
.
.
node_A_2:0m.i2.1L90      -    21  19 SSD      unassigned -        -
24 entries were displayed.

cluster_A::>
```

- (4) Repeat these steps to assign pool 1 drives to the second node on site A (for example, node_A_2).
- (5) Repeat these steps on site B.

Mirroring the root aggregates

You must mirror the root aggregates to provide data protection.

About this task

By default, the root aggregate is created as RAID-DP type aggregate. You can change the root aggregate from RAID-DP to RAID4 type aggregate. The following command modifies the root aggregate for RAID4 type aggregate:

```
storage aggregate modify -aggregate aggr_name -raidtype raid4
```

Caution

On non-ADP systems, the RAID type of the aggregate can be modified from the default RAID-DP to RAID4 before or after the aggregate is mirrored.

Procedure

- 1 **Mirror the root aggregate:** `storage aggregate mirror aggr_name`
The following command mirrors the root aggregate for controller_A_1:

```
controller_A_1::> storage aggregate mirror aggr0_controller_A_1
```

This mirrors the aggregate, so it consists of a local plex and a remote plex located at the remote MetroCluster site.

- 2 Repeat the previous step for each node in the MetroCluster configuration.

Related information

[Logical Storage Management Guide](#)

Creating a mirrored data aggregate on each node

You must create a mirrored data aggregate on each node in the DR group.

Before you begin

- You should know what drives will be used in the new aggregate.
- If you have multiple drive types in your system (heterogeneous storage), you should understand how you can ensure that the correct drive type is selected.

About this task

- Drives are owned by a specific node; when you create an aggregate, all drives in that aggregate must be owned by the same node, which becomes the home node for that aggregate.
In systems using ADP, aggregates are created using partitions in which each drive is partitioned in to P1, P2 and P3 partitions.
- Aggregate names should conform to the naming scheme you determined when you planned your MetroCluster configuration.
[Disks and Aggregates Power Guide](#)

Procedure ▶▶▶

1 Display a list of available spares: `storage disk show -spare -owner node_name`

2 Create the aggregate by using the `storage aggregate create -mirror true` command.

If you are logged in to the cluster on the cluster management interface, you can create an aggregate on any node in the cluster. To ensure that the aggregate is created on a specific node, use the `-node` parameter or specify drives that are owned by that node.

You can specify the following options:

- Aggregate's home node (that is, the node that owns the aggregate in normal operation)
- List of specific drives that are to be added to the aggregate
- Number of drives to include

Caution

In the minimum supported configuration, in which a limited number of drives are available, you must use the `force-small-aggregate` option to allow the creation of a three disk RAID-DP aggregate.

- Checksum style to use for the aggregate
- Type of drives to use
- Size of drives to use
- Drive speed to use
- RAID type for RAID groups on the aggregate
- Maximum number of drives that can be included in a RAID group
- Whether drives with different RPM are allowed

For more information about these options, see the `storage aggregate create` man page.

The following command creates a mirrored aggregate with 10 disks:

```
cluster_A::> storage aggregate create aggr1_node_A_1 -diskcount 10 -node
node_A_1 -mirror true
[Job 15] Job is queued: Create aggr1_node_A_1.
[Job 15] The job is starting.
[Job 15] Job succeeded: DONE
```

- 3** Verify the RAID group and drives of your new aggregate: `storage aggregate show-status - aggregate aggregate-name`



Implementing the MetroCluster configuration

You must run the `metrocluster configure` command to start data protection in a MetroCluster configuration.

Before you begin

- There should be at least two non-root mirrored data aggregates on each cluster. You can verify this with the `storage aggregate show` command.

Caution

If you want to use a single mirrored data aggregate, then see [step 1](#) for instructions.

- The ha-config state of the controllers and chassis must be `mccip`.

About this task

You issue the `metrocluster configure` command once, on any of the nodes, to enable the MetroCluster configuration. You do not need to issue the command on each of the sites or nodes, and it does not matter which node or site you choose to issue the command on.

The `metrocluster configure` command automatically pairs the two nodes with the lowest system IDs in each of the two clusters as disaster recovery (DR) partners. In a four-node MetroCluster configuration, there are two DR partner pairs. The second DR pair is created from the two nodes with higher system IDs.

Procedure ▶▶▶

- 1** Configure the MetroCluster in the following format:

If your MetroCluster configuration has...	Then do this...
Multiple data aggregates	From any node's prompt, configure MetroCluster: <code>metrocluster configure node-name</code>
A single mirrored data aggregate	<p>(1) From any node's prompt, change to the advanced privilege level: <code>set -privilege advanced</code> You need to respond with <code>y</code> when you are prompted to continue into advanced mode and you see the advanced mode prompt (<code>*></code>).</p> <p>(2) Configure the MetroCluster with the <code>-allow-with-one-aggregate true</code> parameter: <code>metrocluster configure -allow-with-one-aggregate true node-name</code></p> <p>(3) Return to the admin privilege level: <code>set -privilege admin</code></p>

4. Configuring the MetroCluster software in ONTAP
Configuring the clusters into a MetroCluster configuration

Caution

The best practice is to have multiple data aggregates. If the first DR group has only one aggregate and you want to add a DR group with one aggregate, you must move the metadata volume off the single data aggregate. For more information on this procedure, see [Moving a metadata volume in MetroCluster configurations](#).

The following command enables the MetroCluster configuration on all of the nodes in the DR group that contains controller_A_1:

```
cluster_A::*> metrocluster configure -node-name controller_A_1

[Job 121] Job succeeded: Configure is successful.
```

2 Verify the networking status on site A: `network port show`

The following example shows the network port usage on a four-node MetroCluster configuration:

```
cluster_A::> network port show
```

Node	Port	IPspace	Broadcast Domain	Link	MTU	Speed (Mbps) Admin/Oper
controller_A_1						
	e0a	Cluster	Cluster	up	9000	auto/1000
	e0b	Cluster	Cluster	up	9000	auto/1000
	e0c	Default	Default	up	1500	auto/1000
	e0d	Default	Default	up	1500	auto/1000
	e0e	Default	Default	up	1500	auto/1000
	e0f	Default	Default	up	1500	auto/1000
	e0g	Default	Default	up	1500	auto/1000
controller_A_2						
	e0a	Cluster	Cluster	up	9000	auto/1000
	e0b	Cluster	Cluster	up	9000	auto/1000
	e0c	Default	Default	up	1500	auto/1000
	e0d	Default	Default	up	1500	auto/1000
	e0e	Default	Default	up	1500	auto/1000
	e0f	Default	Default	up	1500	auto/1000
	e0g	Default	Default	up	1500	auto/1000

14 entries were displayed.

3 Verify the MetroCluster configuration from both sites in the MetroCluster configuration.

(1) Verify the configuration from site A: `metrocluster show`

```
cluster_A::> metrocluster show
```

Configuration: IP fabric

Cluster	Entry Name	State
Local: cluster_A	Configuration state	configured
	Mode	normal
Remote: cluster_B	Configuration state	configured
	Mode	normal

(2) Verify the configuration from site B: `metrocluster show`

```
cluster_B:> metrocluster show

Configuration: IP fabric

Cluster                Entry Name                State
-----
Local: cluster_B       Configuration state       configured
                        Mode                       normal
Remote: cluster_A      Configuration state       configured
                        Mode                       normal
```



Creating unmirrored data aggregates

You can optionally create unmirrored data aggregates for data that does not require the redundant mirroring provided by MetroCluster configurations.

Before you begin

- You should know what drives or array LUNs will be used in the new aggregate.
- If you have multiple drive types in your system (heterogeneous storage), you should understand how you can verify that the correct drive type is selected.

About this task



CAUTION

In MetroCluster IP configurations, remote unmirrored aggregates are not accessible after a switchover.

Caution

The unmirrored aggregates must be local to the node owning them.

- Drives and array LUNs are owned by a specific node; when you create an aggregate, all drives in that aggregate must be owned by the same node, which becomes the home node for that aggregate.
- Aggregate names should conform to the naming scheme you determined when you planned your MetroCluster configuration.
- The *Disks and Aggregates Power Guide* contains more information about mirroring aggregates.

Procedure ▶▶▶

- 1 **Enable unmirrored aggregate deployment:** `metrocluster modify -enable-unmirrored-aggr-deployment true`
- 2 **Verify that disk autoassignment is disabled:** `disk option show`
- 3 **Install and cable the disk shelves that will contain the unmirrored aggregates.**
You can use the procedures in the Installation and Setup documentaiton for your platform and disk shelves.
[Fujitsu Manual Site](#)
- 4 **Manually assign all disks on the new shelf to the appropriate node:** `disk assign -disk disk-id -owner owner-node-name`

5 Create the aggregate: `storage aggregate create`

If you are logged in to the cluster on the cluster management interface, you can create an aggregate on any node in the cluster. To verify that the aggregate is created on a specific node, you should use the `-node` parameter or specify drives that are owned by that node.

You must also ensure that you are only including drives on the unmirrored shelf to the aggregate.

You can specify the following options:

- Aggregate's home node (that is, the node that owns the aggregate in normal operation)
- List of specific drives or array LUNs that are to be added to the aggregate
- Number of drives to include
- Checksum style to use for the aggregate
- Type of drives to use
- Size of drives to use
- Drive speed to use
- RAID type for RAID groups on the aggregate
- Maximum number of drives or array LUNs that can be included in a RAID group
- Whether drives with different RPM are allowed

For more information about these options, see the `storage aggregate create` man page.

The following command creates a unmirrored aggregate with 10 disks:

```
controller_A_1::> storage aggregate create aggr1_controller_A_1 -diskcount 10 -
node controller_A_1
[Job 15] Job is queued: Create aggr1_controller_A_1.
[Job 15] The job is starting.
[Job 15] Job succeeded: DONE
```

6 Verify the RAID group and drives of your new aggregate: `storage aggregate show-status -aggregate aggregate-name`

7 Disable unmirrored aggregate deployment: `metrocluster modify -enable-unmirrored-aggr-deployment false`

8 Verify that disk autoassignment is enabled: `disk option show`



Related information

[Disks and Aggregates Power Guide](#)

Checking the MetroCluster configuration

You can check that the components and relationships in the MetroCluster configuration are working correctly. You should do a check after initial configuration and after making any changes to the MetroCluster configuration. You should also do a check before a negotiated (planned) switchover or a switchback operation.

About this task

If the `metrocluster check run` command is issued twice within a short time on either or both clusters, a conflict can occur and the command might not collect all data. Subsequent `metrocluster check show` commands do not show the expected output.

Procedure ▶▶▶

1 Check the configuration: `metrocluster check run`

The command runs as a background job and might not be completed immediately.

```
cluster_A::> metrocluster check run
The operation has been started and is running in the background. Wait for
it to complete and run "metrocluster check show" to view the results. To
check the status of the running metrocluster check operation, use the command,
"metrocluster operation history show -job-id 2245"
```

```
cluster_A::> metrocluster check show
Last Checked On: 9/13/2018 20:41:37
```

Component	Result
nodes	ok
lifs	ok
config-replication	ok
aggregates	ok
clusters	ok
connections	ok

6 entries were displayed.

2 Display more detailed results from the most recent `metrocluster check run` command:

```
metrocluster check aggregate showmetrocluster check cluster showmetrocluster
check config-replication showmetrocluster check lif showmetrocluster check node
show
```

The `metrocluster check show` commands show the results of the most recent `metrocluster check run` command. You should always run the `metrocluster check run` command prior to using the `metrocluster check show` commands so that the information displayed is current.

4. Configuring the MetroCluster software in ONTAP Configuring the clusters into a MetroCluster configuration

The following example shows the `metrocluster check aggregate show` command output for a healthy four-node MetroCluster configuration:

```
cluster_A::> metrocluster check aggregate show

Last Checked On: 8/5/2014 00:42:58

Node                Aggregate                Check                Result
-----            -
controller_A_1      controller_A_1_aggr0
                    mirroring-status        ok
                    disk-pool-allocation    ok
                    ownership-state         ok
                    controller_A_1_aggr1
                    mirroring-status        ok
                    disk-pool-allocation    ok
                    ownership-state         ok
                    controller_A_1_aggr2
                    mirroring-status        ok
                    disk-pool-allocation    ok
                    ownership-state         ok

controller_A_2      controller_A_2_aggr0
                    mirroring-status        ok
                    disk-pool-allocation    ok
                    ownership-state         ok
                    controller_A_2_aggr1
                    mirroring-status        ok
                    disk-pool-allocation    ok
                    ownership-state         ok
                    controller_A_2_aggr2
                    mirroring-status        ok
                    disk-pool-allocation    ok
                    ownership-state         ok

18 entries were displayed.
```

The following example shows the `metrocluster check cluster show` command output for a healthy four-node MetroCluster configuration. It indicates that the clusters are ready to perform a negotiated switchover if necessary.

```
Last Checked On: 9/13/2017 20:47:04

Cluster          Check          Result
-----
mccint-fas9000-0102
    negotiated-switchover-ready    not-applicable
    switchback-ready                not-applicable
    job-schedules                   ok
    licenses                         ok
    periodic-check-enabled          ok
mccint-fas9000-0304
    negotiated-switchover-ready    not-applicable
    switchback-ready                not-applicable
    job-schedules                   ok
    licenses                         ok
    periodic-check-enabled          ok
10 entries were displayed.
```

Related information

- [Disks and Aggregates Power Guide](#)
- [Network Management Guide](#)

Completing ONTAP configuration

After configuring, enabling, and checking the MetroCluster configuration, you can proceed to complete the cluster configuration by adding additional SVMs, network interfaces and other ONTAP functionality as needed.

Verifying switchover, healing, and switchback

You should verify the switchover, healing, and switchback operations of the MetroCluster configuration.

Procedure ▶▶▶

- 1 Use the procedures for negotiated switchover, healing, and switchback that are mentioned in the *MetroCluster Management and Disaster Recovery Guide*.
[MetroCluster® Management and Disaster Recovery Guide](#)

Configuring the MetroCluster Tiebreaker or ONTAP Mediator software

You can download and install on a third site either the MetroCluster Tiebreaker software, or, starting with ONTAP 9.7, the ONTAP Mediator.

Before you begin

You must have a Linux host available that has network connectivity to both clusters in the MetroCluster configuration. The specific requirements are in the MetroCluster Tiebreaker or ONTAP Mediator documentation. If you are connecting to an existing Tiebreaker or ONTAP Mediator instance, you need the username, password, and IP address of the Tiebreaker or Mediator service.

About this task

If you must install a new instance of the ONTAP Mediator, follow the directions to install and configure the software.

["5. Configuring the ONTAP Mediator service for unplanned automatic switchover" \(page 95\)](#)

If you must install a new instance of the Tiebreaker software, follow the directions to install and configure the software.

[Tiebreaker Software 1.21 Installation and Configuration Guide](#)

You cannot use both the MetroCluster Tiebreaker software and the ONTAP Mediator with the same MetroCluster configuration.

["Considerations for using ONTAP Mediator or MetroCluster Tiebreaker" \(page 8\)](#)

Procedure ▶▶▶

1 Configure the ONTAP Mediator service or the Tiebreaker software:

- If you are using an existing instance of the ONTAP Mediator, add the ONTAP Mediator service to ONTAP using the following command:
`metrocluster configuration-settings mediator add -mediator-address ip-address-of-mediator-host`
- If you are using the Tiebreaker software, refer to the Tiebreaker documentation.
[Tiebreaker Software 1.21 Installation and Configuration Guide](#)



Protecting configuration backup files

You can provide additional protection for the cluster configuration backup files by specifying a remote URL (either HTTP or FTP) where the configuration backup files will be uploaded in addition to the default locations in the local cluster.

Procedure ▶▶▶

1 Set the URL of the remote destination for the configuration backup files: `system configuration backup settings modify URL-of-destination`

The System Administration Guide contains additional information under the section *Managing configuration backups*.



Related information

[System Administration Reference](#)

5. Configuring the ONTAP Mediator service for unplanned automatic switchover

Starting with ONTAP 9.7, the ONTAP Mediator service can assist the MetroCluster IP configuration in performing an automatic unplanned switchover by providing a physically separate repository for status information.

Caution

- The ONTAP Mediator service and MetroCluster Tiebreaker software should not both be used with the same MetroCluster configuration.
- The ONTAP Mediator service can support up to five MetroCluster configurations simultaneously.

Installing and configuring the ONTAP Mediator service

The ONTAP Mediator must be installed on the Linux host and then configured in ONTAP from the MetroCluster nodes.

Network requirements for using Mediator in a MetroCluster configuration

To install the ONTAP Mediator service in a MetroCluster configuration, you must ensure that the configuration meets several network requirements.

- Round trip latency
Round trip latency must be no more than 25 ms.
- MTU
The MTU size must be at least 1400.
- Packet loss
Packet loss must be less than or equal to 0.01%.
- Bandwidth
The link between the Mediator service and the MetroCluster configuration must have at least 1 Gbps of bandwidth.

Guidelines for upgrading the ONTAP Mediator in a MetroCluster configuration

If you are upgrading the ONTAP Mediator you must meet the Linux version requirements and follow guidelines for the upgrade.

- The Mediator service can be upgraded from version 1.0 to 1.1.
- All Mediator versions are supported on MetroCluster IP configurations running ONTAP 9.7 or later.
- Upgrading the host operating system and Mediator together
The following table provides the upgrade guidelines if you are upgrading from RHEL/CentOS 7.6 to a later RHEL/CentOS release in addition upgrading the Mediator version.

5. Configuring the ONTAP Mediator service for unplanned automatic switchover
Installing and configuring the ONTAP Mediator service

Target Linux version	Target Mediator version	Upgrade notes
RHEL/CentOS 7.7	1.1	<ul style="list-style-type: none"> The upgrade must be performed in the following order: <ol style="list-style-type: none"> Upgrade the operating system from RHEL/CentOS version 7.6 to 7.7. <div data-bbox="863 427 1474 678" style="background-color: #fff9c4; padding: 10px; margin: 10px 0;"> <p>Caution</p> <p>The ONTAP Mediator and Mediator-assisted automatic unplanned switchover is not available during the operating system upgrade. The Mediator is offline until the Mediator version 1.0 to 1.1 upgrade process is complete..</p> </div> Reboot the host to apply the kernel module changes. Upgrade the Mediator from version 1.0 to 1.1. "Installing or upgrading the ONTAP Mediator service" (page 98) The <code>storage iscsi-initiator show</code> command will report that the connection to the Mediator service is down during the upgrade. The ONTAP operating system will generate <code>cf.mccip.med.auso.stDisabled</code> EMS events during the upgrade. The ONTAP operating system will generate a <code>cf.mccip.med.auso.stEnabled</code> EMS event when automatic unplanned switchover is re-enabled.
RHEL/CentOS 8.0 or 8.1	1.1	<p>There is no direct upgrade path. You must remove the 1.0 version and install the 1.1 version after the operating system upgrade:</p> <ol style="list-style-type: none"> Delete the Mediator service from the ONTAP configuration: <code>metrocluster configuration-settings mediator remove</code> Uninstall the 1.0 version of the Mediator service. "Uninstalling the ONTAP Mediator service" (page 115) Upgrade the Linux operating system to version 8.0 or 8.1. Install the 1.1 version of the Mediator service. "Installing or upgrading the ONTAP Mediator service" (page 98) Add the Mediator service to the ONTAP configuration: <code>metrocluster configuration-settings add -address mediator-1.1-ip-address</code>

● After the upgrade

After the Mediator and operating system upgrade is complete, you should issue the `storage iscsi-initiator show` command to confirm that the Mediator connections are up.

- Reverting from a Mediator 1.1 installation

A direct revert from Mediator version 1.1 to 1.0 is not supported. You must remove the 1.1 version and reinstall the 1.0 version.

- 1 Delete the Mediator service from the ONTAP configuration: `metrocluster configuration-settings mediator remove`
- 2 Uninstall the 1.1 version of the Mediator service.
["Uninstalling the ONTAP Mediator service " \(page 115\)](#)
- 3 Install the 1.0 version of the Mediator service.
["Installing or upgrading the ONTAP Mediator service" \(page 98\)](#)
- 4 Add the Mediator service to the ONTAP configuration: `metrocluster configuration-settings add -address mediator-1.0-ip-address`

- Recovering from Linux kernel upgrades

The ONTAP Mediator requires the SCST kernel module. If the Linux kernel is updated, this dependency may lead to a loss of service. It is highly recommended that you rebuild the SCST kernel module when any kernel package changes are made.

Caution

- Upgrading from ONTAP Mediator version 1.0 to 1.1 rebuilds the SCST module.
- Kernel module changes are applied after the Linux kernel is rebooted.

You can use either of the following procedures to recover from a kernel upgrade that has resulted in loss of service for the Mediator.

Procedure	Steps
Remove and reinstall the SCST kernel module	<p>You must have the SCST tar bundle used by your version of Mediator:</p> <ul style="list-style-type: none"> • ONTAP Mediator 1.0 requires <code>scst-3.3.0.tar.bz2</code> • ONTAP Mediator 1.1 requires <code>scst-3.4.0.tar.bz2</code> <p>1 Uninstall the SCST module:</p> <p>(1) Download and untar the SCST tar bundle required by your version of Mediator.</p> <p>(2) Run the following commands inside of the <code>scst</code> directory:</p> <pre>systemctl stop mediator-scst make scstadm_uninstall make iscsi_uninstall make usr_uninstall make scst_uninstall depmod</pre> <p>2 Reinstall the SCST module for your version of Mediator by issuing the following commands inside of the <code>scst</code> directory:</p> <pre>make scst_install make usr_install make iscsi_install make scstadm_install depmod patch /etc/init.d/scst < /opt/netapp/lib/ontap_mediator/systemd/scst.patch reboot</pre>
Remove and reinstall ONTAP Mediator	<p>1 Delete the Mediator service from the ONTAP configuration: <code>metrocluster configuration-settings mediator remove</code></p> <p>2 Uninstall the Mediator service. "Uninstalling the ONTAP Mediator service " (page 115)</p> <p>3 Reinstall the Mediator service. "Installing or upgrading the ONTAP Mediator service" (page 98)</p> <p>4 Add the Mediator service to the ONTAP configuration: <code>metrocluster configuration-settings add -address mediator-ip-address</code></p>

Caution

This requires a reconfiguration of the Mediator in ONTAP.

Installing or upgrading the ONTAP Mediator service

To install the ONTAP Mediator service, you must ensure all prerequisites are met, get the installation package and run the installer on the host.

Before you begin

5. Configuring the ONTAP Mediator service for unplanned automatic switchover
Installing and configuring the ONTAP Mediator service

Mediator version	Supported Linux versions
1.2	Red Hat Enterprise Linux or CentOS 7.6, 7.7, 8.0, 8.1

- 64-bit physical installation or virtual machine

- 8 GB RAM
- User: Root access

The best practices for installing Red Hat Enterprise Linux or CentOS and the associated repositories on your system are listed below. Systems installed or configured differently might require additional steps.

- You must install Red Hat Enterprise Linux or CentOS according to Red Hat best practices.
- While installing the ONTAP Mediator service on Red Hat Enterprise Linux or CentOS, the system must have access to the appropriate repository so that the installation program can access and install all the required software dependencies.
- For the `yum` installer to find dependent software in the Red Hat Enterprise Linux repositories, you must have registered the system during the Red Hat Enterprise Linux installation or afterwards by using a valid Red Hat subscription.

See the Red Hat documentation for information about the Red Hat Subscription Manager.

- The following ports must be unused and available for the Mediator:

- 3260
- 31784

- If using a third-party firewall:

- HTTPS access must be enabled.
- It must be configured to allow access on ports 31784 and 3260

When using the default Red Hat or CentOS firewall, the firewall is automatically configured during Mediator installation.

- If the Linux host is in a location without access to the internet, you can either install the packages manually or you must ensure that the required packages are available in a local repository.

You can use the following link for information about setting up a repository.

[Redhat Solution 3176811: Creating a Local Repository and Sharing With Disconnected/Offline/Air-gapped Systems](#)

If you are using Link Aggregation Control Protocol (LACP) in a Linux environment, you must correctly configure the kernel and make sure the `sysctl net.ipv4.conf.all.arp_ignore` is set to 2.

The following packages are required by the ONTAP Mediator service version 1.2:

5. Configuring the ONTAP Mediator service for unplanned automatic switchover
 Installing and configuring the ONTAP Mediator service

All RHEL/CentOS versions	Additional packages for RHEL/CentOS 7.x	Additional packages for RHEL/CentOS 8.x
<ul style="list-style-type: none"> • openssl • openssl-devel • kernel-devel • gcc • libselinux-utils • make • redhat-lsb-core • patch • bzip2 • python36 • python36-devel • perl-Data-Dumper • perl-ExtUtils-MakeMaker • python3-pip 	<ul style="list-style-type: none"> • policycoreutils-python • python36-pip 	<ul style="list-style-type: none"> • elfutils-libelf-devel • policycoreutils-python-utils

- If signature verification is configured, it must be disabled. This can be done in one of two ways:
 - If the UEFI SecureBoot mechanism is configured, disable it.
 - Disable the signature verification mechanism by updating and regenerating the `grub.cfg` file:
 - 1 Open the `/etc/default/grub` file.
 - 2 Add the string `module.sig_enforce=0` to the end of the `GRUB_CMDLINE_LINUX` statement.
 - 3 Regenerate the `grub.cfg` file to implement the change: `update-bootloader || update-grub || grub2-mkconfig -o /boot/grub2/grub.cfg`
 - 4 Reboot the host.

About this task

The Mediator installation package is a self-extracting compressed tar file that includes:

- An RPM file containing all dependencies that cannot be obtained from the supported release's repository.
- An install script.

A valid SSL certification is recommended, as documented in this procedure.

This procedure is used for an installation or an upgrade of an existing installation.

["Guidelines for upgrading the ONTAP Mediator in a MetroCluster configuration" \(page 95\)](#)

Procedure ▶▶▶

- 1 Enable access to the repositories listed above so Mediator can access the required packages during the installation process.

If your operating system is...	You must provide access to these repositories...
RHEL 7.x	rhel-7-server-optional-rpms
CentOS 7.x	C7.6.1810 - Base repository
RHEL 8.x	<ul style="list-style-type: none"> • rhel-8-for-x86_64-baseos-rpms • rhel-8-for-x86_64-appstream-rpms
CentOS 8.0	kernel-devel

5. Configuring the ONTAP Mediator service for unplanned automatic switchover
 Installing and configuring the ONTAP Mediator service

If your operating system is...	Use these steps...
RHEL 7.x	<p>(1) Subscribe to the required repository: <code>subscription-manager repos --enable rhel-7-server-optional-rpms</code></p> <p>The following example shows the execution of this command:</p> <pre data-bbox="560 405 1471 483">[root@localhost ~]# subscription-manager repos --enable rhel-7-server-optional-rpms Repository 'rhel-7-server-optional-rpms' is enabled for this system.</pre> <p>(2) Run the <code>yum repolist</code> command.</p> <p>The following example shows the execution of this command. The <code>rhel-7-server-optional-rpms</code> repository should appear in the list.</p> <pre data-bbox="560 629 1471 1256">[root@localhost ~]# yum repolist Loaded plugins: product-id, search-disabled-repos, subscription-manager rhel-7-server-optional-rpms 3.2 kB 00:00:00 rhel-7-server-rpms 3.5 kB 00:00:00 (1/3): rhel-7-server-optional-rpms/7Server/x86_64/ group 26 kB 00:00:00 (2/3): rhel-7-server-optional-rpms/7Server/x86_64/ updateinfo 2.5 MB 00:00:00 (3/3): rhel-7-server-optional-rpms/7Server/x86_64/ primary_db 8.3 MB 00:00:01 repo id repo name status rhel-7-server-optional-rpms/7Server/x86_64 Red Hat Enterprise Linux 7 Server - Optional (RPMs) 19,447 rhel-7-server-rpms/7Server/x86_64 Red Hat Enterprise Linux 7 Server (RPMs) 26,758 repolist: 46,205 [root@localhost ~]#</pre>
RHEL 8.x	<p>(1) Subscribe to the required repository: <code>subscription-manager repos --enable rhel-8-for-x86_64-baseos-rpms</code> <code>subscription-manager repos --enable rhel-8-for-x86_64-appstream-rpms</code></p> <p>The following example shows the execution of this command:</p> <pre data-bbox="560 1442 1471 1588">[root@localhost ~]# subscription-manager repos --enable rhel-8-for-x86_64-baseos-rpms [root@localhost ~]# subscription-manager repos --enable rhel-8-for-x86_64-appstream-rpms Repository 'rhel-8-for-x86_64-baseos-rpms' is enabled for this system. Repository 'rhel-8-for-x86_64-appstream-rpms' is enabled for this system.</pre> <p>(2) Run the <code>yum repolist</code> command.</p> <p>The newly subscribed repositories should appear in the list.</p>
CentOS 7.x	<p>Add the C7.6.1810 - Base repository. The C7.6.1810 - Base vault repository contains the kernel-level package needed for ONTAP Mediator.</p> <p>(1) Add the following lines to <code>/etc/yum.repos.d/CentOS-Vault.repo</code>.</p> <pre data-bbox="560 1839 1471 2049">[C7.6.1810-base] name=CentOS-7.6.1810 - Base baseurl=http://vault.centos.org/7.6.1810/os/\$ basearch/gpgcheck=1 gpgkey=file:///etc/pki/rpm-gpg/RPM-GPG-KEY-CentOS-7 enabled=1</pre>

5. Configuring the ONTAP Mediator service for unplanned automatic switchover
 Installing and configuring the ONTAP Mediator service

If your operating system is...	Use these steps...
	<p>(2) Run the <code>yum repolist</code> command.</p> <p>The following example shows the execution of this command. The CentOS-7.6.1810 - Base repository should appear in the list.</p> <pre data-bbox="558 392 1476 1232"> Loaded plugins: fastestmirror Loading mirror speeds from cached hostfile * base: distro.ibiblio.org * extras: distro.ibiblio.org * updates: ewr.edge.kernel.org C7.6.1810-base 3.6 kB 00:00:00 (1/2): C7.6.1810-base/x86_64/group_gz 166 kB 00:00:00 (2/2): C7.6.1810-base/x86_64/primary_db 6.0 MB 00:00:04 repo id repo name status C7.6.1810-base/x86_64 CentOS-7.6.1810 - Base base/7/x86_64 CentOS-7 - Base 10,097 extras/7/x86_64 CentOS-7 - Extras 307 updates/7/x86_64 CentOS-7 - Updates 1,010 repolist: 21,433 [root@localhost ~]# </pre>
CentOS 8.0.1905 or later builds	<p>Because the latest version of the 8.0 (CentOS 8.0.1905) core resides in the CentOS Vault, you must provide access to the matching kernel-level package to compile the needed kernel module.</p> <p>Issue the following command to directly install the kernel-devel package: <code>rpm -i http://vault.centos.org/8.0.1905/BaseOS/x86_64/os/Packages/kernel-devel-\$(uname -r).rpm</code></p> <p>If the system displays an error indicating that the package is already installed, remove the package and try again:</p> <p>(1) Remove the kernel-devel package: <code>yum remove kernel-devel</code></p> <p>(2) Repeat the <code>rpm</code> command shown above.</p>

2 Download the Mediator installation package from the ONTAP Mediator page.
[ONTAP Mediator download page](#)

3 Confirm that the Mediator installation package is in the target directory: `ls`

```
[root@mediator-host ~]#ls
./ontap-mediator_1.2
```

If you are at a location without access to the internet, you must ensure that the installer has access to the required packages.

5. Configuring the ONTAP Mediator service for unplanned automatic switchover
Installing and configuring the ONTAP Mediator service

4 If necessary, move the Mediator installation package from the download directory to the installation directory on the Linux Mediator host.

5 Install the Mediator installation package and respond to the prompts as required: `./ontap-mediator_1.2`

The installation process proceeds to create the required accounts and install required packages. If you have a previous version of Mediator installed on the host, you will be prompted to confirm that you want to upgrade.

The following example shows a fresh installation of the Mediator service:

```
[root@red-hat-enterprise-linux ~]# ./ontap-mediator_1.2
ONTAP Mediator: Self Extracting Installer

ONTAP Mediator requires two user accounts. One for the service (fujitsu), and one for use by ONTAP to the mediator API (mediatoradmin).

Would you like to use the default account names: fujitsu + mediatoradmin? (Y(es)/n(o)): y

Enter ONTAP Mediator system service account (mediatoradmin) password:

Re-Enter ONTAP Mediator system service account (mediatoradmin) password:

Checking for default Linux firewall

Linux firewall is running. Open ports 31784 and 3260? y(es)/n(o): y
success
success

#####

Preparing for installation of ONTAP Mediator packages.

Do you wish to continue? y(es)/n(o): y

+ Installing required packages.

Loaded plugins: product-id, search-disabled-repos, subscription-manager
epel/x86_64/
metalink
| 17 kB 00:00:00
epel-
extra
| 4.9 kB 00:00:00
ius
| 1.3 kB 00:00:00
```

5. Configuring the ONTAP Mediator service for unplanned automatic switchover

Installing and configuring the ONTAP Mediator service

```

rhel-7-server-
rpms
  | 3.5 kB 00:00:00
(1/3): ius/x86_64/
primary
  | 129 kB 00:00:00
(2/3): epel-extra/
group_gz
  | 88 kB 00:00:01
(3/3): epel-extra/
primary_db
  | 6.7 MB 00:00:06
ius
      538/538

Package 1:make-3.82-23.el7.x86_64 already installed and latest version
.
.
.
.

=====
Package                               Arch                               Version
Repository                             Size
=====
Installing:
bzip2                                   x86_64                             1.0.6-13.el7      rhel-7-
server-rpms                             52 k
gcc                                     x86_64                             4.8.5-36.el7_6.2  rhel-7-
server-rpms                             16 M
kernel-devel                            x86_64                             3.10.0-957.el7    rhel-7-
server-rpms                             17 M
openssl-devel                           x86_64                             1:1.0.2k-16.el7_6.1 rhel-7-
server-rpms                             1.5 M
patch                                    x86_64                             2.7.1-10.el7_5    rhel-7-
server-rpms                             110 k
perl-Compress-Raw-Zlib                  x86_64                             1:2.061-4.el7     rhel-7-
server-rpms                             57 k
perl-Digest-MD5                         x86_64                             2.52-3.el7        rhel-7-
server-rpms                             30 k
perl-Digest-SHA                        x86_64                             1:5.85-4.el7      rhel-7-
server-rpms                             58 k
perl-ExtUtils-CBuilder                  noarch                              1:0.28.2.6-294.el7_6 rhel-7-
server-rpms                             68 k
perl-ExtUtils-MakeMaker                 noarch                              6.68-3.el7        rhel-7-
server-rpms                             275 k

```


5. Configuring the ONTAP Mediator service for unplanned automatic switchover

Installing and configuring the ONTAP Mediator service

```

perl-JSON-PP                                noarch                2.27202-2.el7         rhel-7-
server-rpms                                  55 k
python36                                     x86_64                 3.6.8-2.el7.ius
ius                                           60 k
python36-devel                               x86_64                 3.6.8-2.el7.ius
ius                                           206 k
python36-pip                                 noarch                 9.0.1-2.el7.ius
ius                                           1.7 M
redhat-lsb-core                              x86_64                 4.1-27.el7            rhel-7-
server-rpms                                  37 k

Updating:
openssl                                       x86_64                 1:1.0.2k-16.el7_6.1   rhel-7-
server-rpms                                  493 k

Installing for dependencies:
at                                             x86_64                 3.1.13-24.el7         rhel-7-
server-rpms                                  51 k
avahi-libs                                   x86_64                 0.6.31-19.el7         rhel-7-
server-rpms                                  62 k
bc                                             x86_64                 1.06.95-13.el7        rhel-7-
server-rpms                                  115 k
cpp                                            x86_64                 4.8.5-36.el7_6.2     rhel-7-
server-rpms                                  6.0 M
cups-client                                  x86_64                 1:1.6.3-35.el7        rhel-7-
server-rpms                                  151 k
cups-libs                                    x86_64                 1:1.6.3-35.el7        rhel-7-
server-rpms                                  357 k
dwz                                           x86_64                 0.11-3.el7            rhel-7-
server-rpms                                  99 k
ed                                             x86_64                 1.9-4.el7             rhel-7-
server-rpms                                  72 k
gdbm-devel                                   x86_64                 1.10-8.el7            rhel-7-
server-rpms                                  47 k
glibc-devel                                 x86_64                 2.17-260.el7_6.6     rhel-7-
server-rpms                                  1.1 M
glibc-headers                               x86_64                 2.17-260.el7_6.6     rhel-7-
server-rpms                                  684 k
kernel-headers                              x86_64                 3.10.0-957.27.2.el7   rhel-7-
server-rpms                                  8.0 M
keyutils-libs-devel                         x86_64                 1.5.8-3.el7           rhel-7-
server-rpms                                  37 k
krb5-devel                                   x86_64                 1.15.1-37.el7_6      rhel-7-
server-rpms                                  271 k
libcom_err-devel                             x86_64                 1.42.9-13.el7         rhel-7-
server-rpms                                  31 k
libdb-devel                                  x86_64                 5.3.21-24.el7         rhel-7-
server-rpms                                  38 k
libkadm5                                     x86_64                 1.15.1-37.el7_6      rhel-7-
server-rpms                                  178 k
libmpc                                       x86_64                 1.0.1-3.el7           rhel-7-
server-rpms                                  51 k

```

5. Configuring the ONTAP Mediator service for unplanned automatic switchover

Installing and configuring the ONTAP Mediator service

libselinux-devel		x86_64	2.5-14.1.e17	rhel-7-
server-rpms	187 k			
libsepol-devel		x86_64	2.5-10.e17	rhel-7-
server-rpms	77 k			
libverto-devel		x86_64	0.2.5-4.e17	rhel-7-
server-rpms	12 k			
m4		x86_64	1.4.16-10.e17	rhel-7-
server-rpms	256 k			
mailx		x86_64	12.5-19.e17	rhel-7-
server-rpms	245 k			
mpfr		x86_64	3.1.1-4.e17	rhel-7-
server-rpms	203 k			
pcre-devel		x86_64	8.32-17.e17	rhel-7-
server-rpms	480 k			
perl		x86_64	4:5.16.3-294.e17_6	rhel-7-
server-rpms	8.0 M			
perl-Carp		noarch	1.26-244.e17	rhel-7-
server-rpms	19 k			
perl-Data-Dumper		x86_64	2.145-3.e17	rhel-7-
server-rpms	47 k			
perl-Digest		noarch	1.17-245.e17	rhel-7-
server-rpms	23 k			
perl-Encode		x86_64	2.51-7.e17	rhel-7-
server-rpms	1.5 M			
perl-Exporter		noarch	5.68-3.e17	rhel-7-
server-rpms	28 k			
perl-ExtUtils-Install		noarch	1.58-294.e17_6	rhel-7-
server-rpms	75 k			
perl-ExtUtils-Manifest		noarch	1.61-244.e17	rhel-7-
server-rpms	31 k			
perl-ExtUtils-ParseXS		noarch	1:3.18-3.e17	rhel-7-
server-rpms	77 k			
perl-File-Path		noarch	2.09-2.e17	rhel-7-
server-rpms	27 k			
perl-File-Temp		noarch	0.23.01-3.e17	rhel-7-
server-rpms	56 k			
perl-Filter		x86_64	1.49-3.e17	rhel-7-
server-rpms	76 k			
perl-Getopt-Long		noarch	2.40-3.e17	rhel-7-
server-rpms	56 k			
perl-HTTP-Tiny		noarch	0.033-3.e17	rhel-7-
server-rpms	38 k			
perl-IPC-Cmd		noarch	1:0.80-4.e17	rhel-7-
server-rpms	34 k			
perl-Locale-Maketext		noarch	1.23-3.e17	rhel-7-
server-rpms	93 k			
perl-Locale-Maketext-Simple		noarch	1:0.21-294.e17_6	rhel-7-
server-rpms	50 k			
perl-Module-CoreList		noarch	1:2.76.02-294.e17_6	rhel-7-
server-rpms	86 k			
perl-Module-Load		noarch	1:0.24-3.e17	rhel-7-
server-rpms	11 k			

5. Configuring the ONTAP Mediator service for unplanned automatic switchover

Installing and configuring the ONTAP Mediator service

perl-Module-Load-Conditional		noarch	0.54-3.e17	rhel-7-
server-rpms	18 k			
perl-Module-Metadata		noarch	1.000018-2.e17	rhel-7-
server-rpms	26 k			
perl-Params-Check		noarch	1:0.38-2.e17	rhel-7-
server-rpms	18 k			
perl-PathTools		x86_64	3.40-5.e17	rhel-7-
server-rpms	83 k			
perl-Perl-OSType		noarch	1.003-3.e17	rhel-7-
server-rpms	20 k			
perl-Pod-Escapes		noarch	1:1.04-294.e17_6	rhel-7-
server-rpms	51 k			
perl-Pod-Perldoc		noarch	3.20-4.e17	rhel-7-
server-rpms	87 k			
perl-Pod-Simple		noarch	1:3.28-4.e17	rhel-7-
server-rpms	216 k			
perl-Pod-Usage		noarch	1.63-3.e17	rhel-7-
server-rpms	27 k			
perl-Scalar-List-Utils		x86_64	1.27-248.e17	rhel-7-
server-rpms	36 k			
perl-Socket		x86_64	2.010-4.e17	rhel-7-
server-rpms	49 k			
perl-Storable		x86_64	2.45-3.e17	rhel-7-
server-rpms	77 k			
perl-Test-Harness		noarch	3.28-3.e17	rhel-7-
server-rpms	302 k			
perl-Text-ParseWords		noarch	3.29-4.e17	rhel-7-
server-rpms	14 k			
perl-Time-HiRes		x86_64	4:1.9725-3.e17	rhel-7-
server-rpms	45 k			
perl-Time-Local		noarch	1.2300-2.e17	rhel-7-
server-rpms	24 k			
perl-constant		noarch	1.27-2.e17	rhel-7-
server-rpms	19 k			
perl-devel		x86_64	4:5.16.3-294.e17_6	rhel-7-
server-rpms	453 k			
perl-libs		x86_64	4:5.16.3-294.e17_6	rhel-7-
server-rpms	688 k			
perl-macros		x86_64	4:5.16.3-294.e17_6	rhel-7-
server-rpms	44 k			
perl-parent		noarch	1:0.225-244.e17	rhel-7-
server-rpms	12 k			
perl-podlators		noarch	2.5.1-3.e17	rhel-7-
server-rpms	112 k			

5. Configuring the ONTAP Mediator service for unplanned automatic switchover

Installing and configuring the ONTAP Mediator service

perl-srpm-macros		noarch	1-8.e17	rhel-7-
server-rpms	4.7 k			
perl-threads		x86_64	1.87-4.e17	rhel-7-
server-rpms	49 k			
perl-threads-shared		x86_64	1.43-6.e17	rhel-7-
server-rpms	39 k			
perl-version		x86_64	3:0.99.07-3.e17	rhel-7-
server-rpms	84 k			
psmisc		x86_64	22.20-15.e17	rhel-7-
server-rpms	141 k			
pyarsing		noarch	1.5.6-9.e17	rhel-7-
server-rpms	94 k			
python-rpm-macros		noarch	3-24.e17	epel-
extra	7.9 k			
python-srpm-macros		noarch	3-24.e17	epel-
extra	7.3 k			
python3-rpm-macros		noarch	3-24.e17	epel-
extra	6.9 k			
python36-libs		x86_64	3.6.8-2.e17.ius	
ius	8.6 M			
python36-setuptools		noarch	39.2.0-4.e17.ius	
ius	621 k			
redhat-lsb-submod-security		x86_64	4.1-27.e17	rhel-7-
server-rpms	15 k			
redhat-rpm-config		noarch	9.1.0-87.e17	rhel-7-
server-rpms	81 k			
spax		x86_64	1.5.2-13.e17	rhel-7-
server-rpms	260 k			
systemtap-sdt-devel		x86_64	3.3-3.e17	rhel-7-
server-rpms	74 k			
time		x86_64	1.7-45.e17	rhel-7-
server-rpms	30 k			
zip		x86_64	3.0-11.e17	rhel-7-
server-rpms	260 k			
zlib-devel		x86_64	1.2.7-18.e17	rhel-7-
server-rpms	50 k			
Updating for dependencies:				
glibc		x86_64	2.17-260.e17_6.6	rhel-7-
server-rpms	3.6 M			
glibc-common		x86_64	2.17-260.e17_6.6	rhel-7-
server-rpms	11 M			
krb5-libs		x86_64	1.15.1-37.e17_6	rhel-7-
server-rpms	803 k			
libgcc		x86_64	4.8.5-36.e17_6.2	rhel-7-
server-rpms	102 k			
libgomp		x86_64	4.8.5-36.e17_6.2	rhel-7-
server-rpms	158 k			
openssl-libs		x86_64	1:1.0.2k-16.e17_6.1	rhel-7-
server-rpms	1.2 M			

5. Configuring the ONTAP Mediator service for unplanned automatic switchover

Installing and configuring the ONTAP Mediator service

```
Transaction Summary
=====
Install 15 Packages (+84 Dependent packages)
Upgrade 1 Package (+6 Dependent packages)

Total download size: 97 M

Is this ok [y/d/N]: y

Downloading packages:

Delta RPMs disabled because /usr/bin/applydeltarpm not installed.

.
.
.
.

+ Installing ONTAP Mediator. (Log: /tmp/ontap_mediator.jx6UrF/ontap-mediator/install.log)
+ Install successful. (Moving log to /opt/netapp/lib/ontap_mediator/log/install.log)

[root@red-hat-enterprise-linux ~]#
```

The following example shows an upgrade of the Mediator service:

```
[root@scspr1845484003 ontap-mediator]# ./ontap_mediator_1.2
ONTAP Mediator will be upgraded from version 1.1.231618 to 1.2.0.

Do you wish to continue? Y(es)/n(o): y
Capturing support_bundle
  Mediator API User Name: mediatoradmin
  Password:
Running plugins, please wait...

Running 11/11: sysctl...
Creating tar archive...
Support bundle has been generated at /opt/netapp/data/support_bundles/scspr1845484003_1582830730_BASIC.tgz
Testing the DB can be upgraded
Verifying database is up-to-date...
Database out-of-date.Migrations are necessary:
Current revision: bdd9252c33ee, Head revision: 6b2a053cd598
Cloning database...
Cloning successful.
Performing database migrations against clone...
/opt/netapp/lib/ontap_mediator/pyenv/lib64/python3.6/site-packages/alembic/util/messaging.py:70: UserWarning:
Skipping unsupported ALTER for creation of implicit constraint
  warnings.warn(msg)
Migrations successful.
Database upgrade operation complete.
+ Upgrading ONTAP Mediator. (Log: /root/ontap-mediator/upgrade_20200227141143.log)
+ Upgrade successful. (Moving log to /opt/netapp/lib/ontap_mediator/log/upgrade_20200227141143.log)
```

- 6 Use the following steps to configure third-part certification.**
Third-party certification is recommended.

5. Configuring the ONTAP Mediator service for unplanned automatic switchover

Installing and configuring the ONTAP Mediator service

- (1) **The certificate must be placed in the following directory:** `/opt/netapp/lib/ontap_mediator/ontap_mediator/server_config`

You must overwrite the following files with your certificate, renaming the certificate files if necessary to `ontap_mediator_server.crt` and `ontap_mediator_server.key`.

```
# SSL Certificates
cert_path: '/opt/netapp/lib/ontap_mediator/ontap_mediator/server_config/ontap_mediator_server.crt'
key_path: '/opt/netapp/lib/ontap_mediator/ontap_mediator/server_config/ontap_mediator_server.key'
```

You must be sure to follow security best practices for your operating system.
If you are not using a third-party certificate, the Mediator installation process will create a self-signed certificate.

- (2) **Along with the server certificates, you must update the CA certificates as well.**
You must overwrite the following files with your certificate, renaming the certificate files if necessary to:

- `ontap_mediator_server.crt`
- `ontap_mediator_server.key`
- `ca.crt`
- `ca.key`
- `ca.srl`

```
# SSL Certificates
cert_path: '/opt/netapp/lib/ontap_mediator/ontap_mediator/server_config/ontap_mediator_server.crt'
key_path: '/opt/netapp/lib/ontap_mediator/ontap_mediator/server_config/ontap_mediator_server.key'
ca_cert_path: '/opt/netapp/lib/ontap_mediator/ontap_mediator/server_config/ca.crt'
ca_key_path: '/opt/netapp/lib/ontap_mediator/ontap_mediator/server_config/ca.key'
ca_serial_path: '/opt/netapp/lib/ontap_mediator/ontap_mediator/server_config/ca.srl'
```

7 Verify the installation.

- (1) **Run the following command to view the status of the ONTAP Mediator services:** `systemctl`

```
[root@scspr1915530002 ~]# systemctl status ontap_mediator mediator-scst

● ontap_mediator.service - ONTAP Mediator
   Loaded: loaded (/opt/netapp/lib/ontap_mediator/systemd/ontap_mediator.service; enabled; vendor preset: disabled)
   Active: active (running) since Thu 2020-06-18 09:55:02 EDT; 3 days ago
     Main PID: 3559 (uwsgi)
   Status: "uWSGI is ready"
```

5. Configuring the ONTAP Mediator service for unplanned automatic switchover

Installing and configuring the ONTAP Mediator service

```
CGroup: /system.slice/ontap_mediator.service

\u251c\u25003559 /opt/netapp/lib/ontap_mediator/pyenv/bin/uwsgi --ini /opt/netapp/lib/ontap_mediator/uwsgi/ontap_mediator.ini

\u251c\u25004510 /opt/netapp/lib/ontap_mediator/pyenv/bin/uwsgi --ini /opt/netapp/lib/ontap_mediator/uwsgi/ontap_mediator.ini

\u2514\u25004512 /opt/netapp/lib/ontap_mediator/pyenv/bin/uwsgi --ini /opt/netapp/lib/ontap_mediator/uwsgi/ontap_mediator.ini
Jun 18 09:54:43 scspr1915530002 systemd[1]: Starting ONTAP Mediator...

Jun 18 09:54:45 scspr1915530002 ontap_mediator[3559]: [uWSGI] getting INI configuration from /opt/netapp/lib/ontap_mediator/uwsgi/ontap_mediator.ini

Jun 18 09:55:02 scspr1915530002 systemd[1]: Started ONTAP Mediator.

mediator-scst.service
Loaded: loaded (/opt/netapp/lib/ontap_mediator/systemd/mediator-scst.service; enabled; vendor preset: disabled)

Active: active (running) since Thu 2020-06-18 09:54:51 EDT; 3 days ago

Process: 3564 ExecStart=/etc/init.d/scst start (code=exited, status=0/SUCCESS)

Main PID: 4202 (iscsi-scstd)

CGroup: /system.slice/mediator-scst.service

\u2514\u25004202 /usr/local/sbin/iscsi-scstd

Jun 18 09:54:43 scspr1915530002 systemd[1]: Starting mediator-scst.service...

Jun 18 09:54:48 scspr1915530002 iscsi-scstd[4200]: max_data_seg_len 1048576, max_queued_cmds 2048

Jun 18 09:54:51 scspr1915530002 scst[3564]: Loading and configuring SCST[ OK ]

Jun 18 09:54:51 scspr1915530002 systemd[1]: Started mediator-scst.service.

[root@scspr1915530002 ~]#
```

(2) To view the ports the ONTAP Mediator service is using, run: `netstat`

```
[root@scspr1905507001 ~]# netstat -anlt | grep -E '3260|31784'

tcp        0      0 0.0.0.0:31784          0.0.0.0:*              LISTEN
tcp        0      0 0.0.0.0:3260          0.0.0.0:*              LISTEN
tcp6       0      0 :::3260               :::*                    LISTEN
```

Results

The ONTAP Mediator service is now installed and running. Further configuration must be performed in the ONTAP storage system to use the Mediator features.

Configuring the ONTAP Mediator service from a MetroCluster IP configuration

The ONTAP Mediator service must be configured on the ONTAP node for use in a MetroCluster IP configuration.

Before you begin

- The ONTAP Mediator must have been successfully installed on a network location that can be reached by both MetroCluster sites.
- You must have the IP address of the host running the ONTAP Mediator service.
- You must have the username and password for the ONTAP Mediator service.
- All nodes of the MetroCluster IP configuration must be online.

About this task

- This task enables automatic unplanned switchover by default.
- This task can be performed on the ONTAP interface of any node in the MetroCluster IP configuration.
- A single installation of the ONTAP Mediator service can be configured with up to five MetroCluster IP configurations.

Procedure ▶▶▶

- 1 Add the ONTAP Mediator service to ONTAP using the following command: `metrocluster configuration-settings mediator add -mediator-address ip-address-of-mediator-host`

Caution

You will be prompted for the username and password for the Mediator admin user account.

- 2 Verify that the automatic switchover feature is enabled: `metrocluster show`
- 3 Verify that the Mediator is now running.

- (1) Run: `storage disk show -container-type mediator`

```
ClusterA:> storage disk show -container-type mediator
```

Disk	Usable		Disk		Container	Container	
	Size	Shelf	Bay	Type	Type	Name	Owner
NET-1.5	-	-	-	VMDISK	mediator	-	ClusterA-02
NET-1.6	-	-	-	VMDISK	mediator	-	ClusterB-01
NET-1.7	-	-	-	VMDISK	mediator	-	ClusterB-02
NET-1.8	-	-	-	VMDISK	mediator	-	ClusterA-01

- (2) Run the following command: `set advanced`

```
ClusterA:> set advanced
```


(3) **Run:** `storage iscsi-initiator show -label mediator`

```
ClusterA:*> storage iscsi-initiator show -label mediator
(storage iscsi-initiator show)

```

Node	Type	Label	Target	Portal	Target Name	Status

ClusterA-01		mailbox	mediator	1.1.1.1	iqn.2012-05.local:mailbox.target.6616cd3f-9ef1-11e9-aada-00a098ccf5d8:a05e1ffb-9ef1-11e9-8f68-00a098cbca9e:1	up/up
ClusterA-02		mailbox	mediator	1.1.1.1	iqn.2012-05.local:mailbox.target.6616cd3f-9ef1-11e9-aada-00a098ccf5d8:a05e1ffb-9ef1-11e9-8f68-00a098cbca9e:1	up/up



Connecting a MetroCluster configuration to a different ONTAP Mediator instance

If you want to connect the MetroCluster nodes to a different ONTAP Mediator instance, you must unconfigure and then reconfigure the Mediator connection in the ONTAP software.

Before you begin

You need the username, password, and IP address of the new ONTAP Mediator instance.

About this task

These commands can be issued from any node in the MetroCluster configuration.

Procedure ▶▶▶

- 1 **Remove the current ONTAP Mediator from the MetroCluster configuration:** `metrocluster configuration-settings mediator remove`
- 2 **Establish the new ONTAP Mediator connection to the MetroCluster configuration:** `metrocluster configuration-settings mediator add -mediator-address ip-address-of-mediator-host`



Changing the ONTAP Mediator password

After you have installed ONTAP Mediator service, you might want to change the password. You can change the password in two ways.

About this task

This task is performed on the Linux host on which the ONTAP Mediator service is installed.

If you are unable to reach this command, you might need to run the command using the full path as shown in the following example: `/usr/local/bin/mediator_change_password`

Procedure ▶▶▶

1 Change the password by choosing one of the following options:

- Run the `mediator_change_password` command and respond to the prompts as shown in the following example:

```
[root@mediator-host ~]# mediator_change_password
Change the Mediator API password by entering the following values:
  Mediator API User Name: mediatoradmin
    Old Password:
    New Password:
    Confirm Password:
The password has been updated successfully.
[root@mediator-host ~]#
```

- Run the following command:

```
MEDIATOR_USERNAME=mediatoradminMEDIATOR_PASSWORD=mediator1MEDIATOR_NEW_PASSWORD=m
ediator2mediator_change_password
```

The example shows the password is changed from 'mediator1' to 'mediator2'.

```
[root@mediator-host ~]# MEDIATOR_USERNAME=mediatoradmin MEDIATOR_PASSWORD=mediator1
MEDIATOR_NEW_PASSWORD=mediator2 mediator_change_password
The password has been updated successfully.
[root@mediator-host ~]#
```

Changing the ONTAP Mediator user name

After the ONTAP Mediator service is installed, you might want to change the user name.

About this task

This task is performed on the Linux host on which the ONTAP Mediator service is installed.

If you are unable to reach this command, you might need to run the command using the full path as shown in the following example: `/usr/local/bin/mediator_username`

Procedure ▶▶▶

1 Change the ONTAP Mediator user name by choosing from one of the following options:

- Run the command `mediator_change_user` and respond to the prompts as shown in the following example:

```
[root@mediator-host ~]# mediator_change_user
Modify the Mediator API username by entering the following values:
    Mediator API User Name: mediatoradmin
        Password:
New Mediator API User Name: mediator
The account username has been modified successfully.
[root@mediator-host ~]#
```

- Run the following command:

`MEDIATOR_USERNAME=mediatorMEDIATOR_PASSWORD=mediator2MEDIATOR_NEW_USERNAME=mediatoradmin mediator_change_user` as shown in the following example:

```
[root@mediator-host ~]# MEDIATOR_USERNAME= mediator MEDIATOR_PASSWORD='mediator2' MEDIATOR_NEW_USERNAME=
mediatoradmin mediator_change_user
The account username has been modified successfully.
[root@mediator-host ~]#
```



Uninstalling the ONTAP Mediator service

If necessary, you can remove the ONTAP Mediator service.

Before you begin

The Mediator must be disconnected from ONTAP before you remove the Mediator service.

About this task

This task is performed on the Linux host on which the ONTAP Mediator service is installed.

If you are unable to reach this command, you might need to run the command using the full path as shown in the following example:

```
/usr/local/bin/uninstall_ontap_mediator
```

Procedure ▶▶▶

1 Uninstall the ONTAP Mediator service as shown in the following example using the command:

`uninstall_ontap_mediator`

```
[root@mediator-host ~]# uninstall_ontap_mediator

ONTAP Mediator: Self Extracting Uninstaller

+ Removing ONTAP Mediator. (Log: /tmp/ontap_mediator.GmRGdA/uninstall_ontap_mediator/remove.log)
+ Remove successful.
[root@mediator-host ~]#
```



6. Testing the MetroCluster configuration

You can test failure scenarios to confirm the correct operation of the MetroCluster configuration.

Verifying negotiated switchover

You can test the negotiated (planned) switchover operation to confirm uninterrupted data availability.

About this task

This test validates that data availability is not affected (except for Microsoft Server Message Block (SMB) and Solaris Fibre Channel protocols) by switching the cluster over to the second data center.

This test should take about 30 minutes.

This procedure has the following expected results:

- The `metrocluster switchover` command will present a warning prompt. If you respond `yes` to the prompt, the site the command is issued from will switch over the partner site.
- Nodes at the partner site should shut down gracefully and remain at the LOADER prompt.

Procedure ▶▶▶

- 1 **Confirm that all nodes are in the configured state and normal mode:** `metrocluster node show`

```
cluster_A::> metrocluster node show
```

Cluster	Configuration State	Mode
Local: cluster_A	configured	normal
Remote: cluster_B	configured	normal

- 2 **Begin the switchover operation:** `metrocluster switchover`

```
cluster_A::> metrocluster switchover
Warning: negotiated switchover is about to start. It will stop all the data
Vservers on cluster "cluster_B" and
automatically re-start them on cluster "cluster_A". It will finally gracefully
shutdown cluster "cluster_B".
```

- 3 **Confirm that the local cluster is in the configured state and switchover mode:** `metrocluster node show`

```
cluster_A::> metrocluster node show
```

Cluster	Configuration State	Mode
Local: cluster_A	configured	switchover
Remote: cluster_B	not-reachable	-
	configured	normal

4 Confirm that the switchover operation was successful: `metrocluster operation show`

```
cluster_A::> metrocluster operation show

cluster_A::> metrocluster operation show
  Operation: switchover
    State: successful
  Start Time: 2/6/2016 13:28:50
  End Time: 2/6/2016 13:29:41
  Errors: -
```

5 Use the `vserver show` and `network interface show` commands to verify that DR SVMs and LIFs have come online.



Verifying healing and manual switchback

You can test the healing and manual switchback operations to verify that data availability is not affected (except for SMB and Solaris FC configurations) by switching back the cluster to the original data center after a negotiated switchover.

About this task

This test should take about 30 minutes.

The expected result of this procedure is that services should be switched back to their home nodes.

The healing steps are not required on systems running ONTAP 9.7 or later, on which healing is performed automatically after a negotiated switchover. On systems running ONTAP 9.7 and later, healing is also performed automatically after unscheduled switchover.

Procedure ▶▶▶

1 The following example shows the successful completion of the command:

```
cluster_A::> metrocluster heal aggregates
[Job 936] Job succeeded: Heal Aggregates is successful.
```

2 The following example shows the successful completion of the command:

```
cluster_A::> metrocluster heal root-aggregates
[Job 937] Job succeeded: Heal Root Aggregates is successful.
```

3 Verify that healing is completed: `metrocluster node show`

The following example shows the successful completion of the command:

```
cluster_A::> metrocluster node show

DR                               Configuration  DR
Group Cluster Node              State         Mirroring Mode
-----
1   cluster_A
    node_A_1      configured   enabled      heal roots completed
    cluster_B
    node_B_2      unreachable -            switched over

42 entries were displayed.metrocluster operation show
```

4 Verify that all aggregates are mirrored: `storage aggregate show`

The following example shows that all aggregates have a RAID Status of mirrored:

```
cluster_A:> storage aggregate show
cluster Aggregates:
Aggregate Size      Available Used% State  #Vols  Nodes      RAID Status
-----
data_cluster
      4.19TB      4.13TB   2% online    8 node_A_1  raid_dp,
      mirrored,
      normal
root_cluster
      715.5GB     212.7GB  70% online    1 node_A_1  raid4,
      mirrored,
      normal

cluster_B Switched Over Aggregates:
Aggregate Size      Available Used% State  #Vols  Nodes      RAID Status
-----
data_cluster_B
      4.19TB      4.11TB   2% online    5 node_A_1  raid_dp,
      mirrored,
      normal
root_cluster_B     -         -        - unknown    - node_A_1  -
```

5 Boot nodes from the disaster site.

6 Check the status of switchback recovery: `metrocluster node show`

```
cluster_A:> metrocluster node show
DR
Group Cluster Node          Configuration  DR
State          Mirroring Mode
-----
1   cluster_A
    node_A_1      configured    enabled    heal roots completed
    cluster_B
    node_B_2      configured    enabled    waiting for switchback
                                     recovery
2 entries were displayed.
```

7 Perform the switchback: `metrocluster switchback`

```
cluster_A:> metrocluster switchback
[Job 938] Job succeeded: Switchback is successful.Verify switchback
```

8 Confirm status of the nodes: `metrocluster node show`

```
cluster_A::> metrocluster node show
DR                               Configuration  DR
Group Cluster Node                State          Mirroring Mode
-----
1   cluster_A
    node_A_1      configured    enabled      normal
    cluster_B
    node_B_2      configured    enabled      normal

2 entries were displayed.
```

9 Confirm status of the metrocluster operation: `metrocluster operation show` The output should show a successful state.

```
cluster_A::> metrocluster operation show
Operation: switchback
State: successful
Start Time: 2/6/2016 13:54:25
End Time: 2/6/2016 13:56:15
Errors: -
```



Verifying operation after power line disruption

You can test the MetroCluster configuration's response to the failure of a PDU.

About this task

The best practice is for each power supply unit (PSU) in a component to be connected to separate power supplies. If both PSUs are connected to the same power distribution unit (PDU) and an electrical disruption occurs, the site could down or a complete shelf might become unavailable. Failure of one power line is tested to confirm that there is no cabling mismatch that could cause a service disruption.

This test should take about 15 minutes.

This test requires turning off power to all left-hand PDUs and then all right-hand PDUs on all of the racks containing the MetroCluster components.

This procedure has the following expected results:

- Errors should be generated as the PDUs are disconnected.
- No failover or loss of service should occur.

Procedure ▶▶▶

- 1 Turn off the power of the PDUs on the left-hand side of the rack containing the MetroCluster components.

- 2 Monitor the result on the console by using the `system environment sensors show -state fault` and `storage shelf show -errors` commands.

```
cluster_A::> system environment sensors show -state fault

Node Sensor                State Value/Units Crit-Low Warn-Low Warn-Hi Crit-Hi
-----
node_A_1
  PSU1                    fault
                        PSU_OFF
  PSU1 Pwr In OK          fault
                        FAULT
node_A_2
  PSU1                    fault
                        PSU_OFF
  PSU1 Pwr In OK          fault
                        FAULT

4 entries were displayed.

cluster_A::> storage shelf show -errors
  Shelf Name: 1.1
  Shelf UID: 50:0a:09:80:03:6c:44:d5
  Serial Number: SHFHU1443000059

Error Type          Description
-----
Power              Critical condition is detected in storage shelf power
supply unit "1". The unit might fail.Reconnect PSU1
```

- 3 Turn the power back on to the left-hand PDUs.
- 4 Make sure that ONTAP clears the error condition.
- 5 Repeat the previous steps with the right-hand PDUs.



Verifying operation after loss of a single storage shelf

You can test the failure of a single storage shelf to verify that there is no single point of failure.

About this task

This procedure has the following expected results:

- An error message should be reported by the monitoring software.
- No failover or loss of service should occur.
- Mirror resynchronization starts automatically after the hardware failure is restored.

Procedure ▶▶▶

1 Check the storage failover status: storage failover show

```
cluster_A::> storage failover show
```

Node	Partner	Possible State	Description
node_A_1	node_A_2	true	Connected to node_A_2
node_A_2	node_A_1	true	Connected to node_A_1

2 entries were displayed.

2 Check the aggregate status: storage aggregate show

```
cluster_A::> storage aggregate show
```

cluster Aggregates:

Aggregate	Size	Available	Used%	State	#Vols	Nodes	RAID Status
node_A_1data01_mirrored	4.15TB	3.40TB	18%	online	3	node_A_1	raid_dp, mirrored, normal
node_A_1root	707.7GB	34.29GB	95%	online	1	node_A_1	raid_dp, mirrored, normal
node_A_2_data01_mirrored	4.15TB	4.12TB	1%	online	2	node_A_2	raid_dp, mirrored, normal
node_A_2_data02_unmirrored	2.18TB	2.18TB	0%	online	1	node_A_2	raid_dp, normal
node_A_2_root	707.7GB	34.27GB	95%	online	1	node_A_2	raid_dp, mirrored, normal

3 Verify that all data SVMs and data volumes are online and serving data: `vserver show -type datanetwork interface show -fields is-home false volume show !vol0,!MDV*`

```
cluster_A:> vserver show -type data

cluster_A:> vserver show -type data
              Admin      Operational Root
Vserver      Type      Subtype      State      State      Volume      Aggregate
-----
SVM1         data      sync-source      running    SVM1_root  node_A_1_data01_mirrored
SVM2         data      sync-source      running    SVM2_root  node_A_2_data01_mirrored

cluster_A:> network interface show -fields is-home false
There are no entries matching your query.

cluster_A:> volume show !vol0,!MDV*
Vserver      Volume      Aggregate      State      Type      Size      Available Used%
-----
SVM1
  SVM1_root
    node_A_1data01_mirrored
      online    RW      10GB      9.50GB      5%
SVM1
  SVM1_data_vol
    node_A_1data01_mirrored
      online    RW      10GB      9.49GB      5%
SVM2
  SVM2_root
    node_A_2_data01_mirrored
      online    RW      10GB      9.49GB      5%
SVM2
  SVM2_data_vol
    node_A_2_data02_unmirrored
      online    RW      1GB      972.6MB      5%
```

4 Identify a shelf in Pool 1 for node `node_A_2` to power off to simulate a sudden hardware failure:
`storage aggregate show -r -node node-name !*root`
The shelf you select must contain drives that are part of a mirrored data aggregate.

6. Testing the MetroCluster configuration Verifying operation after loss of a single storage shelf

In the following example, shelf ID 31 is selected to fail.

```
cluster_A:> storage aggregate show -r -node node_A_2 !*root
Owner Node: node_A_2
Aggregate: node_A_2_data01_mirrored (online, raid_dp, mirrored) (block checksums)
Plex: /node_A_2_data01_mirrored/plex0 (online, normal, active, pool0)
RAID Group /node_A_2_data01_mirrored/plex0/rg0 (normal, block checksums)

          Usable Physical
Position Disk      Pool Type   RPM    Size    Size Status
-----
dparity 2.30.3      0 BSAS   7200 827.7GB 828.0GB (normal)
parity  2.30.4      0 BSAS   7200 827.7GB 828.0GB (normal)
data    2.30.6      0 BSAS   7200 827.7GB 828.0GB (normal)
data    2.30.8      0 BSAS   7200 827.7GB 828.0GB (normal)
data    2.30.5      0 BSAS   7200 827.7GB 828.0GB (normal)

Plex: /node_A_2_data01_mirrored/plex4 (online, normal, active, pool1)
RAID Group /node_A_2_data01_mirrored/plex4/rg0 (normal, block checksums)

          Usable Physical
Position Disk      Pool Type   RPM    Size    Size Status
-----
dparity 1.31.7      1 BSAS   7200 827.7GB 828.0GB (normal)
parity  1.31.6      1 BSAS   7200 827.7GB 828.0GB (normal)
data    1.31.3      1 BSAS   7200 827.7GB 828.0GB (normal)
data    1.31.4      1 BSAS   7200 827.7GB 828.0GB (normal)
data    1.31.5      1 BSAS   7200 827.7GB 828.0GB (normal)

Aggregate: node_A_2_data02_unmirrored (online, raid_dp) (block checksums)
Plex: /node_A_2_data02_unmirrored/plex0 (online, normal, active, pool0)
RAID Group /node_A_2_data02_unmirrored/plex0/rg0 (normal, block checksums)

          Usable Physical
Position Disk      Pool Type   RPM    Size    Size Status
-----
dparity 2.30.12     0 BSAS   7200 827.7GB 828.0GB (normal)
parity  2.30.22     0 BSAS   7200 827.7GB 828.0GB (normal)
data    2.30.21     0 BSAS   7200 827.7GB 828.0GB (normal)
data    2.30.20     0 BSAS   7200 827.7GB 828.0GB (normal)
data    2.30.14     0 BSAS   7200 827.7GB 828.0GB (normal)

15 entries were displayed.
```

5 Physically power off the shelf that you selected.

6 Check the aggregate status again: `storage aggregate showstorage aggregate show -r -node node_A_2 !*root`

6. Testing the MetroCluster configuration

Verifying operation after loss of a single storage shelf

The aggregate with drives on the powered-off shelf should have a degraded RAID status, and drives on the affected plex should have a failed status, as shown in the following example:

```
cluster_A:> storage aggregate show
Aggregate      Size Available Used% State  #Vols  Nodes      RAID Status
-----
node_A_1data01_mirrored
      4.15TB   3.40TB   18% online    3 node_A_1  raid_dp,
      mirrored,
      normal
node_A_1root
      707.7GB  34.29GB   95% online    1 node_A_1  raid_dp,
      mirrored,
      normal
node_A_2_data01_mirrored
      4.15TB   4.12TB    1% online    2 node_A_2  raid_dp,
      mirror
      degraded
node_A_2_data02_unmirrored
      2.18TB   2.18TB    0% online    1 node_A_2  raid_dp,
      normal
node_A_2_root
      707.7GB  34.27GB   95% online    1 node_A_2  raid_dp,
      mirror
      degraded

cluster_A:> storage aggregate show -r -node node_A_2 !*root
Owner Node: node_A_2
Aggregate: node_A_2_data01_mirrored (online, raid_dp, mirror degraded) (block checksums)
Plex: /node_A_2_data01_mirrored/plex0 (online, normal, active, pool0)
RAID Group /node_A_2_data01_mirrored/plex0/rg0 (normal, block checksums)
          Usable Physical
Position Disk          Pool Type   RPM    Size    Size Status
-----
dparity 2.30.3            0 BSAS   7200  827.7GB  828.0GB (normal)
parity  2.30.4            0 BSAS   7200  827.7GB  828.0GB (normal)
data    2.30.6            0 BSAS   7200  827.7GB  828.0GB (normal)
data    2.30.8            0 BSAS   7200  827.7GB  828.0GB (normal)
data    2.30.5            0 BSAS   7200  827.7GB  828.0GB (normal)

Plex: /node_A_2_data01_mirrored/plex4 (offline, failed, inactive, pool1)
RAID Group /node_A_2_data01_mirrored/plex4/rg0 (partial, none checksums)
          Usable Physical
Position Disk          Pool Type   RPM    Size    Size Status
-----
dparity FAILED          - -        -  827.7GB  - (failed)
parity  FAILED          - -        -  827.7GB  - (failed)
data    FAILED          - -        -  827.7GB  - (failed)
data    FAILED          - -        -  827.7GB  - (failed)
data    FAILED          - -        -  827.7GB  - (failed)
```

6. Testing the MetroCluster configuration

Verifying operation after loss of a single storage shelf

```
Aggregate: node_A_2_data02_unmirrored (online, raid_dp) (block checksums)
Plex: /node_A_2_data02_unmirrored/plex0 (online, normal, active, pool0)
RAID Group /node_A_2_data02_unmirrored/plex0/rg0 (normal, block checksums)

                Usable Physical
Position Disk          Pool Type    RPM    Size    Size Status
-----
dparity  2.30.12          0 BSAS    7200  827.7GB  828.0GB (normal)
parity   2.30.22          0 BSAS    7200  827.7GB  828.0GB (normal)
data     2.30.21          0 BSAS    7200  827.7GB  828.0GB (normal)
data     2.30.20          0 BSAS    7200  827.7GB  828.0GB (normal)
data     2.30.14          0 BSAS    7200  827.7GB  828.0GB (normal)

15 entries were displayed.
```

- 7** Verify that the data is being served and that all volumes are still online: `vserver show -type datanetwork interface show -fields is-home false volume show !vol0,!MDV*`

```
cluster_A:> vserver show -type data

cluster_A:> vserver show -type data
                Admin    Operational Root
Vserver  Type  Subtype  State  State  Volume  Aggregate
-----
SVM1     data  sync-source  running  SVM1_root  node_A_1_data01_mirrored
SVM2     data  sync-source  running  SVM2_root  node_A_1_data01_mirrored

cluster_A:> network interface show -fields is-home false
There are no entries matching your query.

cluster_A:> volume show !vol0,!MDV*
Vserver  Volume  Aggregate  State  Type  Size  Available  Used%
-----
SVM1
    SVM1_root
        node_A_1data01_mirrored
            online  RW    10GB  9.50GB  5%
SVM1
    SVM1_data_vol
        node_A_1data01_mirrored
            online  RW    10GB  9.49GB  5%
SVM2
    SVM2_root
        node_A_1data01_mirrored
            online  RW    10GB  9.49GB  5%
SVM2
    SVM2_data_vol
        node_A_2_data02_unmirrored
            online  RW    1GB   972.6MB  5%
```

- 8** Physically power on the shelf.
Resynchronization starts automatically.
- 9** Verify that resynchronization has started: `storage aggregate show`

6. Testing the MetroCluster configuration

Verifying operation after loss of a single storage shelf

The affected aggregate should have a `resyncing` RAID status, as shown in the following example:

```
cluster_A:> storage aggregate show
cluster Aggregates:
Aggregate      Size Available Used% State  #Vols  Nodes      RAID Status
-----
node_A_1_data01_mirrored
                4.15TB   3.40TB   18% online    3 node_A_1  raid_dp,
                mirrored,
                normal
node_A_1_root
                707.7GB  34.29GB  95% online    1 node_A_1  raid_dp,
                mirrored,
                normal
node_A_2_data01_mirrored
                4.15TB   4.12TB   1%  online    2 node_A_2  raid_dp,
                resyncing
node_A_2_data02_unmirrored
                2.18TB   2.18TB   0%  online    1 node_A_2  raid_dp,
                normal
node_A_2_root
                707.7GB  34.27GB  95% online    1 node_A_2  raid_dp,
                resyncing
```

- 10** Monitor the aggregate to confirm that resynchronization is complete: `storage aggregate show`
The affected aggregate should have a `normal` RAID status, as shown in the following example:

```
cluster_A:> storage aggregate show
cluster Aggregates:
Aggregate      Size Available Used% State  #Vols  Nodes      RAID Status
-----
node_A_1data01_mirrored
                4.15TB   3.40TB   18% online    3 node_A_1  raid_dp,
                mirrored,
                normal
node_A_1root
                707.7GB  34.29GB  95% online    1 node_A_1  raid_dp,
                mirrored,
                normal
node_A_2_data01_mirrored
                4.15TB   4.12TB   1%  online    2 node_A_2  raid_dp,
                normal
node_A_2_data02_unmirrored
                2.18TB   2.18TB   0%  online    1 node_A_2  raid_dp,
                normal
node_A_2_root
                707.7GB  34.27GB  95% online    1 node_A_2  raid_dp,
                resyncing
```



7. Considerations when removing MetroCluster configurations

After removing the MetroCluster configuration, all disk connectivity and interconnects should be adjusted to be in a supported state. If you need to remove the MetroCluster configuration, contact technical support.



You cannot reverse the MetroCluster unconfiguration. This process should only be done with the assistance of technical support.

8. Considerations when using ONTAP in a MetroCluster configuration

When using ONTAP in a MetroCluster configuration, you should be aware of certain considerations for licensing, peering to clusters outside the MetroCluster configuration, performing volume operations, NVFAIL operations, and other ONTAP operations.

- Licensing considerations
 - Both sites should be licensed for the same site-licensed features.
 - All nodes should be licensed for the same node-locked features.
- SnapMirror consideration
 - SnapMirror SVM disaster recovery is only supported on MetroCluster configurations running versions of ONTAP 9.7 or later.

MetroCluster operations in ONTAP System Manager

Depending on your ONTAP version, some MetroCluster-specific operations can be performed using ONTAP System Manager.

- Switchover and switchback in MetroCluster IP configurations (starting in ONTAP 9.7).
- Provision and grow of mirrored aggregates in the MetroCluster IP configurations (starting in ONTAP 9.8). Unmirrored aggregates are not supported in System Manager.

FlexCache support in a MetroCluster configuration

Starting with ONTAP 9.7, FlexCache volumes are supported on MetroCluster configurations. You should be aware of requirements for manual repeer after switchover or switchback operations.

- SVM repeer after switchover when FlexCache origin and cache are within the same MetroCluster site
After a negotiated or unplanned switchover, any SVM FlexCache peering relationship within the cluster must be manually configured.
For example, SVMs vs1 (cache) and vs2 (origin) are on site_A. These SVMs are peered.
After switchover, SVMs vs1-mc and vs2-mc are activated at the partner site (site_B). They must be manually repeer for FlexCache to work using the `vserver peer repeer` command.
- SVM repeer after switchover or switchback when a FlexCache destination is on a third cluster and in disconnected mode
For FlexCache relationships to a cluster outside of the MetroCluster configuration, the peering must always be manually reconfigured after a switchover if the involved clusters are in disconnected mode during switchover.
For example:
 - One end of the FlexCache (cache_1 on vs1) resides on MetroCluster site_A has one end of the FlexCache
 - The other end of the FlexCache (origin_1 on vs2) resides on site_C (not in the MetroCluster configuration)When switchover is triggered, and if site_A and site_C are not connected, you must manually repeer the SVMs on site_B (the switchover cluster) and site_C using the `vserver peer repeer` command after the switchover.
When switchback is performed, you must again repeer the SVMs on site_A (the original cluster) and site_C.

Related information

[FlexCache® Volumes for Faster Data Access Power Guide](#)

FabricPool support in MetroCluster configurations

Starting with ONTAP 9.7, MetroCluster configurations support FabricPool storage tiers. For general information on using FabricPools, see the *Disks and Aggregates Power Guide*.

[Disks and Aggregates Power Guide](#)

- Considerations when using FabricPools
 - The clusters must have FabricPool licenses with matching capacity limits.
 - The clusters must have IPSpaces with matching names.
This can be the default IPspace, or an IP space an administrator has created. This IPspace will be used for FabricPool object store configuration setups.
 - For the selected IPspace, each cluster must have an intercluster LIF defined that can reach the external object store
- Configuring an aggregate for use in a mirrored FabricPool

Caution

Before you configure the aggregate you must set up object stores as described in "Setting up object stores for FabricPool in a MetroCluster configuration" in the *Disks and Aggregates Power Guide*.

[Disks and Aggregates Power Guide](#)

To configure an aggregate for use in a FabricPool:

- 1 Create the aggregate or select an existing aggregate.
- 2 Mirror the aggregate as a typical mirrored aggregate within the MetroCluster configuration.
- 3 Create the FabricPool mirror with the aggregate, as described in the *Disks and Aggregates Power Guide*:
[Disks and Aggregates Power Guide](#)
 - (1) Attach a primary object store.
This object store is physically closer to the cluster.
 - (2) Add a mirror object store.
This object store is physically further distant to the cluster than the primary object store.

Job schedules in a MetroCluster configuration

In ONTAP 9.7 and later, user-created job schedules are automatically replicated between clusters in a MetroCluster configuration. If you create, modify, or delete a job schedule on a cluster, the same schedule is automatically created on the partner cluster, using Configuration Replication Service (CRS).

Caution

System-created schedules are not replicated and you must manually perform the same operation on the partner cluster so that job schedules on both clusters are identical.

Cluster peering from the MetroCluster site to a third cluster

Because the peering configuration is not replicated, if you peer one of the clusters in the MetroCluster configuration to a third cluster outside of that configuration, you must also configure the peering on the partner MetroCluster cluster. This is so that peering can be maintained if a switchover occurs.

The non-MetroCluster cluster must be running ONTAP 9.7 or later. If not, peering is lost if a switchover occurs even if the peering has been configured on both MetroCluster partners.

LDAP client configuration replication in a MetroCluster configuration

An LDAP client configuration created on a storage virtual machine (SVM) on a local cluster is replicated to its partner data SVM on the remote cluster. For example, if the LDAP client configuration is created on the admin SVM on the local cluster, then it is replicated to all the admin data SVMs on the remote cluster. This MetroCluster feature is intentional so that the LDAP client configuration is active on all the partner SVMs on the remote cluster.

Networking and LIF creation guidelines for MetroCluster configurations

You should be aware of how LIFs are created and replicated in a MetroCluster configuration. You must also know about the requirement for consistency so that you can make proper decisions when configuring your network.

Related information

[Network Management Guide](#)

[IPspace object replication and subnet configuration requirements \(page 130\)](#)

[Requirements for LIF creation in a MetroCluster configuration \(page 131\)](#)

[LIF replication and placement requirements and issues \(page 131\)](#)

IPspace object replication and subnet configuration requirements

You should be aware of the requirements for replicating IPspace objects to the partner cluster and for configuring subnets and IPv6 in a MetroCluster configuration.

● IPspace replication

You must consider the following guidelines while replicating IPspace objects to the partner cluster:

- The IPspace names of the two sites must match.
- IPspace objects must be manually replicated to the partner cluster.
Any storage virtual machines (SVMs) that are created and assigned to an IPspace before the IPspace is replicated will not be replicated to the partner cluster.

● Subnet configuration

You must consider the following guidelines while configuring subnets in a MetroCluster configuration:

- Both clusters of the MetroCluster configuration must have a subnet in the same IPspace with the same subnet name, subnet, broadcast domain, and gateway.
- The IP ranges of the two clusters must be different.

In the following example, the IP ranges are different:

```
cluster_A::> network subnet show

IPspace: Default
Subnet          Broadcast          Avail/
Name            Subnet            Domain            Gateway          Total    Ranges
-----
subnet1         192.168.2.0/24    Default          192.168.2.1     10/10    192.168.2.11-192.168.2.20

cluster_B::> network subnet show

IPspace: Default
Subnet          Broadcast          Avail/
Name            Subnet            Domain            Gateway          Total    Ranges
-----
subnet1         192.168.2.0/24    Default          192.168.2.1     10/10    192.168.2.21-192.168.2.30
```

- IPv6 configuration
If IPv6 is configured on one site, IPv6 must be configured on the other site as well.

Related information

- [Requirements for LIF creation in a MetroCluster configuration \(page 131\)](#)
- [LIF replication and placement requirements and issues \(page 131\)](#)

Requirements for LIF creation in a MetroCluster configuration

You should be aware of the requirements for creating LIFs when configuring your network in a MetroCluster configuration.

You must consider the following guidelines when creating LIFs:

- Fibre Channel: You must use stretched VSAN or stretched fabrics
- IP/iSCSI: You must use layer 2 stretched network
- ARP broadcasts: You must enable ARP broadcasts between the two clusters
- Duplicate LIFs: You must not create multiple LIFs with the same IP address (duplicate LIFs) in an IPspace
- NFS and SAN configurations: You must use different storage virtual machines (SVMs) for both the unmirrored and mirrored aggregates

- Verify LIF creation
You can confirm the successful creation of a LIF in a MetroCluster configuration by running the `metrocluster check lif show` command. If you encounter any issues while creating the LIF, you can use the `metrocluster check lif repair-placement` command to fix the issues.

Related information

- [IPspace object replication and subnet configuration requirements \(page 130\)](#)
- [LIF replication and placement requirements and issues \(page 131\)](#)

LIF replication and placement requirements and issues

You should be aware of the LIF replication requirements in a MetroCluster configuration. You should also know how a replicated LIF is placed on a partner cluster, and you should be aware of the issues that occur when LIF replication or LIF placement fails.

● Replication of LIFs to the partner cluster

When you create a LIF on a cluster in a MetroCluster configuration, the LIF is replicated on the partner cluster. LIFs are not placed on a one-to-one name basis. For availability of LIFs after a switchover operation, the LIF placement process verifies that the ports are able to host the LIF based on reachability and port attribute checks.

The system must meet the following conditions to place the replicated LIFs on the partner cluster:

Condition	LIF type: FC	LIF type: IP/iSCSI
Node identification	ONTAP attempts to place the replicated LIF on the disaster recovery (DR) partner of the node on which it was created. If the DR partner is unavailable, the DR auxiliary partner is used for placement.	ONTAP attempts to place the replicated LIF on the DR partner of the node on which it was created. If the DR partner is unavailable, the DR auxiliary partner is used for placement.
Port identification	ONTAP identifies the connected FC target ports on the DR cluster.	The ports on the DR cluster that are in the same IPspace as the source LIF are selected for a reachability check. If there are no ports in the DR cluster in the same IPspace, the LIF cannot be placed. All of the ports in the DR cluster that are already hosting a LIF in the same IPspace and subnet are automatically marked as reachable; and can be used for placement. These ports are not included in the reachability check.
Reachability check	Reachability is determined by checking for the connectivity of the source fabric WWN on the ports in the DR cluster. If the same fabric is not present at the DR site, the LIF is placed on a random port on the DR partner.	Reachability is determined by the response to an Address Resolution Protocol (ARP) broadcast from each previously identified port on the DR cluster to the source IP address of the LIF to be placed. For reachability checks to succeed, ARP broadcasts must be allowed between the two clusters. Each port that receives a response from the source LIF will be marked as possible for placement.

8. Considerations when using ONTAP in a MetroCluster configuration
Networking and LIF creation guidelines for MetroCluster configurations

Condition	LIF type: FC	LIF type: IP/iSCSI
Port selection	<p>ONTAP categorizes the ports based on attributes such as adapter type and speed, and then selects the ports with matching attributes.</p> <p>If no ports with matching attributes are found, the LIF is placed on a random connected port on the DR partner.</p>	<p>From the ports that are marked as reachable during the reachability check, ONTAP prefers ports that are in the broadcast domain that is associated with the subnet of the LIF.</p> <p>If there are no network ports available on the DR cluster that are in the broadcast domain that is associated with the subnet of the LIF, then ONTAP selects ports that have reachability to the source LIF.</p> <p>If there are no ports with reachability to the source LIF, a port is selected from the broadcast domain that is associated with the subnet of the source LIF, and if no such broadcast domain exists, a random port is selected.</p> <p>ONTAP categorizes the ports based on attributes such as adapter type, interface type, and speed, and then selects the ports with matching attributes.</p>
LIF placement	From the reachable ports, ONTAP selects the least loaded port for placement.	From the selected ports, ONTAP selects the least loaded port for placement.

- **Placement of replicated LIFs when the DR partner node is down**
When an iSCSI or FC LIF is created on a node whose DR partner has been taken over, the replicated LIF is placed on the DR auxiliary partner node. After a subsequent giveback operation, the LIFs are not automatically moved to the DR partner. This can lead to LIFs being concentrated on a single node in the partner cluster. During a MetroCluster switchover operation, subsequent attempts to map LUNs belonging to the storage virtual machine (SVM) fail.
You should run the `metrocluster check lif show` command after a takeover operation or giveback operation to verify that the LIF placement is correct. If errors exist, you can run the `metrocluster check lif repair-placement` command to resolve the issues.
- **LIF placement errors**
LIF placement errors that are displayed by the `metrocluster check lif show` command are retained after a switchover operation. If the `network interface modify`, `network interface rename`, or `network interface delete` command is issued for a LIF with a placement error, the error is removed and does not appear in the output of the `metrocluster check lif show` command.
- **LIF replication failure**
You can also check whether LIF replication was successful by using the `metrocluster check lif show` command. An EMS message is displayed if LIF replication fails.
You can correct a replication failure by running the `metrocluster check lif repair-placement` command for any LIF that fails to find a correct port. You should resolve any LIF replication failures as soon as possible to verify the availability of LIF during a MetroCluster switchover operation.

Caution

Even if the source SVM is down, LIF placement might proceed normally if there is a LIF belonging to a different SVM in a port with the same IPspace and network in the destination SVM.

Related information

[IPspace object replication and subnet configuration requirements \(page 130\)](#)

[Requirements for LIF creation in a MetroCluster configuration \(page 131\)](#)

Volume creation on a root aggregate

The system does not allow the creation of new volumes on the root aggregate (an aggregate with an HA policy of `CFO`) of a node in a MetroCluster configuration.

Because of this restriction, root aggregates cannot be added to an SVM using the `vserver add-aggregates` command.

SVM disaster recovery in a MetroCluster configuration

The destination SVM must be on the third cluster outside of the MetroCluster configuration.

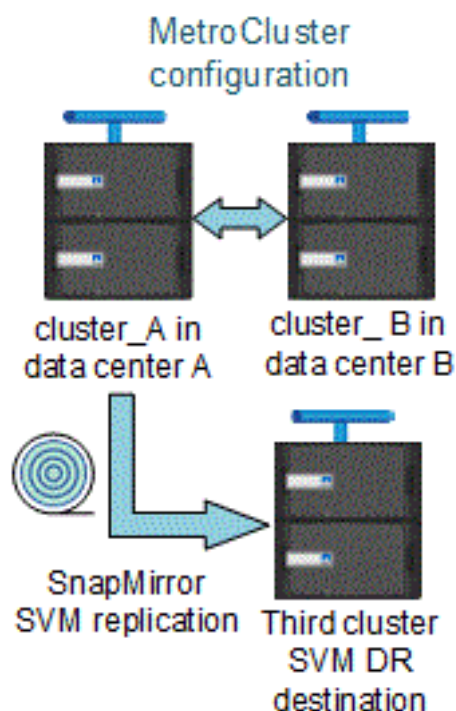
You should be aware of the following requirements and limitations of using SVMs with SnapMirror disaster recovery:

- Only an active SVM within a MetroCluster configuration can be the source of an SVM disaster recovery relationship.

A source can be a sync-source SVM before switchover or a sync-destination SVM after switchover.

- When a MetroCluster configuration is in a steady state, the MetroCluster sync-destination SVM cannot be the source of an SVM disaster recovery relationship, since the volumes are not online.

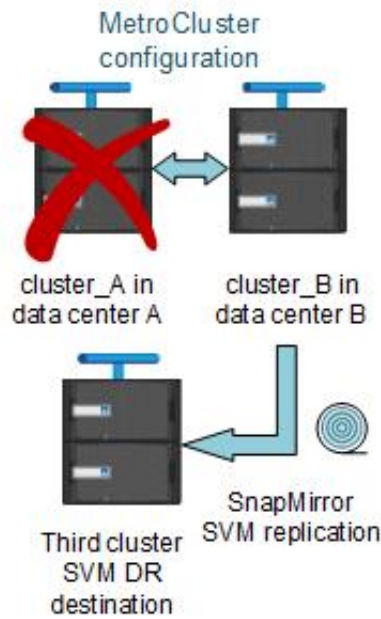
The following image shows the SVM disaster recovery behavior in a steady state:



- When the sync-source SVM is the source of an SVM DR relationship, the source SVM DR relationship information is replicated to the MetroCluster partner.

8. Considerations when using ONTAP in a MetroCluster configuration
SVM disaster recovery in a MetroCluster configuration

This enables the SVM DR updates to continue after a switchover as shown in the following image:



- During the switchover and switchback processes, replication to the SVM DR destination might fail. However, after the switchover or switchback process completes, the next SVM DR scheduled updates will succeed.

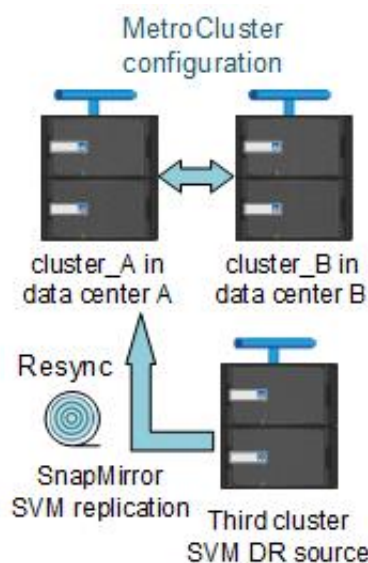
See the section "Replicating the SVM configuration" in the *Data Protection Power Guide* for details on configuring an SVM DR relationship.

[Data Protection Power Guide](#)

SVM resynchronization at a disaster recovery site

During resynchronization, the storage virtual machines (SVMs) disaster recovery (DR) source on the MetroCluster configuration is restored from the destination SVM on the non-MetroCluster site.

During resynchronization, the source SVM (cluster_A) temporarily acts as a destination SVM as shown in the following image:



- If an unplanned switchover occurs during resynchronization

Unplanned switchovers that occur during the resynchronization will halt the resynchronization transfer. If an unplanned switchover occurs, the following conditions are true:

- The destination SVM on the MetroCluster site (which was a source SVM prior to resynchronization) remains as a destination SVM. The SVM at the partner cluster will continue to retain its subtype and remain inactive.
- The SnapMirror relationship must be re-created manually with the sync-destination SVM as the destination.
- The SnapMirror relationship does not appear in the SnapMirror show output after a switchover at the survivor site unless a SnapMirror create operation is executed.

- Performing switchback after an unplanned switchover during resynchronization

To successfully perform the switchback process, the resynchronization relationship must be broken and deleted. Switchback is not permitted if there are any SnapMirror DR destination SVMs in the MetroCluster configuration or if the cluster has an SVM of subtype "*dp-destination*".

Output for the storage aggregate plex show command is indeterminate after a MetroCluster switchover

When you run the `storage aggregate plex show` command after a MetroCluster switchover, the status of `plex0` of the switched over root aggregate is indeterminate and is displayed as `failed`. During this time, the switched over root is not updated. The actual status of this plex can only be determined after the MetroCluster healing phase.

Modifying volumes to set the NVFAIL flag in case of switchover

You can modify a volume so that the NVFAIL flag is set on the volume in the event of a MetroCluster switchover. The NVFAIL flag causes the volume to be fenced off from any modification. This is required for volumes that need to be handled as if committed writes to the volume were lost after the switchover.

Procedure ▶▶▶

- 1 **Enable MetroCluster configuration to trigger NVFAIL on switchover by setting the `vol -dr-force-nvfail` parameter to on:** `vol modify -vserver vserver-name -volume volume-name -dr-force-nvfail on`



Monitoring and protecting the file system consistency using NVFAIL

The `-nvfail` parameter of the `volume modify` command enables ONTAP to detect nonvolatile RAM (NVRAM) inconsistencies when the system is booting or after a switchover operation. It also warns you and protects the system against data access and modification until the volume can be manually recovered.

If ONTAP detects any problems, database or file system instances stop responding or shut down. ONTAP then sends error messages to the console to alert you to check the state of the database or file system. You can enable NVFAIL to warn database administrators of NVRAM inconsistencies among clustered nodes that can compromise database validity.

8. Considerations when using ONTAP in a MetroCluster configuration

Monitoring and protecting the file system consistency using NVFAIL

After the NVRAM data loss during failover or boot recovery, NFS clients cannot access data from any of the nodes until the NVFAIL state is cleared. CIFS clients are unaffected.

How NVFAIL impacts access to NFS volumes or LUNs

The NVFAIL state is set when ONTAP detects NVRAM errors when booting, when a MetroCluster switchover operation occurs, or during an HA takeover operation if the NVFAIL option is set on the volume. If no errors are detected at startup, the file service is started normally. However, if NVRAM errors are detected or NVFAIL processing is enforced on a disaster switchover, ONTAP stops database instances from responding.

When you enable the `NVFAIL` option, one of the processes described in the following table takes place during bootup:

If...	Then...
ONTAP detects no NVRAM errors	File service starts normally.
ONTAP detects NVRAM errors	<ul style="list-style-type: none">• ONTAP returns a stale file handle (<code>ESTALE</code>) error to NFS clients trying to access the database, causing the application to stop responding, crash, or shut down. ONTAP then sends an error message to the system console and log file.• When the application restarts, files are available to CIFS clients even if you have not verified that they are valid. For NFS clients, files remain inaccessible until you reset the <code>in-nvfailed-state</code> option on the affected volume.
If one of the following parameters is used: <ul style="list-style-type: none">• <code>dr-force-nvfail</code> volume option is set• <code>force-nvfail-all</code> switchover command option is set.	<p>You can unset the <code>dr-force-nvfail</code> option after the switchover, if the administrator is not expecting to force NVFAIL processing for possible future disaster switchover operations.</p> <p>For NFS clients, files remain inaccessible until you reset the <code>in-nvfailed-state</code> option on the affected volume.</p> <div style="background-color: #fff9c4; padding: 5px;"><p>Caution</p><p>Using the <code>force-nvfail-all</code> option causes the <code>dr-force-nvfail</code> option to be set on all of the DR volumes processed during the disaster switchover.</p></div>
ONTAP detects NVRAM errors on a volume that contains LUNs	<p>LUNs in that volume are brought offline. The <code>in-nvfailed-state</code> option on the volume must be cleared, and the <code>NVFAIL</code> attribute on the LUNs must be cleared by bringing each LUN in the affected volume online.</p> <p>You can perform the steps to check the integrity of the LUNs and recover the LUN from a Snapshot copy or back up as necessary. After all of the LUNs in the volume are recovered, the <code>in-nvfailed-state</code> option on the affected volume is cleared.</p>

Commands for monitoring data loss events

If you enable the `NVFAIL` option, you receive notification when a system crash caused by NVRAM inconsistencies or a MetroCluster switchover occurs.

By default, the NVFAIL parameter is not enabled.

If you want to...	Use this command...
Create a new volume with NVFAIL enabled	<code>volume create -nvfail on</code>
Enable NVFAIL on an existing volume	<code>volume modify</code> Caution You set the <code>-nvfail</code> option to <code>on</code> to enable NVFAIL on the created volume.
Display whether NVFAIL is currently enabled for a specified volume	<code>volume show</code> Caution You set the <code>-fields</code> parameter to <code>nvfail</code> to display the NVFAIL attribute for a specified volume.

See the man page for each command for more information.

Accessing volumes in NVFAIL state after a switchover

After a switchover, you must clear the NVFAIL state by resetting the `-in-nvfailed-state` parameter of the `volume modify` command to remove the restriction of clients to access data.

Before you begin

The database or file system must not be running or trying to access the affected volume.

About this task

Setting `-in-nvfailed-state` parameter requires advanced-level privilege.

Procedure ▶▶▶

- 1 Recover the volume by using the `volume modify` command with the `-in-nvfailed-state` parameter set to `false`.

What to do next

For instructions about examining database file validity, see the documentation for your specific database software.

If your database uses LUNs, review the steps to make the LUNs accessible to the host after an NVRAM failure.

Recovering LUNs in NVFAIL states after switchover

After a switchover, the host no longer has access to data on the LUNs that are in NVFAIL states. You must perform a number of actions before the database has access to the LUNs.

Before you begin

The database must not be running.

Procedure ▶▶▶

- 1 Clear the NVFAIL state on the affect volume that hosts the LUNs by resetting the `-in-nvfailed-state` parameter of the `volume modify` command.

- 2** Bring the affected LUNs online.
 - 3** Examine the LUNs for any data inconsistencies and resolve them.
This might involve host-based recovery or recovery done on the storage controller using SnapRestore.
 - 4** Bring the database application online after recovering the LUNs.
-



9. Where to find additional information

You can learn more about MetroCluster IP configuration and operation from the ETERNUS AX/HX Series Manual Page library.

- MetroCluster and miscellaneous guides

Guide	Content
Fujitsu Manual Site	<ul style="list-style-type: none"> • All MetroCluster guides
MetroCluster® Management and Disaster Recovery Guide	<ul style="list-style-type: none"> • Understanding the MetroCluster configuration • Switchover, healing and switchback • Disaster recovery
MetroCluster® Service Guide	<ul style="list-style-type: none"> • Guidelines for maintenance in a MetroCluster FC configuration • Hardware replacement or upgrade and firmware upgrade procedures for FC-to-SAS bridges and FC switches • Hot-adding a disk shelf in a fabric-attached or stretch MetroCluster FC configuration • Hot-removing a disk shelf in a fabric-attached or stretch MetroCluster FC configuration • Replacing hardware at a disaster site in a fabric-attached or stretch MetroCluster FC configuration • Expanding a two-node fabric-attached or stretch MetroCluster FC configuration to a four-node MetroCluster configuration. • Expanding a four-node fabric-attached or stretch MetroCluster FC configuration to an eight-node MetroCluster FC configuration.
MetroCluster Upgrade, Transition, and Expansion Guide	<ul style="list-style-type: none"> • Upgrading or refreshing a MetroCluster configuration • Transitioning from a MetroCluster FC configuration to a MetroCluster IP configuration • Expanding a MetroCluster configuration by adding additional nodes
Tiebreaker Software 1.21 Installation and Configuration Guide	<ul style="list-style-type: none"> • Monitoring the MetroCluster configuration with the MetroCluster Tiebreaker software
Fujitsu Manual Site <div style="background-color: #fff9c4; padding: 10px; margin-top: 10px;"> <p>Caution</p> <p>The standard storage shelf maintenance procedures can be used with MetroCluster IP configurations.</p> </div>	<ul style="list-style-type: none"> • Hot-adding a disk shelf • Hot-removing a disk shelf
7-Mode Transition Tool Copy-Based Transition Guide For Transitioning to ONTAP®	<ul style="list-style-type: none"> • Transitioning data from 7-Mode storage systems to clustered storage systems
Concepts Guide	<ul style="list-style-type: none"> • How mirrored aggregates work

Copyright and trademark

Copyright

Copyright 2021 FUJITSU LIMITED. All rights reserved.

No part of this document covered by copyright may be reproduced in any form or by any means - graphic, electronic, or mechanical, including photocopying, recording, taping, or storage in an electronic retrieval system - without prior written permission of the copyright owner.

Software derived from copyrighted Fujitsu material is subject to the following license and disclaimer:

THIS SOFTWARE IS PROVIDED BY FUJITSU "AS IS" AND WITHOUT ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE, WHICH ARE HEREBY DISCLAIMED. IN NO EVENT SHALL FUJITSU BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

Fujitsu reserves the right to change any products described herein at any time, and without notice. Fujitsu assumes no responsibility or liability arising from the use of products described herein, except as expressly agreed to in writing by Fujitsu. The use or purchase of this product does not convey a license under any patent rights, trademark rights, or any other intellectual property rights of Fujitsu.

Trademark

FUJITSU, FUJITSU logo, and ETERNUS are trademarks of Fujitsu. All other trademarks are the property of their respective owners.

<https://www.fujitsu.com/global/products/computing/storage/eternus/trademarks.html>

How to send comments about documentation and receive update notifications

The latest version of this document and the latest information related to this device are available at the following site.

<https://www.fujitsu.com/global/support/products/computing/storage/manuals-list.html>

If necessary, refer to the manuals for your model.

FUJITSU Storage ETERNUS AX/HX Series
MetroCluster® IP Installation and Configuration Guide

A3CA08733-A234-04

Date of issuance: September 2021

Issuance responsibility: FUJITSU LIMITED

- The content of this manual is subject to change without notice.
- This manual was prepared with the utmost attention to detail.
However, Fujitsu shall assume no responsibility for any operational problems as the result of errors, omissions, or the use of information in this manual.
- Fujitsu assumes no liability for damages to third party copyrights or other rights arising from the use of any information in this manual.
- The content of this manual may not be reproduced or distributed in part or in its entirety without prior permission from Fujitsu.


FUJITSU