# Contents

Deciding whether to use this guide ............................................................. 6

**Preparing for the MetroCluster installation** ........................................ 7

- Differences between the ONTAP MetroCluster configurations ............ 7
- Access to remote storage in MetroCluster IP configurations .................. 8
- Considerations for MetroCluster IP configuration ...................................... 8
- Considerations for automatic drive assignment and ADP systems in ONTAP 9.4 and later ................................................................. 9
  - ADP and disk assignment differences by system in MetroCluster IP configurations ........................................... 11
- Considerations for configuring cluster peering ......................................... 15
- Prerequisites for cluster peering ............................................................... 16
- Considerations when using dedicated ports ............................................... 17
- Considerations when sharing data ports .................................................... 17
- Considerations for sharing private layer 2 networks ................................. 18
- MetroCluster ISL requirements in shared networks ................................. 18
- ISL cabling requirements .................................................................... 19
- Required settings on intermediate switches ............................................ 20
- Supported and unsupported shared network configuration types ............ 21
- Considerations for using TDM/xWDM and encryption equipment with MetroCluster IP configurations ....................................................... 27
- Preconfigured settings for new MetroCluster systems from the factory .... 28
  - Hardware setup checklist ................................................................ 28

**Configuring the MetroCluster hardware components** ......................... 30

- Parts of a MetroCluster IP configuration .................................................. 30
- Illustration of the local HA pairs in a MetroCluster configuration ............. 32
- Illustration of the MetroCluster IP and cluster interconnect network .......... 32
- Illustration of the cluster peering network .............................................. 34
- Required MetroCluster IP components and naming conventions ............ 35
- Installing and cabling MetroCluster components .................................... 38
- Racking the hardware components .......................................................... 38
- Cabling the IP switches ....................................................................... 39
- Cabling the cluster peering connections ................................................ 49
- Cabling the management and data connections ....................................... 50
- Configuring the IP switches ................................................................ 50

**Configuring the MetroCluster software in ONTAP** ............................. 68

- Gathering required information ............................................................... 69
- IP network information worksheet for site A ......................................... 69
- IP network information worksheet for site B ......................................... 71
- Similarities and differences between standard cluster and MetroCluster configurations ................................................................. 73
- Restoring system defaults on a previously used controller module ......... 74
Verifying the ha-config state of components ............................................................. 75
Manually assigning drives to pool 0 .......................................................................... 75
Manually assigning drives for pool 0 (ONTAP 9.4 and later) ............................ 76
Manually assigning drives for pool 0 (ONTAP 9.3) ........................................ 77
Setting up ONTAP ..................................................................................................... 79
Configuring the clusters into a MetroCluster configuration ............................... 83
  Disabling automatic drive assignment (if doing manual assignment in
  ONTAP 9.4) ........................................................................................................ 83
  Verifying drive assignment of pool 0 drives ........................................................ 83
  Peering the clusters ............................................................................................ 84
  Creating the DR group ....................................................................................... 90
  Configuring and connecting the MetroCluster IP interfaces ............................ 92
  Verifying or manually performing pool 1 drives assignment ......................... 98
  Enabling automatic drive assignment in ONTAP 9.4 ...................................... 104
  Mirroring the root aggregates .......................................................................... 104
  Creating a mirrored data aggregate on each node .......................................... 105
  Implementing the MetroCluster configuration ................................................ 106
  Checking the MetroCluster configuration ....................................................... 108
  Completing ONTAP configuration ................................................................ 110
Verifying switchover, healing, and switchback .................................................. 110
Installing the MetroCluster Tiebreaker software ............................................... 110
Protecting configuration backup files .................................................................... 111

Testing the MetroCluster configuration .............................................................. 112
  Verifying negotiated switchover ....................................................................... 112
  Verifying healing and manual switchback ....................................................... 113
  Verifying operation after power line disruption ............................................. 115
  Verifying operation after loss of a single storage shelf ................................... 116

Considerations when removing MetroCluster configurations ...................... 122
Considerations when using ONTAP in a MetroCluster configuration
............................................................................................................................ 123
  FlexGroup support in MetroCluster configurations ....................................... 123
  Job schedules in a MetroCluster configuration ............................................... 123
  Cluster peering from the MetroCluster site to a third cluster ....................... 123
  LDAP client configuration replication in a MetroCluster configuration .......... 123
  Networking and LIF creation guidelines for MetroCluster configurations ...... 124
    IPspace object replication and subnet configuration requirements ................. 124
    Requirements for LIF creation in a MetroCluster configuration ................. 125
    LIF replication and placement requirements and issues .............................. 125
    Volume creation on a root aggregate .............................................................. 128
  SVM disaster recovery in a MetroCluster configuration .................................... 128
    SVM resynchronization at a disaster recovery site ...................................... 130
Output for the storage aggregate plex show command is indeterminate after a 130
  MetroCluster switchover .................................................................................. 130
Modifying volumes to set the NVFAIL flag in case of switchover ...................... 131
Monitoring and protecting the file system consistency using NVFAIL .......... 131
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.

ONTAP 9 Documentation Center
Preparing for the MetroCluster installation

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

Differences between 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.

<table>
<thead>
<tr>
<th>Feature</th>
<th>IP configurations</th>
<th>Fabric-attached configurations</th>
<th>Stretch configurations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Four or eight-node</td>
<td>Two-node</td>
</tr>
<tr>
<td>Number of controllers</td>
<td>Four</td>
<td>Four or eight</td>
<td>Two</td>
</tr>
<tr>
<td>Uses an FC switch storage fabric</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Uses an IP switch storage fabric</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Uses FC-to-SAS bridges</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Uses direct-attached SAS storage</td>
<td>Yes (local attached only)</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Supports ADP</td>
<td>Yes (starting in ONTAP 9.4)</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Supports local HA</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Supports automatic switchover</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Supports unmirrored aggregates</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Supports array LUNs</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>
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 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:

Note: 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:

<table>
<thead>
<tr>
<th>Node</th>
<th>Interface</th>
<th>IP address</th>
<th>Subnet</th>
</tr>
</thead>
<tbody>
<tr>
<td>node_A_1</td>
<td>MetroCluster IP interface 1</td>
<td>10.1.1.1</td>
<td>10.1.1/24</td>
</tr>
<tr>
<td></td>
<td>MetroCluster IP interface 2</td>
<td>10.1.2.1</td>
<td>10.1.2/24</td>
</tr>
<tr>
<td>node_A_2</td>
<td>MetroCluster IP interface 1</td>
<td>10.1.1.2</td>
<td>10.1.1/24</td>
</tr>
<tr>
<td></td>
<td>MetroCluster IP interface 2</td>
<td>10.1.2.2</td>
<td>10.1.2/24</td>
</tr>
<tr>
<td>node_B_1</td>
<td>MetroCluster IP interface 1</td>
<td>10.1.1.3</td>
<td>10.1.1/24</td>
</tr>
<tr>
<td></td>
<td>MetroCluster IP interface 2</td>
<td>10.1.2.3</td>
<td>10.1.2/24</td>
</tr>
<tr>
<td>node_B_2</td>
<td>MetroCluster IP interface 1</td>
<td>10.1.1.4</td>
<td>10.1.1/24</td>
</tr>
<tr>
<td></td>
<td>MetroCluster IP interface 2</td>
<td>10.1.2.4</td>
<td>10.1.2/24</td>
</tr>
</tbody>
</table>

**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 the initial MetroCluster configuration. 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 9.4 and later**

Starting with ONTAP 9.4, MetroCluster IP configurations support new installations with AFF systems using ADP (Advanced Drive Partitioning). In most configurations, partitioning and disk assignment is performed automatically during the initial configuration of the MetroCluster sites.

ONTAP 9.4 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.

**Note:** When upgrading from ONTAP 9.4 to 9.5, the system recognizes the existing disk assignments.

**Automatic partitioning**

ADP is performed automatically during initial configuration of the platform.
Note: Starting with ONTAP 9.5, disk autoassignment must be enabled for automatic partitioning for ADP to occur.

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

Manual drive assignment (ONTAP 9.5)
In ONTAP 9.5, manual drive assignment is required on systems with the following shelf configurations:

- Three external shelves per site.
  Two shelves are assigned automatically using a half-shelf assignment policy, but the third shelf must be assigned manually.
- More than four shelves per site and the total number of external shelves is not a multiple of four.
  Extra shelves above the nearest multiple of four are left unassigned and the drives must be assigned manually. For example, if there are five external shelves at the site, shelf five must be assigned manually.

You only need to manually assign a single drive on each unassigned shelf. The rest of the drives on the shelf are then automatically assigned.

Manual drive assignment (ONTAP 9.4)
In ONTAP 9.4, manual drive assignment is required on systems with the following shelf configurations:

- Fewer than four external shelves per site.
  The drives must be assigned manually to ensure symmetrical assignment of the drives, with each pool having an equal number of drives.
- More than four external shelves per site and the total number of external shelves is not a multiple of four.
  Extra shelves above the nearest multiple of four are left unassigned and the drives must be assigned manually.

When manually assigning drives, you should assign disks symmetrically, with an equal number of drives assigned to each pool. For example, if the configuration has two storage shelves at each site, you would one shelf to the local HA pair and one shelf to the remote HA pair:

- Assign half of the disks on site_A-shelf_1 to pool 0 of node_A_1.
- Assign half of the disks on site_A-shelf_1 to pool 0 of node_A_2.
• Assign half of the disks on site_A-shelf_2 to pool 1 of node_B_1.
• Assign half of the disks on site_A-shelf_2 to pool 1 of node_B_2.
• Assign half of the disks on site_B-shelf_1 to pool 0 of node_B_1.
• Assign half of the disks on site_B-shelf_1 to pool 0 of node_B_2.
• Assign half of the disks on site_B-shelf_2 to pool 1 of node_A_1.
• Assign half of the disks on site_B-shelf_2 to pool 1 of node_A_2.

Adding shelves to an existing configuration.

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

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 concepts

*Required MetroCluster IP components and naming conventions* on page 35

Related information

*Disk and aggregate management*

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.

**Note:** In systems using ADP, aggregates are created using partitions in which each drive is partitioned into 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.

*NetApp Hardware Universe*
## ADP and disk assignment on AFF A220 systems

<table>
<thead>
<tr>
<th>Guideline</th>
<th>Shelves per site</th>
<th>Drive assignment rules</th>
<th>ADP layout for root partition</th>
</tr>
</thead>
</table>
| Minimum recommended shelves (per site) | Internal drives only | 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. **Note:** Half of the internal drives remain unassigned before MetroCluster is configured. | Two quarters are used by the local HA pair. The other two quarters are used by the remote HA pair. The root aggregate includes the following partitions in each plex:  
  - Three partitions for data  
  - Two parity partitions  
  - One spare partition                                                                                                                                                                                     |
| Minimum supported shelves (per site) | 16 internal drives | The drives are divided into four equal groups. Each quarter-shelf is automatically assigned to a separate pool.  
Two quarters on a shelf can have the same pool. The pool is chosen based on the node that owns the quarter:  
  - If owned by the local node, pool0 is used.  
  - If owned by the remote node, pool1 is used.  
For example: a shelf with quarters Q1 through Q4 can have following assignments:  
  - Q1: node_A_1 pool0  
  - Q2: node_A_2 pool0  
  - Q3: node_B_1 pool1  
  - Q4: node_B_2 pool1  
**Note:** Half of the internal drives remain unassigned before MetroCluster is configured. | Each of the two plexes in the root aggregate includes the following partitions:  
  - One partition for data  
  - Two parity partitions  
  - One spare partition                                                                                                                                                                              |
## ADP and disk assignment on AFF A300 systems

<table>
<thead>
<tr>
<th>Guideline</th>
<th>Shelves per site</th>
<th>Drive assignment rules</th>
<th>ADP layout for root partition</th>
</tr>
</thead>
</table>
| 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:  
  • 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-self is automatically assigned to a separate pool. | Each of the two plexes in the root aggregate includes the following partitions:  
  • Three partitions for data  
  • Two parity partitions  
  • One spare partition |

## ADP and disk assignment on AFF A700 systems

<table>
<thead>
<tr>
<th>Guideline</th>
<th>Shelves per site</th>
<th>Drive assignment rules</th>
<th>ADP layout for root partition</th>
</tr>
</thead>
</table>
| Minimum recommended shelves (per site) | Four shelves     | Drives are automatically assigned on a shelf-by-shelf basis.                           | Each of the two plexes in the root aggregate includes:  
  • 20 partitions for data  
  • Two parity partitions  
  • Two spare partitions |
| Minimum supported shelves (per site) | One shelf         | The drives are divided into four equal groups (quarters). Each quarter-self is automatically assigned to a separate pool. | Each of the two plexes in the root aggregate includes:  
  • Three partitions for data  
  • Two parity partitions  
  • One spare partition |
### ADP and disk assignment on AFF A800 systems

<table>
<thead>
<tr>
<th>Guideline</th>
<th>Shelves per site</th>
<th>Drive assignment rules</th>
<th>ADP layout for root aggregate</th>
</tr>
</thead>
</table>
| Minimum recommended shelves (per site) | Internal drives and four external shelves | The internal partitions are divided into four equal groups (quarters). Each quarter is automatically assigned to a separate pool. The drives on the external shelves are automatically assigned on a shelf-by-shelf basis, with all of the drives on each shelf assigned to one of the four nodes in the MetroCluster configuration. | The root aggregate is created with 12 root partitions on the internal shelf. Each of the two plexes in the root aggregate includes:  
  - Eight partitions for data  
  - Two parity partitions  
  - Two spare partitions |
| Minimum supported shelves (per site) | 24 internal drives only | The internal partitions are divided into four equal groups (quarters). Each quarter is automatically assigned to a separate pool. | The root aggregate is created with 12 root partitions on the internal shelf. Each of the two plexes in the root aggregate includes:  
  - Three partitions for data  
  - Two parity partitions  
  - One spare partitions |

### Disk assignment on FAS2750 systems

<table>
<thead>
<tr>
<th>Guideline</th>
<th>Shelves per site</th>
<th>Drive assignment rules</th>
<th>ADP layout for root partition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum recommended shelves (per site)</td>
<td>One internal and one external shelf</td>
<td>The internal and external shelves are divided into two equal halves. Each half is automatically assigned to different pool</td>
<td>Not applicable.</td>
</tr>
<tr>
<td>Minimum supported shelves (per site) (active/passive HA configuration)</td>
<td>Internal drives only</td>
<td>Manual assignment required.</td>
<td></td>
</tr>
</tbody>
</table>
### Disk assignment on FAS8200 systems

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

### Disk assignment on FAS9000 systems

<table>
<thead>
<tr>
<th>Guideline</th>
<th>Shelves per site</th>
<th>Drive assignment rules</th>
<th>ADP layout for root partition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum recommended shelves (per site)</td>
<td>Four shelves</td>
<td>Drives are automatically assigned on a shelf-by-shelf basis.</td>
<td>Not applicable.</td>
</tr>
<tr>
<td>Minimum supported shelves (per site)</td>
<td>Two shelves</td>
<td>The drives on the shelves are divided into two equal groups (halves). Each half-shelf is automatically assigned to a separate pool.</td>
<td></td>
</tr>
<tr>
<td>Minimum supported shelves (per site) (active/passive HA configuration)</td>
<td>One shelf</td>
<td>Manual assignment required.</td>
<td></td>
</tr>
</tbody>
</table>

### 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 express configuration*
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.

Note: ONTAP 9 enables you to migrate your peering networks from IPv4 to IPv6 by optionally allowing both protocols to be present simultaneously on the intercluster LIFs. In earlier releases, all intercluster relationships for an entire cluster were either IPv4 or IPv6. This meant that changing protocols was a potentially disruptive event.

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.
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.

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.
Considerations for sharing private layer 2 networks

Starting with ONTAP 9.6, MetroCluster IP configurations with supported Cisco switches can share existing networks for ISLs, rather than using dedicated MetroCluster ISLs. Earlier ONTAP versions require dedicated ISLs.

The MetroCluster IP switches are dedicated to the MetroCluster configuration and cannot be shared. Only the MetroCluster ISL ports on the MetroCluster IP switches can connect to the shared switches.

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

MetroCluster ISL requirements in shared 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.

- Jitter must not be greater than 3 ms.

- 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.

- 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. AFF A220 and FAS2750 systems require VLAN 10 and VLAN 20. 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.

ISL cabling requirements
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.

<table>
<thead>
<tr>
<th>MetroCluster IP switch model</th>
<th>Port type</th>
<th>Number of ISLs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Broadcom-supported BES-53248 switches</td>
<td>Native ports</td>
<td>4 ISLs using 10 or 25-Gbps ports</td>
</tr>
<tr>
<td>Cisco 3132Q-V</td>
<td>Native ports</td>
<td>6 ISLs using 40 Gbps ports</td>
</tr>
<tr>
<td></td>
<td>Breakout cables</td>
<td>16 x 10 Gbps ISLs</td>
</tr>
<tr>
<td>Cisco 3232C</td>
<td>Native ports</td>
<td>6 ISLs using 40 or 100 Gbps ports</td>
</tr>
<tr>
<td></td>
<td>Breakout cables</td>
<td>16 x 10 Gbps ISLs</td>
</tr>
</tbody>
</table>

- 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. Using a single 10 Gbps ISL per network for AFF A800 systems is strongly discouraged.

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.

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.

Example of CoS policies for Cisco Nexus 3000 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:

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

**Supported and unsupported shared network configuration types**

Starting with ONTAP 9.6, some shared ISL network configurations are supported for MetroCluster IP configurations. You should also be aware of unsupported configurations.

**Supported 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.
Supported 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 NetApp. 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.
Supported configuration with two MetroCluster configurations sharing an intermediate network

In this supported 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:
Unsupported configuration with a MetroCluster configuration connected to another MetroCluster configuration's switches

In this unsupported configuration, the MetroCluster switches of one MetroCluster configuration cascade into the switches of a second configuration before connecting to the intermediate network. In the following example, MetroCluster Two switch_A_1 connects to MetroCluster One switch_A_1.

The example is simplified for illustration purposes only:
Unsupported configuration with two MetroCluster configurations with one connecting directly to the intermediate network

In this unsupported topology, two separate MetroCluster configurations are sharing the same intermediate network, and one MetroCluster configuration's nodes are directly connected to the intermediate switch.

The example is simplified for illustration purposes only:
Unsupported configuration with two MetroCluster configurations with one connecting directly to the intermediate network

In this unsupported topology, two separate MetroCluster configurations are sharing the same intermediate network, and one MetroCluster configuration's nodes are directly connected to the intermediate switch. I

The example is simplified for illustration purposes only:
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.

The Hardware Universe tool provides some notes about the requirements that TDM/xWDM equipment must meet to work with a MetroCluster IP configuration.

*NetApp Hardware Universe*

**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 IMT 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 NetApp 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 on page 18.

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.

<table>
<thead>
<tr>
<th>Step</th>
<th>Completed at factory</th>
<th>Completed by you</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mount components in one or more cabinets.</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Position cabinets in the desired location.</td>
<td>No</td>
<td>Yes Position them in the original order so that the supplied cables are long enough.</td>
</tr>
<tr>
<td>Connect multiple cabinets to each other, if applicable.</td>
<td>No</td>
<td>Yes Use the cabinet interconnect kit if it is included in the order. The kit box is labeled.</td>
</tr>
<tr>
<td>Step</td>
<td>Completed at factory</td>
<td>Completed by you</td>
</tr>
<tr>
<td>----------------------------------------------------------------------</td>
<td>----------------------</td>
<td>------------------</td>
</tr>
<tr>
<td>Secure the cabinets to the floor, if applicable.</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Use the universal bolt-down kit if it is included in the order. The kit box is labeled.</td>
</tr>
<tr>
<td>Cable the components within the cabinet.</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Connect the cables between cabinets, if applicable.</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cables are in the accessories box.</td>
</tr>
<tr>
<td>Connect management cables to the customer's network.</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Connect them directly or through the CN1601 management switches, if present.</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Attention:</strong> 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.</td>
</tr>
<tr>
<td>Connect console ports to the customer's terminal server, if applicable.</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Connect the customer's data cables to the cluster.</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Connect the cabinets to power and power on the components.</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Power them on in the following order:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1. PDUs</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Disk shelves</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. Nodes</td>
</tr>
<tr>
<td>Assign IP addresses to the management ports of the cluster switches and to the management ports of the management switches, if present.</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>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.</td>
</tr>
<tr>
<td>Verify cabling by running the Config Advisor tool.</td>
<td>No</td>
<td>Yes</td>
</tr>
</tbody>
</table>
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:
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.
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. The port usage shown is for an AFF A700 or FAS9000 system.
Related concepts

*Considerations for MetroCluster IP configuration* on page 8

**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.
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 AFF systems, all controller modules in the MetroCluster configuration must be configured as AFF 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. Naming must be unique within the MetroCluster configuration.

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. Naming must be unique within the MetroCluster configuration.

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.

Each controller module must be running the same ONTAP version. Supported platform models depend on the ONTAP version:

- New MetroCluster IP installations on FAS systems are not supported in ONTAP 9.4. Existing MetroCluster IP configurations on FAS systems can be upgraded to ONTAP 9.4.
- Starting with ONTAP 9.5, new MetroCluster IP installations on FAS systems are supported.
- Starting with ONTAP 9.4, controller modules configured for ADP are supported.

Example names:
- Site A: cluster_A
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.

<table>
<thead>
<tr>
<th>Platform model</th>
<th>Required Gigabit Ethernet adapter</th>
<th>Required slot for adapter</th>
</tr>
</thead>
<tbody>
<tr>
<td>AFF A700 and FAS9000</td>
<td>X91146A-C</td>
<td>Slot 5</td>
</tr>
<tr>
<td>AFF A800</td>
<td>X1146A</td>
<td>Slot 1</td>
</tr>
<tr>
<td>AFF A300 and FAS8200</td>
<td>X1116A</td>
<td>Slot 1</td>
</tr>
<tr>
<td>AFF A220 and FAS2750</td>
<td>Onboard ports</td>
<td>Slot 0</td>
</tr>
</tbody>
</table>

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

MetroCluster configurations support RAID-DP and RAID4.

Drive location considerations for half-shelf configurations

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.

Drive location considerations for AFF A800 internal drives

For correct implementation of the ADP feature, the AFF A800 system disk slots must be divided into quarters and the disks must be located symmetrically in the quarters.

An AFF A800 system has 48 drive bays. The bays can be divided into quarters:

- Quarter one:
  - Bays 0 - 5
If this system is populated with 16 drives, they must be symmetrically distributed among the four quarters:

- Four drives in the first quarter: 0, 1, 2, 3
- Four drives in the second quarter: 12, 13, 14, 15
- Four drives in the third quarter: 24, 25, 26, 27
- Four drives in the fourth quarter: 36, 37, 38, 39

Related concepts

*Considerations for automatic drive assignment and ADP systems in ONTAP 9.4 and later* on page 9

## 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.

### Steps

1. **Racking the hardware components** on page 38
2. **Cabling the IP switches** on page 39
3. **Cabling the cluster peering connections** on page 49
4. **Cabling the management and data connections** on page 50
5. **Configuring the IP switches** on page 50

### 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.
Steps

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.
   
   AFF A220/FAS2700 Systems Installation and Setup Instructions
   AFF A800 Systems Installation and Setup Instructions
   AFF A700 and FAS9000 Installation and Setup Instructions
   AFF A300 Systems Installation and Setup Instructions
   FAS8200 Systems Installation and Setup Instructions

4. Install the IP switches in the rack or cabinet.

5. Install the disk shelves, power them on, and then set the shelf IDs.

   NetApp Documentation: Disk Shelves
   • 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).

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.

Step

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

   Port assignments for AFF A800, AFF A700, or FAS9000 systems on page 39
   Port assignments for AFF A300 or FAS8200 systems with Cisco 3232C IP switches on page 41
   Port assignments for AFF A300, FAS8200, AFF A220, and FAS2750 systems with Broadcom
      supported BES-53248 IP switches on page 45
   Port assignments for ISLs on page 47

Port assignments for AFF A800, AFF A700, or FAS9000 systems

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

The following tables show the port usage for site A. The same cabling is used for site B.

Note: The switches cannot be configured with ports of different speeds (for example, a mix of 100
Gbps ports and 40 Gbps ports).
### Switch port usage

#### IP_switch_A_1 local connections

<table>
<thead>
<tr>
<th>Port</th>
<th>Port speed</th>
<th>Controller module port</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>100</td>
<td>40 Gbps node_A_1 e0a e4a</td>
<td>Local cluster interconnect</td>
</tr>
<tr>
<td>2</td>
<td>100</td>
<td>40 Gbps node_A_2 e0a e4a</td>
<td>Local cluster interconnect</td>
</tr>
<tr>
<td>3</td>
<td>-</td>
<td>-</td>
<td>Unused</td>
</tr>
<tr>
<td>4</td>
<td>-</td>
<td>-</td>
<td>Unused</td>
</tr>
<tr>
<td>5</td>
<td>-</td>
<td>-</td>
<td>Unused</td>
</tr>
<tr>
<td>6</td>
<td>-</td>
<td>-</td>
<td>Unused</td>
</tr>
<tr>
<td>9</td>
<td>40/100</td>
<td>40 Gbps node_A_1 e0 (100 Gbps) e5a (40 Gbps)</td>
<td>MetroCluster IP interconnect</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>node_A_2 e0 (100 Gbps) e5a (40 Gbps)</td>
<td>MetroCluster IP interconnect</td>
</tr>
</tbody>
</table>

#### IP_switch_A_2 local connections

<table>
<thead>
<tr>
<th>Port</th>
<th>Port speed</th>
<th>Controller module port</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>100</td>
<td>40 Gbps node_A_1 e1a e4b</td>
<td>Local cluster interconnect</td>
</tr>
<tr>
<td>2</td>
<td>100</td>
<td>-</td>
<td>Local cluster interconnect</td>
</tr>
<tr>
<td>3</td>
<td>-</td>
<td>-</td>
<td>Unused</td>
</tr>
<tr>
<td>4</td>
<td>-</td>
<td>-</td>
<td>Unused</td>
</tr>
<tr>
<td>5</td>
<td>-</td>
<td>-</td>
<td>Unused</td>
</tr>
<tr>
<td>6</td>
<td>-</td>
<td>-</td>
<td>Unused</td>
</tr>
<tr>
<td>9</td>
<td>40/100</td>
<td>40 Gbps node_A_1 e1 (100 Gbps) e5b (40 Gbps)</td>
<td>MetroCluster IP interconnect</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>node_A_2 e1 (100 Gbps) e5b (40 Gbps)</td>
<td>MetroCluster IP interconnect</td>
</tr>
</tbody>
</table>
Port assignments for AFF A300 or FAS8200 systems with Cisco 3232C IP switches

The port usage in a MetroCluster IP configuration depends on the switch model and platform type. The following tables show the port usage for site A. The same cabling is used for site B.

Note: The switches cannot be configured with ports of different speeds (for example, a mix of 100 Gbps ports and 40 Gbps ports).

ISL notes
ISL connections can be done either with the 40/100 Gbps ports (15 - 20) or using breakout cables with ports 21 - 24.

Switch port usage
A separate breakout cable is required for each MetroCluster IP interconnect connection (two breakout cables for each node).

The switch ports that support breakout mode are divided into four logical ports. For example, physical port 25 is logically split into four ports:

- 25/1
- 25/2
- 25/3
- 25/4

<table>
<thead>
<tr>
<th>Switch port</th>
<th>Port speed</th>
<th>Controller module port</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>25/1</td>
<td>100 Gbps switch port is connected to a 25 Gbps port on controller using 4x25 Gbps breakout cable</td>
<td>node_A_1</td>
<td>e1a</td>
</tr>
<tr>
<td>26/1</td>
<td>-</td>
<td>node_A_2</td>
<td>e1a</td>
</tr>
<tr>
<td>27/1</td>
<td>-</td>
<td>-</td>
<td>Unused</td>
</tr>
<tr>
<td>28/1</td>
<td>-</td>
<td>-</td>
<td>Unused</td>
</tr>
<tr>
<td>29/1</td>
<td>100 Gbps switch port is connected to a 10 Gbps port on controller using 4x10 Gbps breakout cable</td>
<td>node_A_1</td>
<td>e0a</td>
</tr>
<tr>
<td>30/1</td>
<td>-</td>
<td>node_A_2</td>
<td>e0a</td>
</tr>
</tbody>
</table>
### IP_switch_A_2 Local interconnect connections

<table>
<thead>
<tr>
<th>Switch port</th>
<th>Port speed</th>
<th>Controller module port</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>25/1</td>
<td>100 Gbps switch port is connected to a 25 Gbps port on controller using 4x25 Gbps breakout cable</td>
<td>node_A_1</td>
<td>e1b</td>
</tr>
<tr>
<td>26/1</td>
<td>-</td>
<td>node_A_2</td>
<td>e1b</td>
</tr>
<tr>
<td>27/1</td>
<td>-</td>
<td>-</td>
<td>Unused</td>
</tr>
<tr>
<td>28/1</td>
<td>-</td>
<td>-</td>
<td>Unused</td>
</tr>
<tr>
<td>29/1</td>
<td>100 Gbps switch port is connected to a 10 Gbps port on controller using 4x10 Gbps breakout cable</td>
<td>node_A_1</td>
<td>e0b</td>
</tr>
<tr>
<td>30/1</td>
<td>-</td>
<td>node_A_2</td>
<td>e0b</td>
</tr>
</tbody>
</table>

### IP_switch_A_1 ISL connections

<table>
<thead>
<tr>
<th>Switch port</th>
<th>Port Speed</th>
<th>Connects to switch</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>100 Gbps</td>
<td>IP_switch_A_2</td>
<td>Local cluster ISL</td>
</tr>
<tr>
<td>8</td>
<td>100 Gbps</td>
<td>IP_switch_A_2</td>
<td>Local cluster ISL</td>
</tr>
<tr>
<td>9 - 14</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>15</td>
<td>40/100 Gbps</td>
<td>IP_switch_B_1</td>
<td>MetroCluster ISL</td>
</tr>
<tr>
<td>16</td>
<td>40/100 Gbps</td>
<td>IP_switch_B_1</td>
<td>MetroCluster ISL</td>
</tr>
<tr>
<td>17</td>
<td>40/100 Gbps</td>
<td>IP_switch_B_1</td>
<td>MetroCluster ISL</td>
</tr>
<tr>
<td>18</td>
<td>40/100 Gbps</td>
<td>IP_switch_B_1</td>
<td>MetroCluster ISL</td>
</tr>
<tr>
<td>19</td>
<td>40/100 Gbps</td>
<td>IP_switch_B_1</td>
<td>MetroCluster ISL</td>
</tr>
<tr>
<td>20</td>
<td>40/100 Gbps</td>
<td>IP_switch_B_1</td>
<td>MetroCluster ISL</td>
</tr>
<tr>
<td>21/1</td>
<td>10 Gbps using 4x10 Gbps breakout cables</td>
<td>IP_switch_B_1</td>
<td>MetroCluster ISL (when using the Cisco 3232C switch in breakout mode).</td>
</tr>
</tbody>
</table>

### IP_switch_A_2 ISL connections

<table>
<thead>
<tr>
<th>Switch port</th>
<th>Port Speed</th>
<th>Connects to switch</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>100 Gbps</td>
<td>IP_switch_A_1</td>
<td>Local cluster ISL</td>
</tr>
</tbody>
</table>
Port assignments for AFF A220 and FAS2750 systems with Cisco 3232C or 3132Q-V IP switches

The port usage in a MetroCluster IP configuration depends on the switch model and platform type. The following tables show the port usage for site A. The same cabling is used for site B.

### Switch port usage

These tables do not show unused ports or the ISL ports.

<table>
<thead>
<tr>
<th>Switch</th>
<th>Port</th>
<th>Port Speed</th>
<th>Connects to switch</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>port</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td>100 Gbps</td>
<td>IP_switch_A_1</td>
<td>Local cluster ISL</td>
</tr>
<tr>
<td>9 - 14</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td></td>
<td>40/100 Gbps</td>
<td>IP_switch_B_2</td>
<td>MetroCluster ISL</td>
</tr>
<tr>
<td>16</td>
<td></td>
<td>40/100 Gbps</td>
<td>IP_switch_B_2</td>
<td>MetroCluster ISL</td>
</tr>
<tr>
<td>17</td>
<td></td>
<td>40/100 Gbps</td>
<td>IP_switch_B_2</td>
<td>MetroCluster ISL</td>
</tr>
<tr>
<td>18</td>
<td></td>
<td>40/100 Gbps</td>
<td>IP_switch_B_2</td>
<td>MetroCluster ISL</td>
</tr>
<tr>
<td>19</td>
<td></td>
<td>40/100 Gbps</td>
<td>IP_switch_B_2</td>
<td>MetroCluster ISL</td>
</tr>
<tr>
<td>20</td>
<td></td>
<td>40/100 Gbps</td>
<td>IP_switch_B_2</td>
<td>MetroCluster ISL</td>
</tr>
<tr>
<td>21/1</td>
<td></td>
<td>10 Gbps using 4x10 Gbps breakout cables</td>
<td>IP_switch_B_2</td>
<td>MetroCluster ISL (when using the Cisco 3232C switch in breakout mode).</td>
</tr>
<tr>
<td>22/1</td>
<td></td>
<td></td>
<td>IP_switch_B_2</td>
<td></td>
</tr>
<tr>
<td>23/1</td>
<td></td>
<td></td>
<td>IP_switch_B_2</td>
<td></td>
</tr>
<tr>
<td>24/1</td>
<td></td>
<td></td>
<td>IP_switch_B_2</td>
<td></td>
</tr>
</tbody>
</table>

---

<table>
<thead>
<tr>
<th>Port</th>
<th>Port speed</th>
<th>Controller module port</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Cisco 3232C</td>
<td>Cisco 3132Q-V</td>
</tr>
<tr>
<td>7</td>
<td>100 Gbps</td>
<td>40 Gbps</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cisco 3232C</td>
<td>Cisco 3132Q-V</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Local cluster interconnect</td>
</tr>
<tr>
<td>8</td>
<td>100 Gbps switch port is connected to a 10 Gbps port on controller using 4x10 Gbps breakout cable</td>
<td>10 Gbps</td>
<td>-</td>
</tr>
</tbody>
</table>
### IP_switch_A_1 local connections

<table>
<thead>
<tr>
<th>Port</th>
<th>Port speed</th>
<th>Controller module port</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Cisco 3232C</td>
<td>Cisco 3132Q-V</td>
</tr>
<tr>
<td>9/1</td>
<td>-</td>
<td>-</td>
<td>node_A_1</td>
</tr>
<tr>
<td>9/2</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>9/3</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>9/4</td>
<td>100 Gbps switch port is connected to a 10 Gbps port on controller using 4x10 Gbps breakout cable</td>
<td>10 Gbps</td>
<td>-</td>
</tr>
<tr>
<td>10/1</td>
<td>-</td>
<td>-</td>
<td>node_A_2</td>
</tr>
<tr>
<td>10/2</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>10/3</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>10/4</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

### IP_switch_A_2 local connections

<table>
<thead>
<tr>
<th>Port</th>
<th>Port speed</th>
<th>Controller module port</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Cisco 3232C</td>
<td>Cisco 3132Q-V</td>
</tr>
<tr>
<td>7</td>
<td>100 Gbps</td>
<td>40 Gbps</td>
<td>-</td>
</tr>
<tr>
<td>8</td>
<td>100 Gbps switch port is connected to a 10 Gbps port on controller using 4x10 Gbps breakout cable</td>
<td>10 Gbps</td>
<td>-</td>
</tr>
</tbody>
</table>
### IP_switch_A_2 local connections

<table>
<thead>
<tr>
<th>Port</th>
<th>Port speed</th>
<th>Controller module port</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cisco 3232C</td>
<td>Cisco 3132Q-V</td>
<td>100 Gbps</td>
</tr>
<tr>
<td>9/1</td>
<td>-</td>
<td>-</td>
<td>node_A_1</td>
</tr>
<tr>
<td>9/2</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>9/3</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>9/4</td>
<td>100 Gbps switch port is connected to a 10 Gbps port on controller using 4x10 Gbps breakout cable</td>
<td>10 Gbps</td>
<td>-</td>
</tr>
<tr>
<td>10/1</td>
<td>-</td>
<td>-</td>
<td>node_A_2</td>
</tr>
<tr>
<td>10/2</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>10/3</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>10/4</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

### Port assignments for AFF A300, FAS8200, AFF A220, and FAS2750 systems with Broadcom supported BES-53248 IP switches

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

The following tables show the port usage for site A. The same cabling is used for site B.

**Note:** The switches cannot be configured with ports of different speeds (for example, a mix of 25 Gbps ports and 10 Gbps ports).

### Switch port usage

<table>
<thead>
<tr>
<th>Port</th>
<th>Port speed</th>
<th>Controller module port</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Node</td>
<td>AFF A220 and FAS2750</td>
<td>AFF A300 and FAS8200 systems</td>
</tr>
<tr>
<td>1</td>
<td>node_A_1</td>
<td>-</td>
<td>e0a</td>
</tr>
<tr>
<td>2</td>
<td>node_A_2</td>
<td>-</td>
<td>e0a</td>
</tr>
<tr>
<td>3</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
### IP_switch_A_1 local connections

<table>
<thead>
<tr>
<th>Port</th>
<th>Port speed</th>
<th>Controller module port</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>-</td>
<td>-</td>
<td>Local cluster interconnect</td>
</tr>
<tr>
<td>5</td>
<td>25 Gbps</td>
<td>node_A_1 - e1a</td>
<td>MetroCluster IP interconnect</td>
</tr>
<tr>
<td>6</td>
<td>25 Gbps</td>
<td>node_A_2 - e1a</td>
<td>MetroCluster IP interconnect</td>
</tr>
<tr>
<td>7</td>
<td>-</td>
<td>-</td>
<td>MetroCluster IP interconnect</td>
</tr>
<tr>
<td>8</td>
<td>-</td>
<td>-</td>
<td>MetroCluster IP interconnect</td>
</tr>
<tr>
<td>9</td>
<td>10 Gbps</td>
<td>node_A_1 e0a</td>
<td>Shared cluster / MetroCluster IP interconnect</td>
</tr>
<tr>
<td>10</td>
<td>10 Gbps</td>
<td>node_A_2 e0a</td>
<td>Shared cluster / MetroCluster IP interconnect</td>
</tr>
<tr>
<td>11</td>
<td>-</td>
<td>-</td>
<td>Shared cluster / MetroCluster IP interconnect</td>
</tr>
<tr>
<td>12</td>
<td>-</td>
<td>-</td>
<td>Shared cluster / MetroCluster IP interconnect</td>
</tr>
<tr>
<td>55</td>
<td>100 Gbps</td>
<td>IP_switch_B_1 -</td>
<td>Local interconnect</td>
</tr>
<tr>
<td>56</td>
<td>100 Gbps</td>
<td>IP_switch_B_1 -</td>
<td>Local interconnect</td>
</tr>
</tbody>
</table>

### IP_switch_A_2 local connections

<table>
<thead>
<tr>
<th>Port</th>
<th>Port speed</th>
<th>Controller module port</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>10 Gbps</td>
<td>node_A_1 - e0b</td>
<td>Local cluster interconnect</td>
</tr>
<tr>
<td>2</td>
<td>10 Gbps</td>
<td>node_A_2 - e0b</td>
<td>Local cluster interconnect</td>
</tr>
<tr>
<td>3</td>
<td>-</td>
<td>-</td>
<td>Local cluster interconnect</td>
</tr>
<tr>
<td>4</td>
<td>-</td>
<td>-</td>
<td>Local cluster interconnect</td>
</tr>
</tbody>
</table>
### IP_switch_A_2 local connections

<table>
<thead>
<tr>
<th>Port</th>
<th>Port speed</th>
<th>Node</th>
<th>AFF A220 and FAS2750</th>
<th>AFF A300 and FAS8200 systems</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>25 Gbps</td>
<td>node_A_1</td>
<td>-</td>
<td>e1b</td>
<td>MetroCluster IP interconnect</td>
</tr>
<tr>
<td>6</td>
<td>25 Gbps</td>
<td>node_A_2</td>
<td>-</td>
<td>e1b</td>
<td>MetroCluster IP interconnect</td>
</tr>
<tr>
<td>7</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>MetroCluster IP interconnect</td>
</tr>
<tr>
<td>8</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>MetroCluster IP interconnect</td>
</tr>
<tr>
<td>9</td>
<td>10 Gbps</td>
<td>node_A_1</td>
<td>e0b</td>
<td>-</td>
<td>Shared cluster / MetroCluster IP interconnect</td>
</tr>
<tr>
<td>10</td>
<td>10 Gbps</td>
<td>node_A_2</td>
<td>e0b</td>
<td>-</td>
<td>Shared cluster / MetroCluster IP interconnect</td>
</tr>
<tr>
<td>11</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Shared cluster / MetroCluster IP interconnect</td>
</tr>
<tr>
<td>12</td>
<td>-</td>
<td>--</td>
<td>-</td>
<td>-</td>
<td>Shared cluster / MetroCluster IP interconnect</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>55</td>
<td>100 Gbps</td>
<td>IP_switch _B_2</td>
<td>-</td>
<td>-</td>
<td>Local interconnect</td>
</tr>
<tr>
<td>56</td>
<td>100 Gbps</td>
<td>IP_switch _B_2</td>
<td>-</td>
<td>-</td>
<td>Local interconnect</td>
</tr>
</tbody>
</table>

### Port assignments for ISLs

The ISL port usage in a MetroCluster IP configuration is the same for all switch models.

### ISL usage notes for Cisco 3232C switches

ISL connections can be done either with the 40/100 Gbps ports (15 - 20) or using breakout cables with ports 21 - 24. A separate breakout cable is required for each ISL connection.

The switch ports that support breakout mode are divided into four logical ports. For example, physical port 21 is logically split into four ports:

- 21/1
- 21/2
- 21/3
- 21/4
ISL usage notes for Cisco 3132Q-V switches

One, two or three 40 Gbps ISLs are supported or up to six 10 Gbps MetroCluster ISLs.

If you are using breakout cables for the ISLs, the correct RCF file must be used. A separate breakout cable is required for each ISL connection.

The switch ports that support breakout mode are divided into four logical ports. For example, physical port 21 is logically split into four ports:

- 21/1
- 21/2
- 21/3
- 21/4

ISL port usage

<table>
<thead>
<tr>
<th>Switch port</th>
<th>Port speed for Cisco 3232C switches</th>
<th>Port speed for Cisco 3132Q-V switches</th>
<th>Connects to switch</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>100 Gbps</td>
<td>40 Gbps</td>
<td>IP_switch_A_2</td>
<td>Local cluster ISL</td>
</tr>
<tr>
<td>8</td>
<td>100 Gbps</td>
<td>40 Gbps</td>
<td>IP_switch_A_2</td>
<td>Local cluster ISL</td>
</tr>
<tr>
<td>9 - 14</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>15</td>
<td>40/100 Gbps</td>
<td>40 Gbps</td>
<td>IP_switch_B_1</td>
<td>MetroCluster ISL</td>
</tr>
<tr>
<td>16</td>
<td>40/100 Gbps</td>
<td>40 Gbps</td>
<td>IP_switch_B_1</td>
<td>MetroCluster ISL</td>
</tr>
<tr>
<td>17</td>
<td>40/100 Gbps</td>
<td>40 Gbps</td>
<td>IP_switch_B_1</td>
<td>MetroCluster ISL</td>
</tr>
<tr>
<td>18</td>
<td>40/100 Gbps</td>
<td>40 Gbps</td>
<td>IP_switch_B_1</td>
<td>MetroCluster ISL</td>
</tr>
<tr>
<td>19</td>
<td>40/100 Gbps</td>
<td>40 Gbps</td>
<td>IP_switch_B_1</td>
<td>MetroCluster ISL</td>
</tr>
<tr>
<td>20</td>
<td>40/100 Gbps</td>
<td>40 Gbps</td>
<td>IP_switch_B_1</td>
<td>MetroCluster ISL</td>
</tr>
<tr>
<td>21/1</td>
<td>10 Gbps using 4x10 Gbps breakout cables</td>
<td>10 Gbps using 4x10 Gbps breakout cables</td>
<td>IP_switch_B_1</td>
<td>MetroCluster ISL (switch in breakout mode).</td>
</tr>
<tr>
<td>22/1</td>
<td>IP_switch_B_1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23/1</td>
<td>IP_switch_B_1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24/1</td>
<td>IP_switch_B_1</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Switch port</th>
<th>Port Speed</th>
<th>Port speed</th>
<th>Connects to switch</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>100 Gbps</td>
<td>40 Gbps</td>
<td>IP_switch_A_1</td>
<td>Local cluster ISL</td>
</tr>
<tr>
<td>8</td>
<td>100 Gbps</td>
<td>40 Gbps</td>
<td>IP_switch_A_1</td>
<td>Local cluster ISL</td>
</tr>
<tr>
<td>9 - 14</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>15</td>
<td>40/100 Gbps</td>
<td>40 Gbps</td>
<td>IP_switch_B_2</td>
<td>MetroCluster ISL</td>
</tr>
<tr>
<td>16</td>
<td>40/100 Gbps</td>
<td>40 Gbps</td>
<td>IP_switch_B_2</td>
<td>MetroCluster ISL</td>
</tr>
<tr>
<td>17</td>
<td>40/100 Gbps</td>
<td>40 Gbps</td>
<td>IP_switch_B_2</td>
<td>MetroCluster ISL</td>
</tr>
</tbody>
</table>
### IP_switch_A_2 ISL connections

<table>
<thead>
<tr>
<th>Switch port</th>
<th>Port Speed</th>
<th>Port speed</th>
<th>Connects to switch</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>18</td>
<td>40/100 Gbps</td>
<td>40 Gbps</td>
<td>IP_switch_B_2</td>
<td>MetroCluster ISL</td>
</tr>
<tr>
<td>19</td>
<td>40/100 Gbps</td>
<td>40 Gbps</td>
<td>IP_switch_B_2</td>
<td>MetroCluster ISL</td>
</tr>
<tr>
<td>20</td>
<td>40/100 Gbps</td>
<td>40 Gbps</td>
<td>IP_switch_B_2</td>
<td>MetroCluster ISL</td>
</tr>
<tr>
<td>21/1</td>
<td>10 Gbps using 4x10 Gbps breakout cables</td>
<td>10 Gbps using 4x10 Gbps breakout cables</td>
<td>IP_switch_B_2</td>
<td>MetroCluster ISL (switch in breakout mode).</td>
</tr>
<tr>
<td>22/1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23/1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24/1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### ISL port usage for Broadcom supported BES-53248 IP switches

#### IP_switch_A_1 ISL connections

<table>
<thead>
<tr>
<th>Switch port</th>
<th>Port speed</th>
<th>Connects to switch</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>13</td>
<td>10 / 25 Gbps</td>
<td>IP_switch_B_1</td>
<td>MetroCluster ISL</td>
</tr>
<tr>
<td>14</td>
<td>10 / 25 Gbps</td>
<td>IP_switch_B_1</td>
<td>MetroCluster ISL</td>
</tr>
<tr>
<td>15</td>
<td>10 / 25 Gbps</td>
<td>IP_switch_B_1</td>
<td>MetroCluster ISL</td>
</tr>
<tr>
<td>16</td>
<td>10 / 25 Gbps</td>
<td>IP_switch_B_1</td>
<td>MetroCluster ISL</td>
</tr>
<tr>
<td>55</td>
<td>10 Gbps</td>
<td>IP_switch_A_2</td>
<td>Local cluster ISL</td>
</tr>
<tr>
<td>56</td>
<td>10 Gbps</td>
<td>IP_switch_A_2</td>
<td>Local cluster ISL</td>
</tr>
</tbody>
</table>

#### IP_switch_A_2 ISL connections

<table>
<thead>
<tr>
<th>Switch port</th>
<th>Port speed</th>
<th>Connects to switch</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>13</td>
<td>10 / 25 Gbps</td>
<td>IP_switch_B_2</td>
<td>MetroCluster ISL</td>
</tr>
<tr>
<td>14</td>
<td>10 / 25 Gbps</td>
<td>IP_switch_B_2</td>
<td>MetroCluster ISL</td>
</tr>
<tr>
<td>15</td>
<td>10 / 25 Gbps</td>
<td>IP_switch_B_2</td>
<td>MetroCluster ISL</td>
</tr>
<tr>
<td>16</td>
<td>10 / 25 Gbps</td>
<td>IP_switch_B_2</td>
<td>MetroCluster ISL</td>
</tr>
<tr>
<td>55</td>
<td>10 Gbps</td>
<td>IP_switch_A_1</td>
<td>Local cluster ISL</td>
</tr>
<tr>
<td>56</td>
<td>10 Gbps</td>
<td>IP_switch_A_1</td>
<td>Local cluster ISL</td>
</tr>
</tbody>
</table>

### 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.

**Step**

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 express configuration*

**Related concepts**

*Considerations for configuring cluster peering* on page 15

**Related information**

*Cluster and SVM peering express configuration*

### 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 NetApp CN1601 cluster management switches.

**Step**

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

   *AFF A220/FAS2700 Systems Installation and Setup Instructions*
   *AFF A800 Systems Installation and Setup Instructions*
   *AFF A700 and FAS9000 Installation and Setup Instructions*
   *AFF A300 Systems Installation and Setup Instructions*
   *FAS8200 Systems Installation and Setup Instructions*

### 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.

**Choices**

- Configuring Broadcom IP switches on page 50
- Configuring Cisco IP switches on page 59

### Configuring Broadcom IP switches

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

1. **Resetting the Broadcom IP switch to factory defaults** on page 51
2. **Downloading and installing the Broadcom switch EFOS software and RCF files** on page 54
3. **Downloading and installing the Broadcom RCF files** on page 55
4. **Applying the Broadcom switch RCF files** on page 58

**Resetting the Broadcom IP switch to factory defaults**

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

**About this task**

- You must repeat these steps on each of the IP switches in the MetroCluster IP configuration.
- You must be connected to the switch via the serial console.
- This task resets the configuration of the management network.

**Steps**

1. Change to the elevated command prompt (#):
   ```
   enable
   ```
   **Example**
   ```
   (Routing)> enable
   (Routing) #
   ```

2. Erase the startup configuration:
   ```
   erase startup-config
   ```
   **Example**
   ```
   (Routing) #erase startup-config
   Are you sure you want to clear the configuration? (y/n) y
   (Routing) #
   ```
   This command does not erase the banner.

3. Reboot the switch:
   ```
   reload
   ```
   **Example**
   ```
   (IP_switch_A_1) #reload
   Are you sure you would like to reset the system? (y/n) y
   ```

4. Wait for the switch to reload, and then log in to the switch.
   The default user is “admin”, and no password is set. A prompt similar to the following is displayed:
5. Change to the elevated command prompt:
   `enable`

Example

```
(Routing)> enable
(Routing) #
```

6. Set the serviceport protocol to `none`:
   `serviceport protocol none`

Example

```
(Routing) #serviceport protocol none
Changing protocol mode will reset ip configuration.
Are you sure you want to continue? (y/n) y
(Routing) #
```

7. Assign the IP address to the service port:
   `serviceport ip ip-address netmask gateway`

Example

The following example shows a service port assigned IP address 10.10.10.10 with subnet 255.255.255.0 and gateway 10.10.10.1:

```
(Routing) #serviceport ip 10.10.10.10 255.255.255.0 10.10.10.1
```

8. Verify that the service port is correctly configured:
   `show serviceport`

Example

The following example shows that the port is up and the correct addresses have been assigned:

```
(Routing) #show serviceport
Interface Status............................... Up
IP Address..................................... 10.10.10.10
Subnet Mask.................................... 255.255.255.0
Default Gateway............................... 10.10.10.1
IPv6 Administrative Mode....................... Enabled
IPv6 Prefix is ................................ fe80::dac4:97ff:fe56:87d7/64
IPv6 Default Router............................ fe80::222:bdff:fef8:19ff
Configured IPv4 Protocol....................... None
Configured IPv6 Protocol....................... None
IPv6 AutoConfig Mode.......................... Disabled
Burned In MAC Address......................... D8:C4:97:56:87:D7
(Routing) #
```

9. If desired, configure the domain and name server:
**configure**

**Example**

The following example shows the `ip domain` and `ip name server` commands:

```
(Routing) # configure
(Routing) (Config)#ip domain name lab.netapp.com
(Routing) (Config)#ip name server 10.99.99.1 10.99.99.2
(Routing) (Config)#exit
(Routing) (Config)#
```

10. If desired, configure the time zone and time synchronization (SNTP).

**Example**

The following example shows the `sntp` commands, specifying the IP address of the SNTP server and the relative timezone.

```
(Routing) #
(Routing) (Config)#sntp client mode unicast
(Routing) (Config)#sntp server 10.99.99.5
(Routing) (Config)#clock timezone -7
(Routing) (Config)#exit
(Routing) (Config)#
```

11. Configure the switch name:

```
hostname IP_switch_A_1
```

**Example**

The switch prompt will display the new name:

```
(Routing) # hostname IP_switch_A_1
(IP_switch_A_1) #
```

12. Save the configuration:

```
write memory
```

**Example**

You receive prompts and output similar to the following example:

```
(IP_switch_A_1) #write memory
This operation may take a few minutes.
Management interfaces will not be available during this time.
Are you sure you want to save? (y/n) y
Config file 'startup-config' created successfully .
Configuration Saved!
(IP_switch_A_1) #
```

13. Repeat the previous steps on the other three switches in the MetroCluster IP configuration.
Downloading and installing the Broadcom switch EFOS software and RCF files

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

About this task

This task must be repeated on each switch in the MetroCluster IP configuration.

Steps

1. Copy the switch software to the switch:

   copy sftp://user@50.50.50.50/switchsoftware/efos-3.4.3.1.stk backup

   Example

   In this example, the efos-3.4.3.1.stk operating system file is copied from the SFTP server at 50.50.50.50 to the backup partition. You need to 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://user@50.50.50.50/switchsoftware/efos-3.4.3.1.stk backup
   Remote Password:*************
   Mode........................................... SFTP
   Set Server IP.................................. 50.50.50.50
   Path........................................... /switchsoftware/
   Filename....................................... efos-3.4.3.1.stk
   Data Type...................................... Code
   Destination Filename........................... backup
   Management access will be blocked for the duration of the transfer Are you sure you want to start? (y/n) y
   File transfer in progress. Management access will be blocked for the duration of the transfer. Please wait...
   SFTP Code transfer starting...
   File transfer operation completed successfully.
   (IP_switch_A_1) #

2. Set the switch to boot from the backup partition on the next switch reboot:

   boot system backup

   Example

   (IP_switch_A_1) #boot system backup
   Activating image backup ..
   (IP_switch_A_1) #

3. Verify that the new boot image will be active on the next boot:

   show bootvar

   Example

   (IP_switch_A_1) #show bootvar
   Image Descriptions
### Images currently available on Flash

<table>
<thead>
<tr>
<th>Unit</th>
<th>Active</th>
<th>Backup</th>
<th>Current-active</th>
<th>Next-active</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3.4.3.0</td>
<td>3.4.3.1</td>
<td>3.4.3.0</td>
<td>3.4.3.1</td>
</tr>
</tbody>
</table>

4. Save the configuration:

   ```
   write memory
   ```

   **Example**
   ```
   (IP_switch_A_1) #write memory
   ```
   This operation may take a few minutes. Management interfaces will not be available during this time.
   Are you sure you want to save? (y/n) y
   Configuration Saved!
   ```
   (IP_switch_A_1) #
   ```

5. Reboot the switch:

   ```
   reload
   ```

   **Example**
   ```
   (IP_switch_A_1) #reload
   ```
   Are you sure you would like to reset the system? (y/n) y

6. Wait for the switch to reboot.

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

### Downloading and installing the Broadcom RCF files

You must download and install the switch 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.

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.
Steps

1. Download the MetroCluster IP RCF files for the Broadcom switch.

   Broadcom Cluster and Management Network Switch Reference Configuration File Download for MetroCluster IP

2. Copy the RCF files to the switches:

   a. Copy the RCF files to the first switch:

      ```
      copy sftp://user@FTP-server-IP-address/tftpboot/switch-specific-RCF
      nvram:script BES-53248_RCF_v1.0_Switch-A1.scr
      ```

   Example

   In this example, the BES-53248_RCF_v1.0_Switch-A1.txt RCF file is copied from the SFTP server at 50.50.50.50 to the local bootflash. You need to 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://user@50.50.50.50/RcfFiles/
   BES-53248_RCF_v1.0_IP_switch_A_1.txt nvram:script BES-53248_RCF_v1.0_IP_switch_A_1.scr
   Remote Password:*************
   ```

   ```
   Mode........................................... SFTP
   Set Server IP.................................. 50.50.50.50
   Path........................................... /RcfFiles/
   Filename....................................... BES-53248_RCF_v1.0_IP_switch_A_1.txt
   Data Type...................................... Config Script
   Destination Filename........................... BES-53248_RCF_v1.0_IP_switch_A_1.scr
   ```

   Management access will be blocked for the duration of the transfer
   Are you sure you want to start? (y/n) y

   File transfer in progress. Management access will be blocked for the duration of the transfer. Please wait...

   File transfer operation completed successfully.

   Validating configuration script...
   ```
   config
   set clibanner
   `* NetApp Reference Configuration File (RCF)`
   `*`
   `* Switch : BES-53248`
   ...
   The downloaded RCF is validated. Some output is being logged here.
   ...
   Configuration script validated.
   File transfer operation completed successfully.
   ```
b. Verify that the RCF file has been saved as a script:

```
script list
```

**Example**

```
(IP_switch_A_1) #script list
Configuration Script Name        Size(Bytes)  Date of Modification
-------------------------------  -----------  --------------------
BES-53248_RCF_v1.0_IP_switch_A_1.scr             852   2019 01 29 18:41:25
1 configuration script(s) found.
2046 Kbytes free.
(IP_switch_A_1) #
```

c. Apply the RCF script:

```
script apply BES-53248_RCF_v1.0_IP_switch_A_1.scr
```

**Example**

```
(IP_switch_A_1) #script apply BES-53248_RCF_v1.0_IP_switch_A_1.scr
Are you sure you want to apply the configuration script? (y/n) y

config
set clibanner
"******************************************************************
**************
* NetApp Reference Configuration File (RCF)
*
* Switch    : BES-53248
...
The downloaded RCF is validated. Some output is being logged here.
...
Configuration script 'BES-53248_RCF_v1.0_IP_switch_A_1.scr' applied.
```

```
(IP_switch_A_1) #
```

d. Save the configuration:

```
write memory
```

**Example**

```
(IP_switch_A_1) #write memory
This operation may take a few minutes.
Management interfaces will not be available during this time.
Are you sure you want to save? (y/n) y
```

```
```
Configuration Saved!

(IP_switch_A_1) #

e. Reboot the switch:

   `reload`

**Example**

(IP_switch_A_1) #reload  
Are you sure you would like to reset the system? (y/n) y

f. Repeat the previous steps for each of the other three switches, being sure to copy the matching RCF file to the corresponding switch.

### Applying the Broadcom switch RCF files

You apply the RCF files to the IP switches to configure them correctly for operation in the MetroCluster IP configuration.

#### About this task

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

There are four RCF files, one for each of the four switches in the MetroCluster IP configuration. You must apply the correct RCF file to the matching switch.

<table>
<thead>
<tr>
<th>Switch</th>
<th>RCF file</th>
</tr>
</thead>
<tbody>
<tr>
<td>IP_switch_A_1</td>
<td><code>switch-model_RCF_v1.2-MetroCluster-IP-switch-A-1.txt</code></td>
</tr>
<tr>
<td>IP_switch_B_1</td>
<td><code>switch-model_RCF_v1.2-MetroCluster-IP-switch-B-1.txt</code></td>
</tr>
<tr>
<td>IP_switch_A_2</td>
<td><code>switch-model_RCF_v1.2-MetroCluster-IP-switch-A-2.txt</code></td>
</tr>
<tr>
<td>IP_switch_B_2</td>
<td><code>switch-model_RCF_v1.2-MetroCluster-IP-switch-B-2.txt</code></td>
</tr>
</tbody>
</table>

#### Steps

1. Copy the matching RCF file from the local bootflash to the running configuration on each switch:

   `copy bootflash:switch-specific-RCF.txt running-config`

2. 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:BES53248_RCF_v1.2-MetroCluster-Switch-A-1.txt running-config  
   IP_switch_A-1 # copy running-config startup-config

3. Reload the switch:

   `reload`
Example

```
IP_switch_A_1# reload
```

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

**Configuring Cisco IP switches**

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

**Steps**

1. Resetting the Cisco IP switch to factory defaults on page 59
2. Copying the Cisco switch NX-OS software and RCF files to the MetroCluster IP switches on page 61
3. Installing the Cisco IP switch software on page 63
4. Applying the Cisco switch RCF files on page 65
5. Setting Forward Error Correction for AFF A300 and FAS8200 systems using 25-Gbps connectivity on page 66

**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.

**Steps**

1. Reset the switch to factory defaults:
   a. Erase the existing configuration:
      ```
      write erase
      ```
   b. Reload the switch software:
      ```
      reload
      ```
      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.
   c. 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.
   d. When prompted, enter the user name and password to log in to the switch.
Example

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.

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]:

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]:


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

User Access Verification
IP_switch_A_1 login: admin
Password:
Cisco Nexus Operating System (NX-OS) Software
2. Wait for the switch to reload, and then log in to the switch.

   The switch software has been installed.

   **Example**

   After the switch reboots, 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#
   ```

3. Verify that the switch software has been installed:

   `show version`

   **Example**

   ```
   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
   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#
   ```

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

**Copying the Cisco switch NX-OS software and RCF files to the MetroCluster IP switches**

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

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

You must use the supported switch software version.

NetApp Hardware Universe

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.

<table>
<thead>
<tr>
<th>Switch</th>
<th>RCF file</th>
</tr>
</thead>
<tbody>
<tr>
<td>IP_switch_A_1</td>
<td>switch-model_RCF_v1.2-MetroCluster-IP-switch-A-1.txt</td>
</tr>
<tr>
<td>IP_switch_B_1</td>
<td>switch-model_RCF_v1.2-MetroCluster-IP-switch-B-1.txt</td>
</tr>
<tr>
<td>IP_switch_A_2</td>
<td>switch-model_RCF_v1.2-MetroCluster-IP-switch-A-2.txt</td>
</tr>
<tr>
<td>IP_switch_B_2</td>
<td>switch-model_RCF_v1.2-MetroCluster-IP-switch-B-2.txt</td>
</tr>
</tbody>
</table>

Steps

1. Download the supported NX-OS software file.
   NetApp Downloads: Cisco Ethernet Switch

2. Download the MetroCluster IP RCF files:ntap-dl-rcf-metro-ip
   NetApp Downloads: MetroCluster IP Switch Configuration Files

3. Copy the switch software to the switch:

   ```
   copy sftp://root@server-ip-address/tftpboot/NX-OS-file-name bootflash: vrf management
   ```

   **Example**

   In this example, the `nxos.7.0.3.14.6.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.7.0.3.14.6.bin bootflash: vrf management
   root@10.10.99.99's password: password
   sftp> progress
   Progress meter enabled
   sftp> get /tftpboot/nxos.7.0.3.14.6.bin /bootflash/nxos.7.0.3.14.6.bin
   Fetching /tftpboot/nxos.7.0.3.14.6.bin to /bootflash/nxos.7.0.3.14.6.bin
   /tftpboot/nxos.7.0.3.14.6.bin                 100% 666MB 7.2MB/s 01:32
   sftp> exit
   Copy complete, now saving to disk (please wait)...```

4. Copy the RCF files to the switches:

   a. Copy the RCF files to the first switch:

   ```
   copy sftp://root@FTP-server-IP-address/tftpboot/switch-specific-RCF bootflash: vrf management
   ```

   **Example**

   In this example, the `NX3132_RCF_v1.2-MetroCluster-IP-IP_switch_A_1.txt` RCF file is copied from the SFTP server at 10.10.99.99 to the local bootflash. You need to 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/NX3132_RCF_v1.2-MetroCluster-IP-switch_A_1.txt bootflash: vrf management
root@10.10.99.99's password: password
sftp> progress
Progress meter enabled
sftp> get /tftpboot/NX3132_RCF_v1.2-MetroCluster-IP-switch-A-1.txt /bootflash/
NX3132_RCF_v1.2-MetroCluster-IP-switch-A-1.txt
Fetching /tftpboot/NX3132_RCF_v1.2-MetroCluster-IP-switch-A-1.txt to /bootflash/
NX3132_RCF_v1.2-MetroCluster-IP-switch-A-1.txt
100% 5141 5.0KB/s 00:00
sftp> exit
Copy complete, now saving to disk (please wait)...
IP_switch_A_1#
```

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

5. Verify on each switch that the RCF and switch NX-OS files are present in each switch’s bootflash directory:

```
dir bootflash:

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

**Installing the Cisco IP switch software**

You must install the supported version of the switch NX-OS operating system.

**About this task**

This task must be repeated on each switch in the MetroCluster configuration.

**Steps**

1. 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.

**Example**

The following example shows the software installation on IP_switch_A_1:
Installer will perform compatibility check first. Please wait.

Verifying image bootflash:/nxos.7.0.3.I4.6.bin for boot variable "nxos".

<table>
<thead>
<tr>
<th>Module</th>
<th>bootable</th>
<th>Impact</th>
<th>Install-type</th>
<th>Reason</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>yes</td>
<td>disruptive</td>
<td>reset</td>
<td>default upgrade is not hitless</td>
</tr>
</tbody>
</table>

Images will be upgraded according to following table:

<table>
<thead>
<tr>
<th>Module</th>
<th>Image</th>
<th>Running-Version (pri:alt)</th>
<th>New-Version</th>
<th>Upg-Required</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>nxos</td>
<td>7.0(3)I4(1)</td>
<td>7.0(3)I4(6)</td>
<td>yes</td>
</tr>
<tr>
<td>1</td>
<td>bios</td>
<td>v04.24(04/21/2016)</td>
<td>v04.24(04/21/2016)</td>
<td>no</td>
</tr>
</tbody>
</table>

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

Install is in progress, please wait.

Switch will be reloaded for disruptive upgrade.

2. 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#
```

3. Verify that the switch software has been installed:

```
show version
```
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.7.0.3.14.6.bin
NXOS compile time: 3/9/2017 22:08:00 [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#
```

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

**Applying the Cisco switch RCF files**

You apply the RCF files to the IP switches to configure them correctly for operation in the MetroCluster IP configuration.

**About this task**

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

There are four RCF files, one for each of the four switches in the MetroCluster IP configuration. You must apply the correct RCF file to the matching switch.

<table>
<thead>
<tr>
<th>Switch</th>
<th>RCF file</th>
</tr>
</thead>
<tbody>
<tr>
<td>IP_switch_A_1</td>
<td>switch-model_RCF_v1.2-MetroCluster-IP-switch-A-1.txt</td>
</tr>
<tr>
<td>IP_switch_B_1</td>
<td>switch-model_RCF_v1.2-MetroCluster-IP-switch-B-1.txt</td>
</tr>
<tr>
<td>IP_switch_A_2</td>
<td>switch-model_RCF_v1.2-MetroCluster-IP-switch-A-2.txt</td>
</tr>
<tr>
<td>IP_switch_B_2</td>
<td>switch-model_RCF_v1.2-MetroCluster-IP-switch-B-2.txt</td>
</tr>
</tbody>
</table>

**Steps**

1. Copy the matching RCF file from the local bootflash to the running configuration on each switch:

   `copy bootflash:switch-specific-RCF.txt running-config`
2. Copy the RCF files from the running configuration to the startup configuration on each switch:

```bash
copy running-config startup-config
```

You should see output similar to the following:

```bash
IP_switch_A_1# copy bootflash:NX3132_RCF_v1.2-MetroCluster-Switch-A-1.txt running-config
IP_switch-A-1# copy running-config startup-config
```

3. Reload the switch:

```bash
reload
```

**Example**

```bash
IP_switch_A_1# reload
```

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

### Setting Forward Error Correction for AFF A300 and FAS8200 systems using 25-Gbps connectivity

If your AFF A300 or FAS8200 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.

*Port assignments for AFF A300 or FAS8200 systems with Cisco 3232C IP switches* on page 41

**About this task**

This task only applies to AFF A300 and FAS8200 platforms using 25-Gbps connectivity.

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

**Steps**

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:
   
   a. Enter configuration mode:
      ```bash
      config t
      ```
   
   b. Specify the 25-Gbps interface to configure:
      ```bash
      interface interface-ID
      ```
   
   c. Set fec to off:
      ```bash
      fec off
      ```
   
   d. Repeat the previous steps for each 25-Gbps port on the switch.
   
   e. Exit configuration mode:
      ```bash
      exit
      ```

**Example**

The following example shows the commands for interface Ethernet1/25/1 on switch IP_switch_A_1:
2. Repeat the previous step on the other three switches in the MetroCluster IP configuration.
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.

Steps

1. Gathering required information on page 69
2. Similarities and differences between standard cluster and MetroCluster configurations on page 73
3. Restoring system defaults on a previously used controller module on page 74
4. Verifying the ha-config state of components on page 75
5. Manually assigning drives to pool 0 on page 75
6. Setting up ONTAP on page 79
7. Configuring the clusters into a MetroCluster configuration on page 83
8. Verifying switchover, healing, and switchback on page 110
9. Installing the MetroCluster Tiebreaker software on page 110
10. Protecting configuration backup files on page 111
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.

<table>
<thead>
<tr>
<th>Cluster switch</th>
<th>Host name</th>
<th>IP address</th>
<th>Network mask</th>
<th>Default gateway</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interconnect 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interconnect 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Management 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Management 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Site A cluster creation information

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

<table>
<thead>
<tr>
<th>Type of information</th>
<th>Your values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cluster name</td>
<td>Example used in this guide: site_A</td>
</tr>
<tr>
<td>DNS domain</td>
<td></td>
</tr>
<tr>
<td>DNS name servers</td>
<td></td>
</tr>
<tr>
<td>Location</td>
<td></td>
</tr>
<tr>
<td>Administrator password</td>
<td></td>
</tr>
</tbody>
</table>

Site A node information

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

<table>
<thead>
<tr>
<th>Node</th>
<th>Port</th>
<th>IP address</th>
<th>Network mask</th>
<th>Default gateway</th>
</tr>
</thead>
<tbody>
<tr>
<td>Node 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Example used in this guide:</td>
<td></td>
<td>controller_A_1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Node 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Example used in this guide:</td>
<td></td>
<td>controller_A_2</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
**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* on page 8

<table>
<thead>
<tr>
<th>Node</th>
<th>Port</th>
<th>IP address of MetroCluster IP LIF</th>
<th>Network mask</th>
<th>Default gateway</th>
</tr>
</thead>
<tbody>
<tr>
<td>Node 1 MetroCluster IP LIF 1</td>
<td>e5a</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Node 1 MetroCluster IP LIF 2</td>
<td>e5b</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Node 2 MetroCluster IP LIF 1</td>
<td>e5a</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Node 2 MetroCluster IP LIF 2</td>
<td>e5b</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**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 1 IC LIF 1 | Port | IP address of intercluster LIF | Network mask | Default gateway |
| 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.

<table>
<thead>
<tr>
<th>Node</th>
<th>Host name</th>
<th>IP address</th>
<th>Network mask</th>
<th>Default gateway</th>
</tr>
</thead>
<tbody>
<tr>
<td>NTP server 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NTP server 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Site A AutoSupport information**

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

<table>
<thead>
<tr>
<th>Type of information</th>
<th>Your values</th>
</tr>
</thead>
<tbody>
<tr>
<td>From email address</td>
<td></td>
</tr>
<tr>
<td>Mail hosts</td>
<td>IP addresses or names</td>
</tr>
</tbody>
</table>
### 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:

<table>
<thead>
<tr>
<th>Node</th>
<th>IP address</th>
<th>Network mask</th>
<th>Default gateway</th>
</tr>
</thead>
<tbody>
<tr>
<td>Node 1</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### 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.

<table>
<thead>
<tr>
<th>Cluster switch</th>
<th>Host name</th>
<th>IP address</th>
<th>Network mask</th>
<th>Default gateway</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interconnect 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interconnect 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Management 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Management 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Site B cluster creation information

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

<table>
<thead>
<tr>
<th>Type of information</th>
<th>Your values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cluster name</td>
<td></td>
</tr>
<tr>
<td>Example used in this guide: site_B</td>
<td></td>
</tr>
<tr>
<td>DNS domain</td>
<td></td>
</tr>
<tr>
<td>DNS name servers</td>
<td></td>
</tr>
<tr>
<td>Location</td>
<td></td>
</tr>
<tr>
<td>Administrator password</td>
<td></td>
</tr>
</tbody>
</table>

### Site B node information

For each node in the cluster, you need a management IP address, a network mask, and a default gateway.
### 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* on page 8

<table>
<thead>
<tr>
<th>Node</th>
<th>Port</th>
<th>IP address of MetroCluster IP LIF</th>
<th>Network mask</th>
<th>Default gateway</th>
</tr>
</thead>
<tbody>
<tr>
<td>Node 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MetroCluster IP LIF 1</td>
<td>e5a</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Node 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MetroCluster IP LIF 2</td>
<td>e5b</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Node 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MetroCluster IP LIF 1</td>
<td>e5a</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Node 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MetroCluster IP LIF 2</td>
<td>e5b</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### 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.

<table>
<thead>
<tr>
<th>Node</th>
<th>Port</th>
<th>IP address of intercluster LIF</th>
<th>Network mask</th>
<th>Default gateway</th>
</tr>
</thead>
<tbody>
<tr>
<td>Node 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IC LIF 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Node 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IC LIF 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Node 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IC LIF 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Node 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IC LIF 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Site B time server information

You must synchronize the time, which requires one or more NTP time servers.
### Site B AutoSupport information

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

<table>
<thead>
<tr>
<th>Type of information</th>
<th>Your values</th>
</tr>
</thead>
<tbody>
<tr>
<td>From email address</td>
<td></td>
</tr>
<tr>
<td>Mail hosts</td>
<td>IP addresses or names</td>
</tr>
<tr>
<td>Transport protocol</td>
<td>HTTP, HTTPS, or SMTP</td>
</tr>
<tr>
<td>Proxy server</td>
<td></td>
</tr>
<tr>
<td>Recipient email addresses or distribution lists</td>
<td>Full-length messages</td>
</tr>
<tr>
<td></td>
<td>Concise messages</td>
</tr>
<tr>
<td></td>
<td>Partners</td>
</tr>
</tbody>
</table>

### 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:

<table>
<thead>
<tr>
<th>Node</th>
<th>IP address</th>
<th>Network mask</th>
<th>Default gateway</th>
</tr>
</thead>
<tbody>
<tr>
<td>Node 1 (controller_B_1)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### 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.

<table>
<thead>
<tr>
<th>Configuration step</th>
<th>Standard cluster configuration</th>
<th>MetroCluster configuration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Configure management, cluster, and data LIFs on each node.</td>
<td>Same in both types of clusters</td>
<td></td>
</tr>
<tr>
<td>Configure the root aggregate.</td>
<td>Same in both types of clusters</td>
<td></td>
</tr>
<tr>
<td>Set up the cluster on one node in the cluster.</td>
<td>Same in both types of clusters</td>
<td></td>
</tr>
<tr>
<td>Join the other node to the cluster.</td>
<td>Same in both types of clusters</td>
<td></td>
</tr>
<tr>
<td>Create a mirrored root aggregate.</td>
<td>Optional</td>
<td>Required</td>
</tr>
<tr>
<td>Peer the clusters.</td>
<td>Optional</td>
<td>Required</td>
</tr>
</tbody>
</table>
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.

Steps

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.

       Note: 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.

Steps
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.
Choices

- Manually assigning drives for pool 0 (ONTAP 9.4 and later) on page 76
- Manually assigning drives for pool 0 (ONTAP 9.3) on page 77

Manually assigning drives for pool 0 (ONTAP 9.4 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 9.4 or later.

To determine if your system requires manual disk assignment, you should review Considerations for automatic drive assignment and ADP systems in ONTAP 9.4 and later on page 9.

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)

Steps

1. Display the boot menu:

   boot_ontap menu

2. Select option 9a.

Example

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.

   a. When configuring node_A_1, manually assign drives from slot 0 to 11 to pool0 of node A1 from site_A-shelf_1.

   b. When configuring node_A_2, manually assign drives from slot 12 to 23 to pool0 of node A2 from site_A-shelf_1.

   c. When configuring node_B_1, manually assign drives from slot 0 to 11 to pool0 of node B1 from site_B-shelf_1.

   d. 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* on page 79.

### Manually assigning drives for pool 0 (ONTAP 9.3)

If you have at least two disk shelves for each node, you use ONTAP’s auto-assignment functionality to automatically assign the local (pool 0) disks. While the node is in Maintenance mode, you must first assign a single disk on the appropriate shelves to pool 0. ONTAP then automatically assigns the rest of the disks on the shelf to the same pool. This task is not required on systems received from the factory, which have pool 0 to contain the pre-configured root aggregate.

**About this task**

This procedure applies to configurations running ONTAP 9.3.

This procedure is not required if you received your MetroCluster configuration from the factory. Nodes from the factory are configured with pool 0 disks and root aggregates.

This procedure can be used only if you have at least two disk shelves for each node, which allows shelf-level autoassignment of disks. If you cannot use shelf-level autoassignment, you must manually assign your local disks so that each node has a local pool of disks (pool 0).

These steps must be performed in Maintenance mode.

Examples in this section assume the following disk shelves:

- node_A_1 owns disks on:
• site_A-shelf_1 (local)
• site_B-shelf_2 (remote)

• node_A_2 is connected to:
  ◦ site_A-shelf_3 (local)
  ◦ site_B-shelf_4 (remote)

• node_B_1 is connected to:
  ◦ site_B-shelf_1 (local)
  ◦ site_A-shelf_2 (remote)

• node_B_2 is connected to:
  ◦ site_B-shelf_3 (local)
  ◦ site_A-shelf_4 (remote)

Steps
1. Manually assign a single disk for root aggregate on each node:
   
   ```
   disk assign disk-id -p 0 -s local-node-sysid
   ```

   The manual assignment of these disks allows the ONTAP autoassignment feature to assign the rest of the disks on each shelf.

   a. On node_A_1, manually assign one disk from local site_A-shelf_1 to pool 0.
   b. On node_A_2, manually assign one disk from local site_A-shelf_3 to pool 0.
   c. On node_B_1, manually assign one disk from local site_B-shelf_1 to pool 0.
   d. On node_B_2, manually assign one disk from local site_B-shelf_3 to pool 0.

2. Boot each node at site A, using option 4 on the boot menu:

   You should complete this step on a node before proceeding to the next node.

   a. Exit Maintenance mode:
      
      ```
      halt
      ```

   b. Display the boot menu:
      
      ```
      boot_ontap menu
      ```

   c. Select option 4 from the boot menu and proceed.

3. Boot each node at site B, using option 4 on the boot menu:

   You should complete this step on a node before proceeding to the next node.

   a. Exit Maintenance mode:
      
      ```
      halt
      ```

   b. Display the boot menu:
      
      ```
      boot_ontap menu
      ```

   c. Select option 4 from the boot menu and proceed.
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.

About this task

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

Steps

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

   a. Enable the AutoSupport tool by following the directions provided by the system.

Example

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 NetApp 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
```
b. Configure the node management interface by responding to the prompts.

**Example**

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.
```

c. Create the cluster by responding to the prompts.

**Example**

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:

<table>
<thead>
<tr>
<th>Port</th>
<th>MTU</th>
<th>IP</th>
<th>Netmask</th>
</tr>
</thead>
<tbody>
<tr>
<td>e0a</td>
<td>1500</td>
<td>169.254.18.124</td>
<td>255.255.0.0</td>
</tr>
<tr>
<td>e1a</td>
<td>1500</td>
<td>169.254.184.44</td>
<td>255.255.0.0</td>
</tr>
</tbody>
</table>

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 ela: 172.17.10.229
Enter the cluster name: cluster_A

Creating cluster cluster_A
Starting cluster support services ...
Cluster cluster_A has been created.
```

d. Add licenses, set up a Cluster Administration SVM, and enter DNS information by responding to the prompts.
Example

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 []: sv1

e. Enable storage failover and set up the node by responding to the prompts.

Example

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

f. 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 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.

**Example**

The following example shows two cluster ports on node_A_1:

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

Node: node_A_1

<table>
<thead>
<tr>
<th>Port</th>
<th>IPspace</th>
<th>Broadcast Domain</th>
<th>Link MTU</th>
<th>Admin/Oper</th>
<th>Status</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>e4a</td>
<td>Cluster</td>
<td>Cluster</td>
<td>up</td>
<td>9000</td>
<td>auto/40000 healthy</td>
<td>false</td>
</tr>
<tr>
<td>e4e</td>
<td>Cluster</td>
<td>Cluster</td>
<td>up</td>
<td>9000</td>
<td>auto/40000 healthy</td>
<td>false</td>
</tr>
</tbody>
</table>

Node: node_A_2

<table>
<thead>
<tr>
<th>Port</th>
<th>IPspace</th>
<th>Broadcast Domain</th>
<th>Link MTU</th>
<th>Admin/Oper</th>
<th>Status</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>e4a</td>
<td>Cluster</td>
<td>Cluster</td>
<td>up</td>
<td>9000</td>
<td>auto/40000 healthy</td>
<td>false</td>
</tr>
<tr>
<td>e4e</td>
<td>Cluster</td>
<td>Cluster</td>
<td>up</td>
<td>9000</td>
<td>auto/40000 healthy</td>
<td>false</td>
</tr>
</tbody>
</table>
```

4 entries were displayed.

6. Repeat these steps on the partner cluster.

**After you finish**

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.

Disabling automatic drive assignment (if doing manual assignment in ONTAP 9.4)

In ONTAP 9.4, if your MetroCluster IP configuration has fewer than four external storage shelves per site, you must disable automatic drive assignment on all nodes and manually assign drives.

About this task
This task is not required in ONTAP 9.5 and later.
This task does not apply to an AFF A800 system with an internal shelf and no external shelves.

Considerations for automatic drive assignment and ADP systems in ONTAP 9.4 and later on page 9

Step
1. Disable automatic drive assignment:

   ```bash
   storage disk option modify -node node_name -autoassign off
   ```

   You need to issue this command on all nodes in the MetroCluster IP configuration.

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 9.4 and later on page 9

Step
1. Verify that pool 0 drives are assigned automatically:

   ```bash
disk show
   ```

Example

The following example shows the cluster_A output for an AFF A800 system with no external shelves.

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_A1:*> disk show

<table>
<thead>
<tr>
<th>Disk</th>
<th>Usable</th>
<th>Disk</th>
<th>Shelf</th>
<th>Bay</th>
<th>Container</th>
<th>Container</th>
</tr>
</thead>
<tbody>
<tr>
<td>node_A_1:0n.12</td>
<td>1.75TB</td>
<td>0</td>
<td>12</td>
<td></td>
<td>SSD-NVM</td>
<td>shared</td>
</tr>
<tr>
<td>node_A_1:0n.13</td>
<td>1.75TB</td>
<td>0</td>
<td>13</td>
<td></td>
<td>SSD-NVM</td>
<td>shared</td>
</tr>
<tr>
<td>node_A_1:0n.14</td>
<td>1.75TB</td>
<td>0</td>
<td>14</td>
<td></td>
<td>SSD-NVM</td>
<td>shared</td>
</tr>
<tr>
<td>node_A_1:0n.15</td>
<td>1.75TB</td>
<td>0</td>
<td>15</td>
<td></td>
<td>SSD-NVM</td>
<td>shared</td>
</tr>
<tr>
<td>node_A_1:0n.16</td>
<td>1.75TB</td>
<td>0</td>
<td>16</td>
<td></td>
<td>SSD-NVM</td>
<td>shared</td>
</tr>
<tr>
<td>node_A_1:0n.17</td>
<td>1.75TB</td>
<td>0</td>
<td>17</td>
<td></td>
<td>SSD-NVM</td>
<td>shared</td>
</tr>
<tr>
<td>node_A_1:0n.18</td>
<td>1.75TB</td>
<td>0</td>
<td>18</td>
<td></td>
<td>SSD-NVM</td>
<td>shared</td>
</tr>
<tr>
<td>node_A_1:0n.19</td>
<td>1.75TB</td>
<td>0</td>
<td>19</td>
<td></td>
<td>SSD-NVM</td>
<td>shared</td>
</tr>
<tr>
<td>node_A_2:0n.0</td>
<td>1.75TB</td>
<td>0</td>
<td>0</td>
<td></td>
<td>SSD-NVM</td>
<td>shared</td>
</tr>
</tbody>
</table>
```
The following example shows the cluster_B output:

```
cluster_B::> disk show
           Disk             Size       Shelf Bay Type    Type        Name      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      aggr0     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      aggr0_node_B_1_0 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.
```

**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.

**Steps**

1. Configuring intercluster LIFs on page 85
2. Creating a cluster peer relationship on page 89
Configuring the MetroCluster software in ONTAP | 85

Related concepts

- Considerations when using dedicated ports on page 17
- Considerations when sharing data ports on page 17

Related information

- Cluster and SVM peering express configuration

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.

Choices

- Configuring intercluster LIFs on dedicated ports on page 85
- Configuring intercluster LIFs on shared data ports on page 87

Configuring intercluster LIFs on dedicated ports

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

Steps

1. List the ports in the cluster:

   network port show

   For complete command syntax, see the man page.

Example

The following example shows the network ports in cluster01:

```
cluster01::> network port show

Node   Port      IPspace      Broadcast Domain Link   MTU    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.

Example

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
------- -------------------- --------- ---------
```

---
3. Create a failover group for the dedicated ports:

```bash
network interface failover-groups create -vserver system_SVM -failover-group failover_group -targets physical_or_logical_ports
```

**Example**

The following example assigns ports `e0e` and `e0f` to the failover group `intercluster01` on the system SVM `cluster01`:

```bash
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:

```bash
network interface failover-groups show
```

For complete command syntax, see the man page.

**Example**

```bash
class01::> network interface failover-groups show
Failover Vserver          Group            Targets
---------------- ---------------- --------------------------------------------
Cluster  Cluster          cluster01-01:e0a, cluster01-01:e0b,
            cluster01-02:e0a, cluster01-02:e0b
Cluster  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
Cluster  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:

```bash
network interface create -vserver system_SVM -lif LIF_name -role intercluster -home-node node -home-port port -address port_IP -netmask netmask -failover-group failover_group
```

For complete command syntax, see the man page.

**Example**

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

```bash
cluster01::> network interface create -vserver cluster01 -lif cluster01_icl01 -role 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 -role 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:
network interface show –role intercluster

For complete command syntax, see the man page.

Example

<table>
<thead>
<tr>
<th>Vserver</th>
<th>Logical Interface</th>
<th>Status</th>
<th>Network Address/Mask</th>
<th>Current Node</th>
<th>Current Is Port Home</th>
</tr>
</thead>
<tbody>
<tr>
<td>cluster01</td>
<td>cluster01_icl01</td>
<td>up/up</td>
<td>192.168.1.201/24</td>
<td>cluster01-01</td>
<td>e0e true</td>
</tr>
<tr>
<td>cluster01</td>
<td>cluster01_icl02</td>
<td>up/up</td>
<td>192.168.1.202/24</td>
<td>cluster01-02</td>
<td>e0f true</td>
</tr>
</tbody>
</table>

7. Verify that the intercluster LIFs are redundant:

network interface show –role intercluster –failover

For complete command syntax, see the man page.

Example

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.

<table>
<thead>
<tr>
<th>Vserver</th>
<th>Interface</th>
<th>Node:Port</th>
<th>Failover Policy</th>
<th>Failover Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>cluster01-01</td>
<td>cluster01-01_icl01</td>
<td>cluster01-01:e0e</td>
<td>local-only</td>
<td>intercluster01</td>
</tr>
<tr>
<td></td>
<td>Failover Targets:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>cluster01-01:e0e,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>cluster01-01:e0f</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cluster01-02</td>
<td>cluster01-02_icl02</td>
<td>cluster01-02:e0e</td>
<td>local-only</td>
<td>intercluster01</td>
</tr>
<tr>
<td></td>
<td>Failover Targets:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>cluster01-02:e0e,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>cluster01-02:e0f</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Related concepts

*Considerations when using dedicated ports* on page 17

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.

Steps

1. List the ports in the cluster:

   **network port show**

   For complete command syntax, see the man page.

Example

The following example shows the network ports in `cluster01`:

<table>
<thead>
<tr>
<th>Node Port</th>
<th>IPspace</th>
<th>Broadcast Domain</th>
<th>Link</th>
<th>MTU</th>
<th>Speed (Mbps)</th>
</tr>
</thead>
<tbody>
<tr>
<td>cluster01-01</td>
<td>e0a</td>
<td>Cluster</td>
<td>Cluster</td>
<td>up</td>
<td>1500</td>
</tr>
<tr>
<td></td>
<td>e0b</td>
<td>Cluster</td>
<td>Cluster</td>
<td>up</td>
<td>1500</td>
</tr>
<tr>
<td></td>
<td>e0c</td>
<td>Default</td>
<td>Default</td>
<td>up</td>
<td>1500</td>
</tr>
<tr>
<td></td>
<td>e0d</td>
<td>Default</td>
<td>Default</td>
<td>up</td>
<td>1500</td>
</tr>
<tr>
<td>cluster01-02</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
2. Create intercluster LIFs on the system SVM:

```
network interface create -vserver system_SVM -lif LIF_name -role intercluster -home-node node -home-port port -address port_IP -netmask netmask
```

For complete command syntax, see the man page.

**Example**

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

```
cluster01::> network interface create -vserver cluster01 -lif cluster01_icl01 -role 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 -role 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:

```
network interface show –role intercluster
```

For complete command syntax, see the man page.

**Example**

```
cluster01::> network interface show –role 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   cluster01_icl02 up/up      192.168.1.202/24   cluster01-02  e0c     true
```

4. Verify that the intercluster LIFs are redundant:

```
network interface show –role intercluster -failover
```

For complete command syntax, see the man page.

**Example**

```
cluster01::> network interface show -role 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
cluster01   cluster01_icl02 cluster01-02:e0c   local-only      192.168.1.202/24
```

**Related concepts**

*Considerations when sharing data ports* on page 17
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.
- The clusters must be running ONTAP 9.3 or later.

About this task

Steps

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-addrs peer_LIF_IPs -ipspace ipspace
   ```

   If you specify both `-generate-passphrase` and `-peer-addrs`, only the cluster whose intercluster LIFs are specified in `-peer-addrs` can use the generated password.

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

   Example

   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-addrs peer_LIF_IPs -ipspace ipspace
   ```

   For complete command syntax, see the man page.

   Example

   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-addrs 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`

   **Example**

   ```
   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
   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`

   **Example**

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

**Note:** The DR relationships cannot be changed after the DR groups are created.
Steps

1. Verify that the nodes are ready for creation of the DR group by entering the following command on each:

   \texttt{metrocluster configuration-settings show-status}

   \textbf{Example}

   The command output should show that the nodes are ready:

   \begin{verbatim}
   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.
   \end{verbatim}

2. Create the DR group:

   \texttt{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

```bash
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

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

*Considerations for MetroCluster IP configuration* on page 8

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

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

<table>
<thead>
<tr>
<th>Node</th>
<th>Interface</th>
<th>IP address</th>
<th>Subnet</th>
</tr>
</thead>
<tbody>
<tr>
<td>node_A_1</td>
<td>MetroCluster IP interface 1</td>
<td>10.1.1.1</td>
<td>10.1.1/24</td>
</tr>
<tr>
<td></td>
<td>MetroCluster IP interface 2</td>
<td>10.1.2.1</td>
<td>10.1.2/24</td>
</tr>
<tr>
<td>node_A_2</td>
<td>MetroCluster IP interface 1</td>
<td>10.1.1.2</td>
<td>10.1.1/24</td>
</tr>
<tr>
<td></td>
<td>MetroCluster IP interface 2</td>
<td>10.1.2.2</td>
<td>10.1.2/24</td>
</tr>
<tr>
<td>node_B_1</td>
<td>MetroCluster IP interface 1</td>
<td>10.1.1.3</td>
<td>10.1.1/24</td>
</tr>
<tr>
<td></td>
<td>MetroCluster IP interface 2</td>
<td>10.1.2.3</td>
<td>10.1.2/24</td>
</tr>
<tr>
<td>node_B_2</td>
<td>MetroCluster IP interface 1</td>
<td>10.1.1.4</td>
<td>10.1.1/24</td>
</tr>
<tr>
<td></td>
<td>MetroCluster IP interface 2</td>
<td>10.1.2.4</td>
<td>10.1.2/24</td>
</tr>
</tbody>
</table>

The port usage in the following examples is for an AFF A700 or a FAS9000 system.

- AFF A800 systems use ports e0b and e1b for the MetroCluster IP interfaces.
- AFF A300 and FAS8200 systems use ports e1a and e1b for the MetroCluster IP interfaces.
- AFF A220 and FAS2750 systems use VLAN ports e0a-10 and e0b-20 for the MetroCluster IP interfaces.
  These physical ports are also used as cluster interfaces. The VLANs are configured automatically.

#### Steps

1. Confirm that each node has disk autoassignment enabled:

   ```bash
   storage disk option show
   ```

   Disk autoassignment will assign pool 0 and pool 1 disks on a shelf-by-shelf basis.

   The `Auto Assign` column indicates whether disk autoassignment is enabled.
2. Verify you can create MetroCluster IP interfaces on the nodes:

   `metrocluster configuration-settings show-status`

**Example**

All nodes should be ready:

```
Cluster       Node         Configuration Settings Status
----------    -----------  ---------------------------------
cluster_A     node_A_1     ready for interface create
              node_A_2     ready for interface create
cluster_B     node_B_1     ready for interface create
              node_B_2     ready for interface create
```

3. Create the interfaces on node_A_1.

   **Note:** The port usage in the following examples is for an AFF A700 or a FAS9000 system (e5a and e5b). You must configure the interfaces on the correct ports for your platform model, as given in above.

   a. Configure the interface on port e5a on node_A_1:

   ```bash
   metrocluster configuration-settings interface create -cluster-name cluster_name -home-node node_name -home-port e5a -address ip-address -netmask 255.255.255.0
   ```

   **Example**

   The following example shows the creation of the interface on port e5a on node_A_1 with IP address 10.1.1.1:

   ```bash
   cluster_A::> metrocluster configuration-settings interface create -cluster-name cluster_A -home-node node_A_1 -home-port e5a -address 10.1.1.1 -netmask 255.255.255.0
   [Job 28] Job succeeded: Interface Create is successful.
   ```

   b. Configure the interface on port e5b on node_A_1:

   ```bash
   metrocluster configuration-settings interface create -cluster-name cluster_name -home-node node_name -home-port e5b -address ip-address -netmask 255.255.255.0
   ```

   **Example**

   The following example shows the creation of the interface on port e5b on node_A_1 with IP address 10.1.2.1:
4. Create the interfaces on node_A_2.

   **Note:** The port usage in the following examples is for an AFF A700 or a FAS9000 system (e5a and e5b). You must configure the interfaces on the correct ports for your platform model, as given in above.

   a. Configure the interface on port e5a on node_A_2:

   ```bash
   metrocluster configuration-settings interface create -cluster-name cluster_A -home-node node_A_2 -home-port e5a -address 10.1.1.2 -netmask 255.255.255.0
   [Job 28] Job succeeded: Interface Create is successful.
   ```

   **Example**

   The following example shows the creation of the interface on port e5a on node_A_2 with IP address 10.1.1.2:

   ```bash
   cluster_A::> metrocluster configuration-settings interface create -cluster-name cluster_A -home-node node_A_2 -home-port e5a -address 10.1.1.2 -netmask 255.255.255.0
   [Job 28] Job succeeded: Interface Create is successful.
   ```

   b. Configure the interface on port e5b on node_A_2:

   ```bash
   metrocluster configuration-settings interface create -cluster-name cluster_A -home-node node_A_2 -home-port e5b -address 10.1.2.2 -netmask 255.255.255.0
   ```

   **Example**

   The following example shows the creation of the interface on port e5b on node_A_2 with IP address 10.1.2.2:

   ```bash
   cluster_A::> metrocluster configuration-settings interface create -cluster-name cluster_A -home-node node_A_2 -home-port e5b -address 10.1.2.2 -netmask 255.255.255.0
   [Job 28] Job succeeded: Interface Create is successful.
   ```

5. Create the interfaces on node_B_1.

   **Note:** The port usage in the following examples is for an AFF A700 or a FAS9000 system (e5a and e5b). You must configure the interfaces on the correct ports for your platform model, as given in above.

   a. Configure the interface on port e5a on node_B_1:

   ```bash
   metrocluster configuration-settings interface create -cluster-name cluster_A -home-node node_B_1 -home-port e5a -address 10.1.1.3 -netmask 255.255.255.0
   ```

   **Example**

   The following example shows the creation of the interface on port e5a on node_B_1 with IP address 10.1.1.3:
b. Configure the interface on port e5b on node_B_1:

```
metrocluster configuration-settings interface create -cluster-name cluster_A -home-node node_B_1 -home-port e5b -address ip-address -netmask 255.255.255.0
```

**Example**

The following example shows the creation of the interface on port e5b on node_B_1 with IP address 10.1.2.3:

```
cluster_A::> metrocluster configuration-settings interface create -cluster-name cluster_A -home-node node_B_1 -home-port e5b -address 10.1.2.3 -netmask 255.255.255.0
[Job 28] Job succeeded: Interface Create is successful.
cluster_A::>
```

6. Create the interfaces on node_B_2.

**Note:** The port usage in the following examples is for an AFF A700 or a FAS9000 system (e5a and e5b). You must configure the interfaces on the correct ports for your platform model, as given in above.

a. Configure the interface on port e5a on node_B_2:

```
metrocluster configuration-settings interface create -cluster-name cluster_A -home-node node_B_1 -home-port e5a -address ip-address -netmask 255.255.255.0
```

**Example**

The following example shows the creation of the interface on port e5a on node_B_2 with IP address 10.1.1.4:

```
cluster_A::> metrocluster configuration-settings interface create -cluster-name cluster_A -home-node node_B_2 -home-port e5a -address 10.1.1.4 -netmask 255.255.255.0
[Job 28] Job succeeded: Interface Create is successful.
cluster_A::>
```

b. Configure the interface on port e5b on node_B_2:

```
metrocluster configuration-settings interface create -cluster-name cluster_A -home-node node_B_1 -home-port e5b -address ip-address -netmask 255.255.255.0
```

**Example**

The following example shows the creation of the interface on port e5b on node_B_2 with IP address 10.1.2.4:

```
cluster_A::> metrocluster configuration-settings interface create -cluster-name cluster_A -home-node node_B_2 -home-port e5b -address 10.1.2.4 -netmask 255.255.255.0
[Job 28] Job succeeded: Interface Create is successful.
cluster_A::>
```
7. Verify that the interfaces have been configured:

```
metrocluster configuration-settings interface show
```

**Example**

The following example shows that the configuration state for each interface is completed.

```
cluster_A::> metrocluster configuration-settings interface show
```

```
DR Group Cluster Node    Network Address Netmask         Gateway   State
----- ------- ------- --------------- --------------- --------- ----------
1     cluster_A  node_A_1 Home Port: e5a
       10.1.1.1     255.255.255.0   -         completed
       Home Port: e5b
       10.1.2.1     255.255.255.0   -         completed
       node_A_2 Home Port: e5a
       10.1.1.2
       Home Port: e5b
       10.1.2.2     255.255.255.0   -         completed

cluster_B

node_B_1 Home Port: e5a
       10.1.1.3     255.255.255.0   -         completed
       Home Port: e5b
       10.1.2.3     255.255.255.0   -         completed
       node_B_2 Home Port: e5a
       10.1.1.4
       Home Port: e5b
       10.1.2.4     255.255.255.0   -         completed

8 entries were displayed.
```

8. Verify that the nodes are ready to connect the MetroCluster interfaces:

```
metrocluster configuration-settings show-status
```

**Example**

The following example shows all nodes in the ready for connection state:

```
cluster_A::> metrocluster configuration-settings show-status
```

```
Cluster       Node         Configuration Settings Status
----------    -----------  ---------------------------------
cluster_A
node_A_1     ready for connection connect
node_A_2
cluster_B
node_B_1     ready for connection connect
node_B_2
```

4 entries were displayed.

9. Establish the connections:

```
metrocluster configuration-settings connection connect
```

The IP addresses cannot be changed after you issue this command.

**Example**

The following example shows cluster_A is successfully connected:

```
cluster_A::> metrocluster configuration-settings connection connect
[Job 53] Job succeeded: Connect is successful.
```

10. Verify that the connections have been established:

```
metrocluster configuration-settings show-status
```
Example

The configuration settings status for all nodes should be completed:

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Node</th>
<th>Configuration Settings Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>cluster_A</td>
<td>node_A_1</td>
<td>completed</td>
</tr>
<tr>
<td>cluster_A</td>
<td>node_A_2</td>
<td>completed</td>
</tr>
<tr>
<td>cluster_B</td>
<td>node_B_1</td>
<td>completed</td>
</tr>
<tr>
<td>cluster_B</td>
<td>node_B_2</td>
<td>completed</td>
</tr>
</tbody>
</table>

4 entries were displayed.

11. Verify that the iSCSI connections have been established:

a. Change to the advanced privilege level:

```
set -privilege advanced
```

You need to respond with y when you are prompted to continue into advanced mode and you see the advanced mode prompt (*>).

b. Display the connections:

```
storage iscsi-initiator show
```

On systems running ONTAP 9.5, there are eight MCCIP initiators on each cluster that should appear in the output.

On systems running ONTAP 9.4 and earlier, there are four MCCIP initiators on each cluster that should appear in the output.

Example

The following example shows the eight MCCIP initiators on a cluster running ONTAP 9.5:

```
cluster_A::*> storage iscsi-initiator show
----- ----- ------- ------------------      ----------------------------------------
cluster_A-01 dr_auxiliary
      mccip-aux-a-initiator
10.227.16.113:65200 prod506.com.company:abab44 up/up
10.227.16.113:65200 prod507.com.company:abab44 up/up
mccip-aux-a-initiator2
10.227.16.113:65200 prod506.com.company:abab44 up/up
10.227.16.113:65200 prod507.com.company:abab44 up/up
mccip-aux-b-initiator
10.227.95.166:65200 prod506.com.company:abab44 up/up
10.227.95.166:65200 prod507.com.company:abab44 up/up
mccip-aux-b-initiator2
10.227.95.166:65200 prod506.com.company:abab44 up/up
10.227.95.166:65200 prod507.com.company:abab44 up/up

cluster_A-02 dr_auxiliary
      mccip-aux-a-initiator
10.227.16.112:65200 prod506.com.company:cdcd88 up/up
10.227.16.112:65200 prod507.com.company:cdcd88 up/up
mccip-aux-a-initiator2
10.227.16.112:65200 prod506.com.company:cdcd88 up/up
10.227.16.112:65200 prod507.com.company:cdcd88 up/up
mccip-aux-b-initiator
10.227.95.165:65200 prod506.com.company:cdcd88 up/up
10.227.95.165:65200 prod507.com.company:cdcd88 up/up
mccip-aux-b-initiator2
10.227.95.165:65200 prod506.com.company:cdcd88 up/up
10.227.95.165:65200 prod507.com.company:cdcd88 up/up
```
c. Return to the admin privilege level:

   `set -privilege admin`

12. Verify that the nodes are ready for final implementation of the MetroCluster configuration:

   `metrocluster node show`

**Example**

```
cluster_A::> metrocluster node show
DR Group Cluster Node      | Configuration | DR Mirroring Mode
----- ------- -------------- | ------------- | ----------------|
          cluster_A        |               |
   -       node_A_1       | ready to config | -
   -       node_A_2       | ready to config | -
2 entries were displayed.
cluster_A::>
```

```
cluster_B::> metrocluster node show
DR Group Cluster Node      | Configuration | DR Mirroring Mode
----- ------- -------------- | ------------- | ----------------|
          cluster_B        |               |
   -       node_B_1       | ready to config | -
   -       node_B_2       | ready to config | -
2 entries were displayed.
cluster_B::>
```

**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.

**About this task**

<table>
<thead>
<tr>
<th>Configuration type</th>
<th>Procedure</th>
</tr>
</thead>
<tbody>
<tr>
<td>The systems meet the requirements for automatic drive assignment or, if running ONTAP 9.3, were received from the factory.</td>
<td><strong>Verifying disk assignment for pool 1 disks on page 99</strong></td>
</tr>
<tr>
<td>The configuration includes either three shelves, or, if it contains more than four shelves, has an uneven multiple of four shelves (for example, seven shelves), and is running ONTAP 9.5.</td>
<td><strong>Manually assigning drives for pool 1 (ONTAP 9.4 or later) on page 101</strong></td>
</tr>
<tr>
<td>The configuration does not include four storage shelves per site and is running ONTAP 9.5.</td>
<td><strong>Manually assigning drives for pool 1 (ONTAP 9.4 or later) on page 101</strong></td>
</tr>
<tr>
<td>The systems were not received from the factory and are running ONTAP 9.4</td>
<td><strong>Manually assigning disks for pool 1 (ONTAP 9.3) on page 102</strong></td>
</tr>
</tbody>
</table>
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 `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 9.4 and later on page 9

Step

1. Verify pool 1 disks are auto-assigned:

   `disk show`

   The following output shows the output for an AFF A800 system with no external shelves.

Example

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
Usable     Disk          Container   Container
Disk                Size       Shelf Bay Type    Type        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
Usable     Disk          Container   Container
Disk                Size       Shelf Bay Type    Type        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.

cluster_B::> disk show
Usable     Disk          Container   Container
Disk                Size       Shelf Bay Type    Type        Name      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.i1.0L14 1.75TB     0     18  SSD-NVM shared      -         node_A_1
node_B_1:0m.i1.0L16 1.75TB     0     19  SSD-NVM shared      -         node_A_1
node_B_1:0m.i1.0L19 1.75TB     0     20  SSD-NVM shared      -         node_A_1
node_B_1:0m.i1.0L25 1.75TB     0     21  SSD-NVM shared      -         node_A_1
node_B_1:0m.i1.0L28 1.75TB     0     22  SSD-NVM shared      -         node_A_1
node_B_1:0m.i1.0L29 1.75TB     0     14  SSD-NVM shared      -         node_A_1
node_B_1:0m.i1.0L31 1.75TB     0     15  SSD-NVM shared      -         node_A_1
node_B_1:0m.i1.0L32 1.75TB     0     16  SSD-NVM shared      -         node_A_1
node_B_1:0m.i1.0L34 1.75TB     0     17  SSD-NVM shared      -         node_A_1
node_B_1:0m.i1.0L35 1.75TB     0     18  SSD-NVM shared      -         node_A_1
node_B_1:0m.i1.0L36 1.75TB     0     19  SSD-NVM shared      -         node_A_1
node_B_1:0m.i1.0L37 1.75TB     0     20  SSD-NVM shared      -         node_A_1
node_B_1:0m.i1.0L38 1.75TB     0     21  SSD-NVM shared      -         node_A_1
node_B_1:0m.i1.0L39 1.75TB     0     22  SSD-NVM shared      -         node_A_1
node_B_1:0m.i1.0L40 1.75TB     0     23  SSD-NVM shared      -         node_A_1
node_B_1:0m.i1.0L41 1.75TB     0     24  SSD-NVM shared      -         node_A_1
node_B_1:0m.i1.0L42 1.75TB     0     25  SSD-NVM shared      -         node_A_1
node_B_1:0m.i1.0L43 1.75TB     0     26  SSD-NVM shared      -         node_A_1
node_B_1:0m.i1.0L44 1.75TB     0     27  SSD-NVM shared      -         node_A_1
node_B_1:0m.i1.0L45 1.75TB     0     28  SSD-NVM shared      -         node_A_1
node_B_1:0m.i1.0L46 1.75TB     0     29  SSD-NVM shared      -         node_A_1
node_B_1:0m.i1.0L47 1.75TB     0     30  SSD-NVM shared      -         node_A_1
node_B_1:0m.i1.0L48 1.75TB     0     31  SSD-NVM shared      -         node_A_1
node_B_1:0m.i1.0L49 1.75TB     0     32  SSD-NVM shared      -         node_A_1
node_B_1:0m.i1.0L50 1.75TB     0     33  SSD-NVM shared      -         node_A_1
node_B_1:0m.i1.0L51 1.75TB     0     34  SSD-NVM shared      -         node_A_1
node_B_1:0m.i1.0L52 1.75TB     0     35  SSD-NVM shared      -         node_A_1
node_B_1:0m.i1.0L53 1.75TB     0     36  SSD-NVM shared      -         node_A_1
node_B_1:0m.i1.0L54 1.75TB     0     37  SSD-NVM shared      -         node_A_1
node_B_1:0m.i1.0L55 1.75TB     0     38  SSD-NVM shared      -         node_A_1
node_B_1:0m.i1.0L56 1.75TB     0     39  SSD-NVM shared      -         node_A_1
node_B_1:0m.i1.0L57 1.75TB     0     40  SSD-NVM shared      -         node_A_1
node_B_1:0m.i1.0L58 1.75TB     0     41  SSD-NVM shared      -         node_A_1
node_B_1:0m.i1.0L59 1.75TB     0     42  SSD-NVM shared      -         node_A_1
8 entries were displayed.
```
<table>
<thead>
<tr>
<th>Disk Size</th>
<th>Shelf</th>
<th>Bay</th>
<th>Type</th>
<th>Name</th>
<th>Owner</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.75TB</td>
<td>0</td>
<td>43</td>
<td>SSD-NVM shared</td>
<td>node_B_1</td>
<td></td>
</tr>
<tr>
<td>0.894TB</td>
<td>0</td>
<td>24</td>
<td>SSD-NVM shared</td>
<td>node_B_1</td>
<td></td>
</tr>
<tr>
<td>0.894TB</td>
<td>0</td>
<td>25</td>
<td>SSD-NVM shared</td>
<td>node_B_1</td>
<td></td>
</tr>
<tr>
<td>0.894TB</td>
<td>0</td>
<td>26</td>
<td>SSD-NVM shared</td>
<td>node_B_1</td>
<td></td>
</tr>
<tr>
<td>0.894TB</td>
<td>0</td>
<td>27</td>
<td>SSD-NVM shared</td>
<td>node_B_1</td>
<td></td>
</tr>
<tr>
<td>0.894TB</td>
<td>0</td>
<td>28</td>
<td>SSD-NVM shared</td>
<td>node_B_1</td>
<td></td>
</tr>
<tr>
<td>0.894TB</td>
<td>0</td>
<td>29</td>
<td>SSD-NVM shared</td>
<td>node_B_1</td>
<td></td>
</tr>
<tr>
<td>0.894TB</td>
<td>0</td>
<td>30</td>
<td>SSD-NVM shared</td>
<td>node_B_1</td>
<td></td>
</tr>
<tr>
<td>0.894TB</td>
<td>0</td>
<td>31</td>
<td>SSD-NVM shared</td>
<td>node_B_1</td>
<td></td>
</tr>
<tr>
<td>0.894TB</td>
<td>0</td>
<td>32</td>
<td>SSD-NVM shared</td>
<td>node_B_1</td>
<td></td>
</tr>
<tr>
<td>0.894TB</td>
<td>0</td>
<td>33</td>
<td>SSD-NVM shared</td>
<td>node_B_1</td>
<td></td>
</tr>
<tr>
<td>0.894TB</td>
<td>0</td>
<td>34</td>
<td>SSD-NVM shared</td>
<td>node_B_1</td>
<td></td>
</tr>
<tr>
<td>0.894TB</td>
<td>0</td>
<td>35</td>
<td>SSD-NVM shared</td>
<td>node_B_1</td>
<td></td>
</tr>
<tr>
<td>0.894TB</td>
<td>0</td>
<td>36</td>
<td>SSD-NVM shared</td>
<td>node_B_1</td>
<td></td>
</tr>
<tr>
<td>0.894TB</td>
<td>0</td>
<td>37</td>
<td>SSD-NVM shared</td>
<td>node_B_1</td>
<td></td>
</tr>
<tr>
<td>0.894TB</td>
<td>0</td>
<td>38</td>
<td>SSD-NVM shared</td>
<td>node_B_1</td>
<td></td>
</tr>
<tr>
<td>0.894TB</td>
<td>0</td>
<td>39</td>
<td>SSD-NVM shared</td>
<td>node_B_1</td>
<td></td>
</tr>
<tr>
<td>0.894TB</td>
<td>0</td>
<td>40</td>
<td>SSD-NVM shared</td>
<td>node_B_1</td>
<td></td>
</tr>
<tr>
<td>0.894TB</td>
<td>0</td>
<td>41</td>
<td>SSD-NVM shared</td>
<td>node_B_1</td>
<td></td>
</tr>
<tr>
<td>0.894TB</td>
<td>0</td>
<td>42</td>
<td>SSD-NVM shared</td>
<td>node_B_1</td>
<td></td>
</tr>
<tr>
<td>0.894TB</td>
<td>0</td>
<td>43</td>
<td>SSD-NVM shared</td>
<td>node_B_1</td>
<td></td>
</tr>
<tr>
<td>0.894TB</td>
<td>0</td>
<td>44</td>
<td>SSD-NVM shared</td>
<td>node_B_1</td>
<td></td>
</tr>
<tr>
<td>0.894TB</td>
<td>0</td>
<td>45</td>
<td>SSD-NVM shared</td>
<td>node_B_1</td>
<td></td>
</tr>
<tr>
<td>0.894TB</td>
<td>0</td>
<td>46</td>
<td>SSD-NVM shared</td>
<td>node_B_1</td>
<td></td>
</tr>
<tr>
<td>0.894TB</td>
<td>0</td>
<td>47</td>
<td>SSD-NVM shared</td>
<td>node_B_1</td>
<td></td>
</tr>
<tr>
<td>0.894TB</td>
<td>0</td>
<td>48</td>
<td>SSD-NVM shared</td>
<td>node_B_1</td>
<td></td>
</tr>
<tr>
<td>0.894TB</td>
<td>0</td>
<td>49</td>
<td>SSD-NVM shared</td>
<td>node_B_1</td>
<td></td>
</tr>
<tr>
<td>0.894TB</td>
<td>0</td>
<td>50</td>
<td>SSD-NVM shared</td>
<td>node_B_1</td>
<td></td>
</tr>
<tr>
<td>0.894TB</td>
<td>0</td>
<td>51</td>
<td>SSD-NVM shared</td>
<td>node_B_1</td>
<td></td>
</tr>
<tr>
<td>0.894TB</td>
<td>0</td>
<td>52</td>
<td>SSD-NVM shared</td>
<td>node_B_1</td>
<td></td>
</tr>
<tr>
<td>0.894TB</td>
<td>0</td>
<td>53</td>
<td>SSD-NVM shared</td>
<td>node_B_1</td>
<td></td>
</tr>
<tr>
<td>0.894TB</td>
<td>0</td>
<td>54</td>
<td>SSD-NVM shared</td>
<td>node_B_1</td>
<td></td>
</tr>
<tr>
<td>0.894TB</td>
<td>0</td>
<td>55</td>
<td>SSD-NVM shared</td>
<td>node_B_1</td>
<td></td>
</tr>
<tr>
<td>0.894TB</td>
<td>0</td>
<td>56</td>
<td>SSD-NVM shared</td>
<td>node_B_1</td>
<td></td>
</tr>
<tr>
<td>0.894TB</td>
<td>0</td>
<td>57</td>
<td>SSD-NVM shared</td>
<td>node_B_1</td>
<td></td>
</tr>
<tr>
<td>0.894TB</td>
<td>0</td>
<td>58</td>
<td>SSD-NVM shared</td>
<td>node_B_1</td>
<td></td>
</tr>
<tr>
<td>0.894TB</td>
<td>0</td>
<td>59</td>
<td>SSD-NVM shared</td>
<td>node_B_1</td>
<td></td>
</tr>
<tr>
<td>0.894TB</td>
<td>0</td>
<td>60</td>
<td>SSD-NVM shared</td>
<td>node_B_1</td>
<td></td>
</tr>
<tr>
<td>0.894TB</td>
<td>0</td>
<td>61</td>
<td>SSD-NVM shared</td>
<td>node_B_1</td>
<td></td>
</tr>
<tr>
<td>0.894TB</td>
<td>0</td>
<td>62</td>
<td>SSD-NVM shared</td>
<td>node_B_1</td>
<td></td>
</tr>
<tr>
<td>0.894TB</td>
<td>0</td>
<td>63</td>
<td>SSD-NVM shared</td>
<td>node_B_1</td>
<td></td>
</tr>
<tr>
<td>0.894TB</td>
<td>0</td>
<td>64</td>
<td>SSD-NVM shared</td>
<td>node_B_1</td>
<td></td>
</tr>
<tr>
<td>0.894TB</td>
<td>0</td>
<td>65</td>
<td>SSD-NVM shared</td>
<td>node_B_1</td>
<td></td>
</tr>
<tr>
<td>0.894TB</td>
<td>0</td>
<td>66</td>
<td>SSD-NVM shared</td>
<td>node_B_1</td>
<td></td>
</tr>
<tr>
<td>0.894TB</td>
<td>0</td>
<td>67</td>
<td>SSD-NVM shared</td>
<td>node_B_1</td>
<td></td>
</tr>
<tr>
<td>0.894TB</td>
<td>0</td>
<td>68</td>
<td>SSD-NVM shared</td>
<td>node_B_1</td>
<td></td>
</tr>
<tr>
<td>0.894TB</td>
<td>0</td>
<td>69</td>
<td>SSD-NVM shared</td>
<td>node_B_1</td>
<td></td>
</tr>
<tr>
<td>0.894TB</td>
<td>0</td>
<td>70</td>
<td>SSD-NVM shared</td>
<td>node_B_1</td>
<td></td>
</tr>
<tr>
<td>0.894TB</td>
<td>0</td>
<td>71</td>
<td>SSD-NVM shared</td>
<td>node_B_1</td>
<td></td>
</tr>
<tr>
<td>0.894TB</td>
<td>0</td>
<td>72</td>
<td>SSD-NVM shared</td>
<td>node_B_1</td>
<td></td>
</tr>
<tr>
<td>0.894TB</td>
<td>0</td>
<td>73</td>
<td>SSD-NVM shared</td>
<td>node_B_1</td>
<td></td>
</tr>
<tr>
<td>0.894TB</td>
<td>0</td>
<td>74</td>
<td>SSD-NVM shared</td>
<td>node_B_1</td>
<td></td>
</tr>
<tr>
<td>0.894TB</td>
<td>0</td>
<td>75</td>
<td>SSD-NVM shared</td>
<td>node_B_1</td>
<td></td>
</tr>
<tr>
<td>0.894TB</td>
<td>0</td>
<td>76</td>
<td>SSD-NVM shared</td>
<td>node_B_1</td>
<td></td>
</tr>
<tr>
<td>0.894TB</td>
<td>0</td>
<td>77</td>
<td>SSD-NVM shared</td>
<td>node_B_1</td>
<td></td>
</tr>
<tr>
<td>0.894TB</td>
<td>0</td>
<td>78</td>
<td>SSD-NVM shared</td>
<td>node_B_1</td>
<td></td>
</tr>
<tr>
<td>0.894TB</td>
<td>0</td>
<td>79</td>
<td>SSD-NVM shared</td>
<td>node_B_1</td>
<td></td>
</tr>
<tr>
<td>0.894TB</td>
<td>0</td>
<td>80</td>
<td>SSD-NVM shared</td>
<td>node_B_1</td>
<td></td>
</tr>
<tr>
<td>0.894TB</td>
<td>0</td>
<td>81</td>
<td>SSD-NVM shared</td>
<td>node_B_1</td>
<td></td>
</tr>
<tr>
<td>0.894TB</td>
<td>0</td>
<td>82</td>
<td>SSD-NVM shared</td>
<td>node_B_1</td>
<td></td>
</tr>
<tr>
<td>0.894TB</td>
<td>0</td>
<td>83</td>
<td>SSD-NVM shared</td>
<td>node_B_1</td>
<td></td>
</tr>
<tr>
<td>0.894TB</td>
<td>0</td>
<td>84</td>
<td>SSD-NVM shared</td>
<td>node_B_1</td>
<td></td>
</tr>
<tr>
<td>0.894TB</td>
<td>0</td>
<td>85</td>
<td>SSD-NVM shared</td>
<td>node_B_1</td>
<td></td>
</tr>
<tr>
<td>0.894TB</td>
<td>0</td>
<td>86</td>
<td>SSD-NVM shared</td>
<td>node_B_1</td>
<td></td>
</tr>
<tr>
<td>0.894TB</td>
<td>0</td>
<td>87</td>
<td>SSD-NVM shared</td>
<td>node_B_1</td>
<td></td>
</tr>
<tr>
<td>0.894TB</td>
<td>0</td>
<td>88</td>
<td>SSD-NVM shared</td>
<td>node_B_1</td>
<td></td>
</tr>
<tr>
<td>0.894TB</td>
<td>0</td>
<td>89</td>
<td>SSD-NVM shared</td>
<td>node_B_1</td>
<td></td>
</tr>
<tr>
<td>0.894TB</td>
<td>0</td>
<td>90</td>
<td>SSD-NVM shared</td>
<td>node_B_1</td>
<td></td>
</tr>
<tr>
<td>0.894TB</td>
<td>0</td>
<td>91</td>
<td>SSD-NVM shared</td>
<td>node_B_1</td>
<td></td>
</tr>
<tr>
<td>0.894TB</td>
<td>0</td>
<td>92</td>
<td>SSD-NVM shared</td>
<td>node_B_1</td>
<td></td>
</tr>
<tr>
<td>0.894TB</td>
<td>0</td>
<td>93</td>
<td>SSD-NVM shared</td>
<td>node_B_1</td>
<td></td>
</tr>
<tr>
<td>0.894TB</td>
<td>0</td>
<td>94</td>
<td>SSD-NVM shared</td>
<td>node_B_1</td>
<td></td>
</tr>
<tr>
<td>0.894TB</td>
<td>0</td>
<td>95</td>
<td>SSD-NVM shared</td>
<td>node_B_1</td>
<td></td>
</tr>
<tr>
<td>0.894TB</td>
<td>0</td>
<td>96</td>
<td>SSD-NVM shared</td>
<td>node_B_1</td>
<td></td>
</tr>
<tr>
<td>0.894TB</td>
<td>0</td>
<td>97</td>
<td>SSD-NVM shared</td>
<td>node_B_1</td>
<td></td>
</tr>
<tr>
<td>0.894TB</td>
<td>0</td>
<td>98</td>
<td>SSD-NVM shared</td>
<td>node_B_1</td>
<td></td>
</tr>
<tr>
<td>0.894TB</td>
<td>0</td>
<td>99</td>
<td>SSD-NVM shared</td>
<td>node_B_1</td>
<td></td>
</tr>
<tr>
<td>0.894TB</td>
<td>0</td>
<td>100</td>
<td>SSD-NVM shared</td>
<td>node_B_1</td>
<td></td>
</tr>
</tbody>
</table>
Manually assigning drives for pool 1 (ONTAP 9.4 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 9.4 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 9.4 and later on page 9.

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)

Step

1. From each node in the MetroCluster IP configuration, assign remote drives to pool 1.
   a. Display the list of unassigned drives:

```
disk show -host-adapter 0m -container-type unassigned
```

Example

```
<table>
<thead>
<tr>
<th>Disk</th>
<th>Usable Size</th>
<th>Shelf</th>
<th>Bay</th>
<th>Disk Type</th>
<th>Container Type</th>
<th>Container Name</th>
<th>Owner</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6.23.0</td>
<td></td>
<td>23</td>
<td>0</td>
<td>SSD</td>
<td>unassigned</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>6.23.1</td>
<td></td>
<td>23</td>
<td>1</td>
<td>SSD</td>
<td>unassigned</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>node_A_2:0m.i1.2L51</td>
<td></td>
<td>21</td>
<td>14</td>
<td>SSD</td>
<td>unassigned</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>node_A_2:0m.i1.2L64</td>
<td></td>
<td>21</td>
<td>10</td>
<td>SSD</td>
<td>unassigned</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
```
b. 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.

c. Confirm that the drives were assigned to pool 1:
   disk show -host-adapter 0m -container-type unassigned

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

**Example**

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::>

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

e. Repeat these steps on site B.

**Manually assigning disks for pool 1 (ONTAP 9.3)**

If you have at least two disk shelves for each node, you use ONTAP's auto-assignment functionality to automatically assign the remote (pool1) disks. You must first assign a disk on the shelf to pool1. ONTAP then automatically assigns the rest of the disks on the shelf to the same pool.

**About this task**

This procedure applies to configurations running ONTAP 9.3.

This procedure can be used only if you have at least two disk shelves for each node, which allows shelf-level autoassignment of disks.

If you cannot use shelf-level autoassignment, you must manually assign your remote disks so that each node has a remote pool of disks (pool 1).

The ONTAP automatic disk assignment feature assigns the disks on a shelf-by-shelf basis. For example:

- All the disks on site_B-shelf_2 are autoassigned to pool1 of node_A_1
- All the disks on site_B-shelf_4 are autoassigned to pool1 of node_A_2
- All the disks on site_A-shelf_2 are autoassigned to pool1 of node_B_1
- All the disks on site_A-shelf_4 are autoassigned to pool1 of node_B_2

You must "seed" the autoassignment by specifying a single disk on each shelf.
Step

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

   a. Display the list of unassigned disks:

```
disk show -host-adapter 0m -container-type unassigned
```

   Example

```
cluster_A::> disk show -host-adapter 0m -container-type unassigned

<table>
<thead>
<tr>
<th>Disk Name</th>
<th>Owner</th>
<th>Size</th>
<th>Shelf</th>
<th>Bay</th>
<th>Type</th>
<th>Container Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.23.0</td>
<td></td>
<td>-</td>
<td>23</td>
<td>0</td>
<td>SSD</td>
<td>unassigned</td>
</tr>
<tr>
<td>node_A_2:0m.i1.2L51</td>
<td></td>
<td>-</td>
<td>21</td>
<td>14</td>
<td>SSD</td>
<td>unassigned</td>
</tr>
<tr>
<td>node_A_2:0m.i1.2L64</td>
<td></td>
<td>-</td>
<td>21</td>
<td>10</td>
<td>SSD</td>
<td>unassigned</td>
</tr>
</tbody>
</table>

48 entries were displayed.
```

   b. Select a remote disk (0m) and assign ownership of the disk to pool1 of the first node (for example, node_A_1):

```
disk assign -disk disk-id -pool 1 -owner owner-node-name
```

The `disk-id` must identify a disk on a remote shelf of `owner-node-name`.

The ONTAP disk autoassignment feature assigns all disks on the remote shelf that contains the specified disk.

   c. After waiting at least 60 seconds for disk autoassignment to take place, verify that the remote disks on the shelf were auto-assigned to pool 1:

```
disk show -host-adapter 0m -container-type unassigned
```

   Note: The iSCSI connection used to access the remote disks appears as device 0m.

Example

The following output shows that the disks on shelf 23 have now been assigned and no longer appear:

```
cluster_A::> disk show -host-adapter 0m -container-type unassigned

<table>
<thead>
<tr>
<th>Disk Name</th>
<th>Owner</th>
<th>Size</th>
<th>Shelf</th>
<th>Bay</th>
<th>Type</th>
<th>Container Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>node_A_2:0m.i1.2L51</td>
<td></td>
<td>21</td>
<td>14</td>
<td>SSD</td>
<td>unassigned</td>
<td></td>
</tr>
<tr>
<td>node_A_2:0m.i1.2L64</td>
<td></td>
<td>21</td>
<td>10</td>
<td>SSD</td>
<td>unassigned</td>
<td></td>
</tr>
<tr>
<td>node_A_2:0m.i1.2L72</td>
<td></td>
<td>21</td>
<td>23</td>
<td>SSD</td>
<td>unassigned</td>
<td></td>
</tr>
<tr>
<td>node_A_2:0m.i1.2L74</td>
<td></td>
<td>21</td>
<td>1</td>
<td>SSD</td>
<td>unassigned</td>
<td></td>
</tr>
<tr>
<td>node_A_2:0m.i1.2L83</td>
<td></td>
<td>21</td>
<td>22</td>
<td>SSD</td>
<td>unassigned</td>
<td></td>
</tr>
<tr>
<td>node_A_2:0m.i1.2L90</td>
<td></td>
<td>21</td>
<td>7</td>
<td>SSD</td>
<td>unassigned</td>
<td></td>
</tr>
<tr>
<td>node_A_2:0m.i1.3L52</td>
<td></td>
<td>21</td>
<td>6</td>
<td>SSD</td>
<td>unassigned</td>
<td></td>
</tr>
<tr>
<td>node_A_2:0m.i1.3L59</td>
<td></td>
<td>21</td>
<td>13</td>
<td>SSD</td>
<td>unassigned</td>
<td></td>
</tr>
<tr>
<td>node_A_2:0m.i1.3L66</td>
<td></td>
<td>21</td>
<td>17</td>
<td>SSD</td>
<td>unassigned</td>
<td></td>
</tr>
<tr>
<td>node_A_2:0m.i1.3L73</td>
<td></td>
<td>21</td>
<td>12</td>
<td>SSD</td>
<td>unassigned</td>
<td></td>
</tr>
<tr>
<td>node_A_2:0m.i1.3L80</td>
<td></td>
<td>21</td>
<td>5</td>
<td>SSD</td>
<td>unassigned</td>
<td></td>
</tr>
</tbody>
</table>
```
d. Repeat these steps to assign pool 1 disks to the second node on site A (for example, node_A_2).

e. Repeat these steps on site B.

Enabling automatic drive assignment in ONTAP 9.4

In ONTAP 9.4, if you disabled automatic drive assignment as directed previously in this procedure, you must reenable it on all nodes.

About this task

Considerations for automatic drive assignment and ADP systems in ONTAP 9.4 and later on page 9

Step

1. Enable automatic drive assignment:

   storage disk option modify -node node_name -autoassign on

   You must issue this command on all nodes in the MetroCluster IP configuration.

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

Note: 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.

Steps

1. Mirror the root aggregate:

   storage aggregate mirror aggr_name

Example

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*

**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 into P1, P2, and P3 partitions.
- Aggregate names should conform to the naming scheme you determined when you planned your MetroCluster configuration.
  
  *Disk and aggregate management*

**Steps**

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
     
     **Note:** 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.

**Example**

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.

  **Note:** If you want to use a single mirrored data aggregate, then see step 1 on page 106 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.

**Steps**

1. Configure the MetroCluster in the following format:

```
<table>
<thead>
<tr>
<th>If your MetroCluster configuration has...</th>
<th>Then do this...</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multiple data aggregates</td>
<td>From any node's prompt, configure MetroCluster:</td>
</tr>
<tr>
<td></td>
<td><code>metrocluster configure node-name</code></td>
</tr>
</tbody>
</table>
```
If your MetroCluster configuration has...  Then do this...

A single mirrored data aggregate

<table>
<thead>
<tr>
<th>a.</th>
<th>From any node's prompt, change to the advanced privilege level:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><code>set -privilege advanced</code></td>
</tr>
<tr>
<td></td>
<td>You need to respond with y when you are prompted to continue</td>
</tr>
<tr>
<td></td>
<td>into advanced mode and you see the advanced mode prompt (*&gt;).</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>b.</th>
<th>Configure the MetroCluster with the <code>-allow-with-one-aggregate</code> true parameter:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><code>metrocluster configure -allow-with-one-aggregate true node-name</code></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>c.</th>
<th>Return to the admin privilege level:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><code>set -privilege admin</code></td>
</tr>
</tbody>
</table>

Example

Note: 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
```

Example

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

```
cluster_A::> network port show

<table>
<thead>
<tr>
<th>Node</th>
<th>Port</th>
<th>IPspace</th>
<th>Broadcast Domain</th>
<th>Link</th>
<th>MTU</th>
<th>Speed (Mbps)</th>
<th>Admin/Oper</th>
</tr>
</thead>
<tbody>
<tr>
<td>controller_A_1</td>
<td>e0a</td>
<td>Cluster</td>
<td>Cluster</td>
<td>up</td>
<td>9000</td>
<td>auto/1000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>e0b</td>
<td>Cluster</td>
<td>Cluster</td>
<td>up</td>
<td>9000</td>
<td>auto/1000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>e0c</td>
<td>Default</td>
<td>Default</td>
<td>up</td>
<td>1500</td>
<td>auto/1000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>e0d</td>
<td>Default</td>
<td>Default</td>
<td>up</td>
<td>1500</td>
<td>auto/1000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>e0e</td>
<td>Default</td>
<td>Default</td>
<td>up</td>
<td>1500</td>
<td>auto/1000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>e0f</td>
<td>Default</td>
<td>Default</td>
<td>up</td>
<td>1500</td>
<td>auto/1000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>e0g</td>
<td>Default</td>
<td>Default</td>
<td>up</td>
<td>1500</td>
<td>auto/1000</td>
<td></td>
</tr>
<tr>
<td>controller_A_2</td>
<td>e0a</td>
<td>Cluster</td>
<td>Cluster</td>
<td>up</td>
<td>9000</td>
<td>auto/1000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>e0b</td>
<td>Cluster</td>
<td>Cluster</td>
<td>up</td>
<td>9000</td>
<td>auto/1000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>e0c</td>
<td>Default</td>
<td>Default</td>
<td>up</td>
<td>1500</td>
<td>auto/1000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>e0d</td>
<td>Default</td>
<td>Default</td>
<td>up</td>
<td>1500</td>
<td>auto/1000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>e0e</td>
<td>Default</td>
<td>Default</td>
<td>up</td>
<td>1500</td>
<td>auto/1000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>e0f</td>
<td>Default</td>
<td>Default</td>
<td>up</td>
<td>1500</td>
<td>auto/1000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>e0g</td>
<td>Default</td>
<td>Default</td>
<td>up</td>
<td>1500</td>
<td>auto/1000</td>
<td></td>
</tr>
</tbody>
</table>

14 entries were displayed.

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

Steps

1. Check the configuration:

   `metrocluster check run`

   Example

   The command runs as a background job and might not be completed immediately.
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 show
  metrocluster check cluster show
  metrocluster check config-replication show
  metrocluster check lif show
  metrocluster 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.

Example

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

<table>
<thead>
<tr>
<th>Node</th>
<th>Aggregate</th>
<th>Check</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>controller_A_1</td>
<td>controller_A_1_aggr0</td>
<td>mirroring-status ok</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>disk-pool-allocation ok</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>ownership-state ok</td>
<td></td>
</tr>
<tr>
<td>controller_A_1</td>
<td>controller_A_1_aggr1</td>
<td>mirroring-status ok</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>disk-pool-allocation ok</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>ownership-state ok</td>
<td></td>
</tr>
<tr>
<td>controller_A_1</td>
<td>controller_A_1_aggr2</td>
<td>mirroring-status ok</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>disk-pool-allocation ok</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>ownership-state ok</td>
<td></td>
</tr>
<tr>
<td>controller_A_2</td>
<td>controller_A_2_aggr0</td>
<td>mirroring-status ok</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>disk-pool-allocation ok</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>ownership-state ok</td>
<td></td>
</tr>
<tr>
<td>controller_A_2</td>
<td>controller_A_2_aggr1</td>
<td>mirroring-status ok</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>disk-pool-allocation ok</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>ownership-state ok</td>
<td></td>
</tr>
<tr>
<td>controller_A_2</td>
<td>controller_A_2_aggr2</td>
<td>mirroring-status ok</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>disk-pool-allocation ok</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>ownership-state ok</td>
<td></td>
</tr>
</tbody>
</table>

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.

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Check</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>mccint-fas9000-0102</td>
<td>negotiated-switchover-ready</td>
<td>not-applicable</td>
</tr>
<tr>
<td></td>
<td>switchback-ready</td>
<td>not-applicable</td>
</tr>
<tr>
<td></td>
<td>job-schedules</td>
<td>ok</td>
</tr>
<tr>
<td></td>
<td>licenses</td>
<td>ok</td>
</tr>
<tr>
<td></td>
<td>periodic-check-enabled</td>
<td>ok</td>
</tr>
<tr>
<td>mccint-fas9000-0304</td>
<td>negotiated-switchover-ready</td>
<td>not-applicable</td>
</tr>
<tr>
<td></td>
<td>switchback-ready</td>
<td>not-applicable</td>
</tr>
<tr>
<td></td>
<td>job-schedules</td>
<td>ok</td>
</tr>
<tr>
<td></td>
<td>licenses</td>
<td>ok</td>
</tr>
<tr>
<td></td>
<td>periodic-check-enabled</td>
<td>ok</td>
</tr>
</tbody>
</table>

10 entries were displayed.

Related information

- Disk and aggregate management
- Network and LIF management

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.

Step

1. Use the procedures for negotiated switchover, healing, and switchback that are mentioned in the MetroCluster Management and Disaster Recovery Guide.

Installing the MetroCluster Tiebreaker software

You can download and install Tiebreaker software to monitor the two clusters and the connectivity status between them from a third site. Doing so enables each partner in a cluster to distinguish between an ISL failure (when inter-site links are down) and a site failure.

Before you begin

You must have a Linux host available that has network connectivity to both clusters in the MetroCluster configuration.
Steps

1. Go to MetroCluster Tiebreaker Software Download page.
   
   NetApp Downloads: MetroCluster Tiebreaker for Linux

2. Follow the directions to download the Tiebreaker software and documentation.

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.

Step

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.

   System administration

Related information

   System administration
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 from where the command is issued should switch over to the partner site.
- Nodes at the partner site should shut down gracefully and remain at the LOADER> prompt.

Steps

1. Confirm that all nodes are in the configured state and normal mode:
   `metrocluster node show`

   **Example**

   ```
   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`

   **Example**

   ```
   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`

   **Example**

   ```
   cluster_A::> metrocluster node show
   Cluster                   Configuration State    Mode
   ------------------------- ---------------------- --------
   Local: cluster_A          configured             switchover
   Remote: cluster_B         configured             normal
   ```
4. Confirm that the switchover operation was successful:

   `metrocluster operation show`

   **Example**

   ```
   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 9.5 or later, on which healing is performed automatically after a negotiated switchover. On systems running ONTAP 9.6 and later, healing is also performed automatically after unscheduled switchover.

#### Steps

1. If the system is running ONTAP 9.4 or earlier, heal the data aggregate:

   `metrocluster heal aggregates`

   **Example**

   The following example shows the successful completion of the command:

   ```
   cluster_A::> metrocluster heal aggregates
   [Job 936] Job succeeded: Heal Aggregates is successful.
   ```

2. If it is required by your configuration, heal the root aggregate:

   `metrocluster heal root-aggregates`

   This step is required on the following configurations:

   - MetroCluster FC configurations.
   - MetroCluster IP configurations running ONTAP 9.4 or earlier.

   **Example**

   The following example shows the successful completion of the command:
3. Verify that healing is completed:

`metrocluster node show`

**Example**

The following example shows the successful completion of the command:

```
cluster_A::> metrocluster node show
DR Group Cluster Node  Configuration  DR
----- ------- ------------------ -------------- --------- -------------- -------------
1  cluster_A node_A_1        configured     enabled heal roots completed
    cluster_B node_B_2        unreachable - switched over
42 entries were displayed.
```

If the automatic healing operation fails for any reason, you must issue the `metrocluster heal` commands manually as done in ONTAP versions prior to ONTAP 9.5. You can use the `metrocluster operation show` and `metrocluster operation history show` commands to monitor the status of healing and determine the cause of a failure.

4. Verify that all aggregates are mirrored:

`storage aggregate show`

**Example**

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`

**Example**

```
cluster_A::> metrocluster node show
DR Group Cluster Node  Configuration  DR
----- ------- ------------------ -------------- --------- -------------
1  cluster_A node_A_1        configured     enabled heal roots completed
```

---

---
7. Perform the switchback:

```
metrocluster switchback
```

**Example**

```
cluster_A::> metrocluster switchback
[Job 938] Job succeeded: Switchback is successful. Verify switchback
```

8. Confirm status of the nodes:

```
metrocluster node show
```

**Example**

```
cluster_A::> metrocluster node show
DR   Group  Cluster Node   Configuration  DR
----- ------- -------------- -------------- ---------
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
```

**Example**

```
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 bridge is switched off.
- No failover or loss of service should occur.
- Only one path from the controller module to the drives behind the bridge is available.

**Steps**

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.

**Example**

```
cluster_A::> system environment sensors show -state fault

--------- --------------------- ------ ----------- -------- --------
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.

**Steps**

1. Check the storage failover status:
   ```
   storage failover show
   ```

   **Example**
   ```
   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
   ```

   **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, 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 data
   network interface show -fields is-home false
   volume show !vol0,!MDV*
   ```

   **Example**
   ```
   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*

<table>
<thead>
<tr>
<th>Vserver</th>
<th>Volume</th>
<th>Aggregate</th>
<th>State</th>
<th>Type</th>
<th>Size</th>
<th>Available</th>
<th>Used%</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVM1</td>
<td>SVM1_root</td>
<td>node_A_1data01_mirrored</td>
<td>online</td>
<td>RW</td>
<td>10GB</td>
<td>9.50GB</td>
<td>5%</td>
</tr>
<tr>
<td>SVM1</td>
<td>SVM1_data_vol</td>
<td>node_A_1data01_mirrored</td>
<td>online</td>
<td>RW</td>
<td>10GB</td>
<td>9.49GB</td>
<td>5%</td>
</tr>
<tr>
<td>SVM2</td>
<td>SVM2_root</td>
<td>node_A_2_data01_mirrored</td>
<td>online</td>
<td>RW</td>
<td>10GB</td>
<td>9.49GB</td>
<td>5%</td>
</tr>
<tr>
<td>SVM2</td>
<td>SVM2_data_vol</td>
<td>node_A_2_data02_unmirrored</td>
<td>online</td>
<td>RW</td>
<td>1GB</td>
<td>972.6MB</td>
<td>5%</td>
</tr>
</tbody>
</table>

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.

   **Example**

   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 show`
   `storage aggregate show -r -node node_A_2 !*root`

   **Example**

   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:
7. Verify that the data is being served and that all volumes are still online:

```
vserver show -type data
network interface show -fields is-home false
volume show !vol0,!MDV*
```

Example

```
cluster_A::> vserver show -type data
cluster_A::> vserver show -type data
cluster_A::> vserver show -type data
cluster_A::> network interface show -fields is-home false
There are no entries matching your query.
cluster_A::> volume show !vol0,!MDV*
```

Testing the MetroCluster configuration | 119
8. Physically power on the shelf.
   Resynchronization starts automatically.

9. Verify that resynchronization has started:
   ```
   storage aggregate show
   ```

   **Example**
   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,
                                 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,
                                 normal
   ```

10. Monitor the aggregate to confirm that resynchronization is complete:
    ```
    storage aggregate show
    ```

    **Example**
    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_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,
                                 normal
    node_A_2_data02_unmirrored 2.18TB    2.18TB    0% online       1 node_A_2       raid_dp,
<table>
<thead>
<tr>
<th>node_A_2_root</th>
<th>707.7GB</th>
<th>34.27GB</th>
<th>95% online</th>
<th>1 node_A_2</th>
<th>raid_dp, resyncing</th>
</tr>
</thead>
</table>

Testing the MetroCluster configuration | 121
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.

Attention: You cannot reverse the MetroCluster unconfiguration. This process should only be done with the assistance of technical support.
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 9.5 or later.

FlexGroup support in MetroCluster configurations

Starting with ONTAP 9.6 MetroCluster configurations support FlexGroup volumes.

Job schedules in a MetroCluster configuration

In ONTAP 9.3 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).

Note: 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 8.3 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 concepts

- IPspace object replication and subnet configuration requirements on page 124
- Requirements for LIF creation in a MetroCluster configuration on page 125
- LIF replication and placement requirements and issues on page 125

Related information

- Network and LIF management

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 Domain Gateway          Avail/ Total Ranges
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 Domain Gateway          Avail/ Total Ranges
Name      Subnet            Domain    Gateway           Total    Ranges
--------- ---------------- --------- ------------      -------
```
IPv6 configuration

If IPv6 is configured on one site, IPv6 must be configured on the other site as well.

Related concepts

Requirements for LIF creation in a MetroCluster configuration on page 125
LIF replication and placement requirements and issues on page 125

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

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 concepts

IPspace object replication and subnet configuration requirements on page 124
LIF replication and placement requirements and issues on page 125

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:
<table>
<thead>
<tr>
<th>Condition</th>
<th>LIF type: FC</th>
<th>LIF type: IP/iSCSI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Node identification</td>
<td>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.</td>
<td>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.</td>
</tr>
<tr>
<td>Port identification</td>
<td>ONTAP identifies the connected FC target ports on the DR cluster.</td>
<td>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.</td>
</tr>
<tr>
<td>Reachability check</td>
<td>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.</td>
<td>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.</td>
</tr>
</tbody>
</table>
### 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.

<table>
<thead>
<tr>
<th>Condition</th>
<th>LIF type: FC</th>
<th>LIF type: IP/iSCSI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Port selection</td>
<td>ONTAP categorizes the ports based on attributes such as adapter type and speed, and then selects the ports with matching attributes. If no ports with matching attributes are found, the LIF is placed on a random connected port on the DR partner.</td>
<td>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. 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. 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. ONTAP categorizes the ports based on attributes such as adapter type, interface type, and speed, and then selects the ports with matching attributes.</td>
</tr>
<tr>
<td>LIF placement</td>
<td>From the reachable ports, ONTAP selects the least loaded port for placement.</td>
<td>From the selected ports, ONTAP selects the least loaded port for placement.</td>
</tr>
</tbody>
</table>
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.

**Note:** 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 concepts

- [IPspace object replication and subnet configuration requirements](#) on page 124
- [Requirements for LIF creation in a MetroCluster configuration](#) on page 125

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

Starting with ONTAP 9.5, active storage virtual machines (SVMs) in a MetroCluster configuration can be used as sources with the SnapMirror SVM disaster recovery feature. 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. 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
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.

About this task

Note: In ONTAP versions earlier than 9.0, the NVFAIL flag is used for each switchover. In ONTAP 9.0 and later versions, the unplanned switchover (USO) is used.

Step

1. Enable MetroCluster configuration to trigger NVFAIL on switchover by setting the `vol -dr-force-nvfail` parameter to `on`:

   ```bash
   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.

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:

<table>
<thead>
<tr>
<th>If...</th>
<th>Then...</th>
</tr>
</thead>
<tbody>
<tr>
<td>ONTAP detects no NVRAM errors</td>
<td>File service starts normally.</td>
</tr>
<tr>
<td>If...</td>
<td>Then...</td>
</tr>
<tr>
<td>-------</td>
<td>--------</td>
</tr>
</tbody>
</table>
| ONTAP detects NVRAM errors | • ONTAP returns a stale file handle (ESTALE) 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 in-nvfailed-state option on the affected volume. |
| If one of the following parameters is used:  
• dr-force-nvfail volume option is set  
• force-nvfail-all switchover command option is set. | You can unset the dr-force-nvfail option after the switchover, if the administrator is not expecting to force NVFAIL processing for possible future disaster switchover operations. For NFS clients, files remain inaccessible until you reset the in-nvfailed-state option on the affected volume.  
**Note:** Using the force-nvfail-all option causes the dr-force-nvfail option to be set on all of the DR volumes processed during the disaster switchover. |
| ONTAP detects NVRAM errors on a volume that contains LUNs | LUNs in that volume are brought offline. The in-nvfailed-state option on the volume must be cleared, and the NVFAIL attribute on the LUNs must be cleared by bringing each LUN in the affected volume online.  
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 in-nvfailed-state option on the affected volume is cleared. |

**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.

<table>
<thead>
<tr>
<th>If you want to...</th>
<th>Use this command...</th>
</tr>
</thead>
<tbody>
<tr>
<td>Create a new volume with NVFAIL enabled</td>
<td><code>volume create -nvfail on</code></td>
</tr>
</tbody>
</table>
| Enable NVFAIL on an existing volume | `volume modify`  
**Note:** You set the -nvfail option to on to enable NVFAIL on the created volume. |
If you want to... | Use this command...
---|---
Display whether NVFAIL is currently enabled for a specified volume | `volume show`

**Note:** You set the `--fields` parameter to `nvfail` 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.

**Step**

1. Recover the volume by using the `volume modify` command with the `-in-nvfailed-state` parameter set to `false`.

**After you finish**

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.

**Steps**

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.
### Where to find additional information

You can learn more about MetroCluster IP configuration and operation from the NetApp documentation library.

#### MetroCluster and miscellaneous guides

<table>
<thead>
<tr>
<th>Guide</th>
<th>Content</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ONTAP 9 Documentation Center</strong></td>
<td>• All MetroCluster guides</td>
</tr>
<tr>
<td><strong>Fabric-attached MetroCluster installation and configuration</strong></td>
<td>• Fabric-attached MetroCluster architecture</td>
</tr>
<tr>
<td></td>
<td>• Cabling the configuration</td>
</tr>
<tr>
<td></td>
<td>• Configuring the FC-to-SAS bridges</td>
</tr>
<tr>
<td></td>
<td>• Configuring the FC switches</td>
</tr>
<tr>
<td></td>
<td>• Configuring the MetroCluster in ONTAP</td>
</tr>
<tr>
<td><strong>Stretch MetroCluster installation and configuration</strong></td>
<td>• Stretch MetroCluster architecture</td>
</tr>
<tr>
<td></td>
<td>• Cabling the configuration</td>
</tr>
<tr>
<td></td>
<td>• Configuring the FC-to-SAS bridges</td>
</tr>
<tr>
<td></td>
<td>• Configuring the MetroCluster in ONTAP</td>
</tr>
<tr>
<td><strong>MetroCluster management and disaster recovery</strong></td>
<td>• Understanding the MetroCluster configuration</td>
</tr>
<tr>
<td></td>
<td>• Switchover, healing and switchback</td>
</tr>
<tr>
<td></td>
<td>• Disaster recovery</td>
</tr>
<tr>
<td><strong>MetroCluster Tiebreaker Software Installation and Configuration Guide</strong></td>
<td>• Monitoring the MetroCluster configuration with the MetroCluster Tiebreaker software</td>
</tr>
<tr>
<td><strong>Note:</strong> The standard storage shelf maintenance procedures can be used with MetroCluster IP configurations.</td>
<td>• Hot-adding a disk shelf</td>
</tr>
<tr>
<td></td>
<td>• Hot-removing a disk shelf</td>
</tr>
<tr>
<td><strong>Copy-based transition</strong></td>
<td>• Transitioning data from 7-Mode storage systems to clustered storage systems</td>
</tr>
<tr>
<td><strong>ONTAP concepts</strong></td>
<td>• How mirrored aggregates work</td>
</tr>
</tbody>
</table>
Copyright © 2019 NetApp, Inc. All rights reserved. Printed in the U.S.

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 NetApp material is subject to the following license and disclaimer:

THIS SOFTWARE IS PROVIDED BY NETAPP "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 NETAPP 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.

NetApp reserves the right to change any products described herein at any time, and without notice. NetApp assumes no responsibility or liability arising from the use of products described herein, except as expressly agreed to in writing by NetApp. 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 NetApp.

The product described in this manual may be protected by one or more U.S. patents, foreign patents, or pending applications.

Data contained herein pertains to a commercial item (as defined in FAR 2.101) and is proprietary to NetApp, Inc. The U.S. Government has a non-exclusive, non-transferrable, non-sublicensable, worldwide, limited irrevocable license to use the Data only in connection with and in support of the U.S. Government contract under which the Data was delivered. Except as provided herein, the Data may not be used, disclosed, reproduced, modified, performed, or displayed without the prior written approval of NetApp, Inc. United States Government license rights for the Department of Defense are limited to those rights identified in DFARS clause 252.227-7015(b).
Trademark

NETAPP, the NETAPP logo, and the marks listed on the NetApp Trademarks page are trademarks of NetApp, Inc. Other company and product names may be trademarks of their respective owners.

How to send comments about documentation and receive update notifications

You can help us to improve the quality of our documentation by sending us your feedback. You can receive automatic notification when production-level (GA/FCS) documentation is initially released or important changes are made to existing production-level documents.

If you have suggestions for improving this document, send us your comments by email.

doccomments@netapp.com

To help us direct your comments to the correct division, include in the subject line the product name, version, and operating system.

If you want to be notified automatically when production-level documentation is released or important changes are made to existing production-level documents, follow Twitter account @NetAppDoc.

You can also contact us in the following ways:

- NetApp, Inc., 1395 Crossman Ave., Sunnyvale, CA 94089 U.S.
- Telephone: +1 (408) 822-6000
- Fax: +1 (408) 822-4501
- Support telephone: +1 (888) 463-8277