ONTAP® 9

MetroCluster® Management and Disaster Recovery Guide

December 2020 | 215-11167_2020-12_en-us
doccomments@netapp.com

Updated for ONTAP 9.8

NetApp®

Contents

Understanding MetroCluster data protection and disaster recovery

It is helpful to understand how MetroCluster protects data and provides transparent recovery from failures so that you can manage your switchover and switchback activities easily and efficiently.

MetroCluster uses mirroring to protect the data in a cluster. It provides disaster recovery through a single MetroCluster command that activates a secondary on the survivor site to serve the mirrored data originally owned by a primary site affected by disaster.

How eight- and four-node MetroCluster configurations provide local failover and switchover

Eight- and four-node MetroCluster configurations protect data on both a local level and cluster level. If you are setting up a MetroCluster configuration, you need to know how MetroCluster configurations protect your data.

MetroCluster configurations protect data by using two physically separated, mirrored clusters. Each cluster synchronously mirrors the data and storage virtual machine (SVM) configuration of the other. When a disaster occurs at one site, an administrator can activate the mirrored SVM and begin serving the mirrored data from the surviving site. Additionally, the nodes in each cluster are configured as an HA pair, providing a level of local failover.

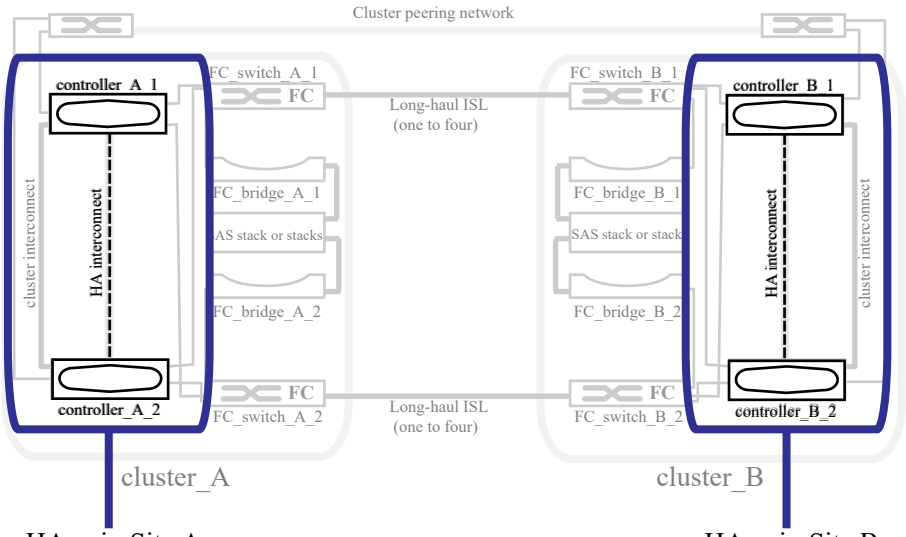
Attention:

How local HA data protection works in a MetroCluster configuration

You need to understand how HA pairs work in the MetroCluster configuration.

The two clusters in the peered network provide bidirectional disaster recovery, where each cluster can be the source and backup of the other cluster. Each cluster includes two nodes, which are configured as an HA pair. In the case of a failure or required maintenance within a single node's configuration, storage failover can transfer that node's operations to its local HA partner.

The following illustration shows a MetroCluster FC configuration. The HA functionality is the same in MetroCluster IP configurations, except that the HA interconnect is provided by the cluster switches.



Related information

High-availability configuration

How MetroCluster configurations provide data and configuration replication

MetroCluster configurations use a variety of ONTAP features to provide synchronous replication of data and configuration between the two MetroCluster sites.

Configuration protection with the configuration replication service

The ONTAP configuration replication service (CRS) protects the MetroCluster configuration by automatically replicating the information to the DR partner.

The CRS synchronously replicates local node configuration to the DR partner in the partner cluster. This replication is carried out over the cluster peering network.

The information replicated includes the cluster configuration and the SVM configuration.

Replication of SVMs during MetroCluster operations

The ONTAP configuration replication service (CRS) provides redundant data server configuration and mirroring of data volumes that belong to the SVM. If a switchover occurs, the source SVM is brought down and the destination SVM, located on the surviving cluster, becomes active.

> **Note:** Destination SVMs in the MetroCluster configuration have the suffix "`-mc`" automatically appended to their name to help identify them. A MetroCluster configuration appends the suffix "`-mc`" to the name of the destination SVMs, if the SVM name contains a period, the suffix "`-mc`" is applied prior to the first period. For example, if the SVM name is SVM.DNS.NAME, then the suffix "`-mc`" is appended as SVM-MC.DNS.NAME.

The following example shows the SVMs for a MetroCluster configuration, where SVM_cluster_A is an SVM on the source site and SVM_cluster_A-mc is a sync-destination aggregate on the disaster recovery site.

- SVM_cluster_A serves data on cluster A.

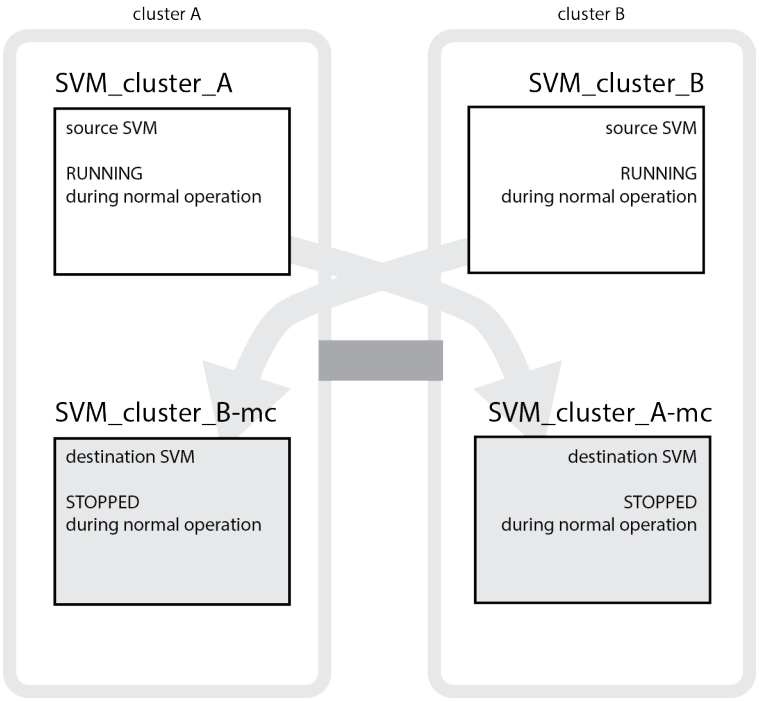
  It is a sync-source SVM that represents the SVM configuration (LIFs, protocols, and services) and data in volumes belonging to the SVM. The configuration and data are replicated to SVM_cluster_A-mc, a sync-destination SVM located on cluster B.
- SVM_cluster_B serves data on cluster B.

It is a sync-source SVM that represents configuration and data to SVM_cluster_B-mc located on cluster A.

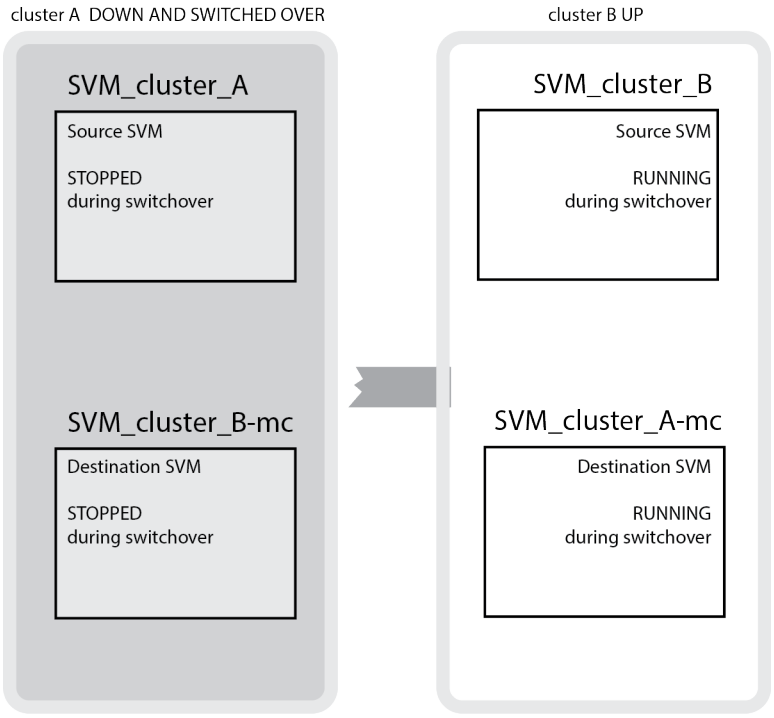
- SVM_cluster_B-mc is a sync-destination SVM that is stopped during normal, healthy operation of the MetroCluster configuration.

  In a successful switchover from cluster B to cluster A, SVM_cluster_B is stopped and SVM_cluster_B-mc is activated and begins serving data from cluster A.
- SVM_cluster_A-mc is a sync-destination SVM that is stopped during normal, healthy operation of the MetroCluster configuration.

  In a successful switchover from cluster A to cluster B, SVM_cluster_A is stopped and SVM_cluster_A-mc is activated and begins serving data from cluster B.



If a switchover occurs, the remote plex on the surviving cluster comes online and the secondary SVM begins serving the data.



The availability of remote plexes after switchover depends on the MetroCluster configuration type:

- For MetroCluster FC configurations, after switchover, both local and remote plexes remain online if the disaster site storage is accessible via the ISLs.
  If the ISLs have failed and the disaster site storage is not available, the sync-destination SVM begins serving data from the surviving site.
- For MetroCluster IP configurations the availability of the remote plexes depends on the ONTAP version:
  - Starting with ONTAP 9.5, both local and remote plexes remain online if the disaster site nodes remain booted up.
  - Prior to ONTAP 9.5, storage is available only from local plex on the surviving site.
    The sync-destination SVM begins serving data from the surviving site.

Related information

System administration

How MetroCluster configurations use SyncMirror to provide data redundancy

Mirrored aggregates using SyncMirror functionality provide data redundancy and contain the volumes owned by the source and destination storage virtual machine (SVM). Data is replicated into disk pools on the partner cluster. Unmirrored aggregates are also supported.

Attention:

The following table shows the state (online or offline) of an unmirrored aggregate after a switchover:

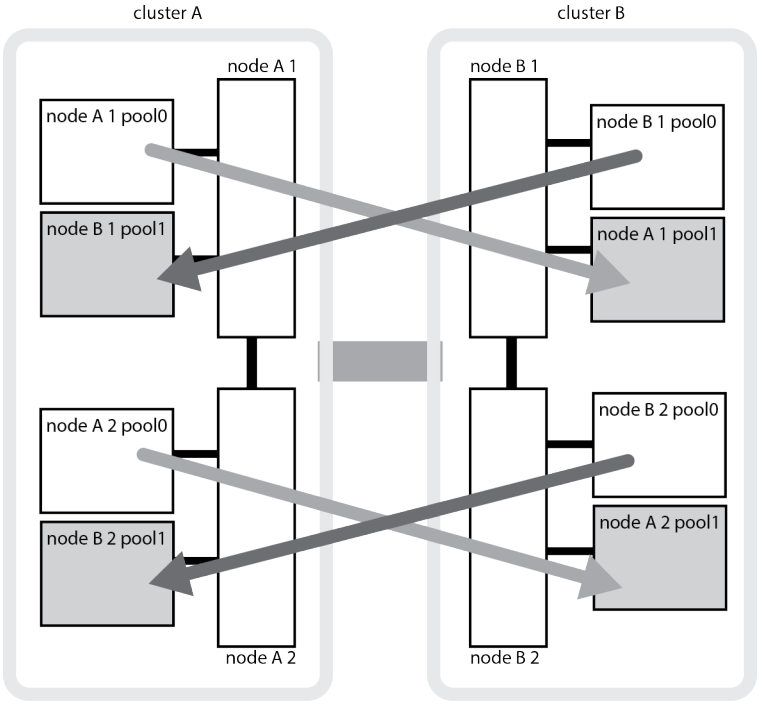
Type of switchover	State
Negotiated switchover (NSO)	Online
Automatic unplanned switchover (AUSO)	Online

Type of switchover	State
Unplanned switchover (USO)	• If storage is not available: Offline • If storage is available: Online

Note: After a switchover, if the unmirrored aggregate is at the DR partner node and there is an inter-switch link (ISL) failure, then that local node might fail.

The following illustration shows how disk pools are mirrored between the partner clusters. Data in local plexes (in pool0) is replicated to remote plexes (in pool1).

Attention: If hybrid aggregates are used, performance degradation can occur after a SyncMirror plex has failed due to the solid-state disk (SSD) layer filling up.

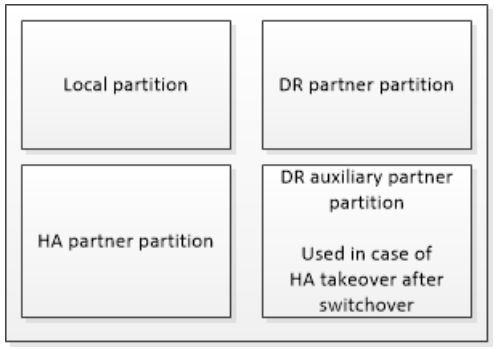
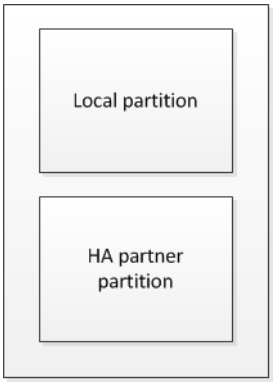


How NVRAM or NVMEM cache mirroring and dynamic mirroring work in MetroCluster configurations

The nonvolatile memory (NVRAM or NVMEM, depending on the platform model) in the storage controllers is mirrored both locally to a local HA partner and remotely to a remote disaster recovery (DR) partner on the partner site. In the event of a local failover or switchover, this configuration enables data in the nonvolatile cache to be preserved.

In an HA pair that is not part of a MetroCluster configuration, each storage controller maintains two nonvolatile cache partitions: one for itself and one for its HA partner.

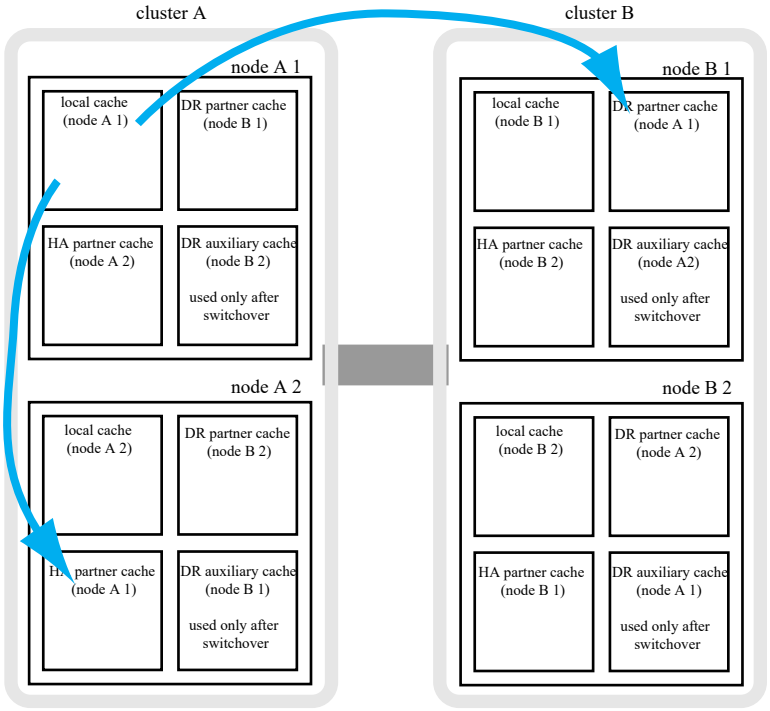
In a four-node MetroCluster configuration, the nonvolatile cache of each storage controller is divided into four partitions. In a two-node MetroCluster configuration, the HA partner partition and DR auxiliary partition are not used, because the storage controllers are not configured as an HA pair.

Nonvolatile caches for a storage controller	
In a MetroCluster configuration	**In a non-MetroCluster HA pair**
	

The nonvolatile caches store the following content:

- The *local partition* holds data that the storage controller has not yet written to disk.
- The *HA partner partition* holds a copy of the local cache of the storage controller's HA partner.
  In a two-node MetroCluster configuration, there is no HA partner partition because the storage controllers are not configured as an HA pair.
- The *DR partner partition* holds a copy of the local cache of the storage controller's DR partner.
  The DR partner is a node in the partner cluster that is paired with the local node.
- The *DR auxiliary partner partition* holds a copy of the local cache of the storage controller's DR auxiliary partner.
  The DR auxiliary partner is the HA partner of the local node's DR partner. This cache is needed if there is an HA takeover (either when the configuration is in normal operation or after a MetroCluster switchover).
  In a two-node MetroCluster configuration, there is no DR auxiliary partner partition because the storage controllers are not configured as an HA pair.

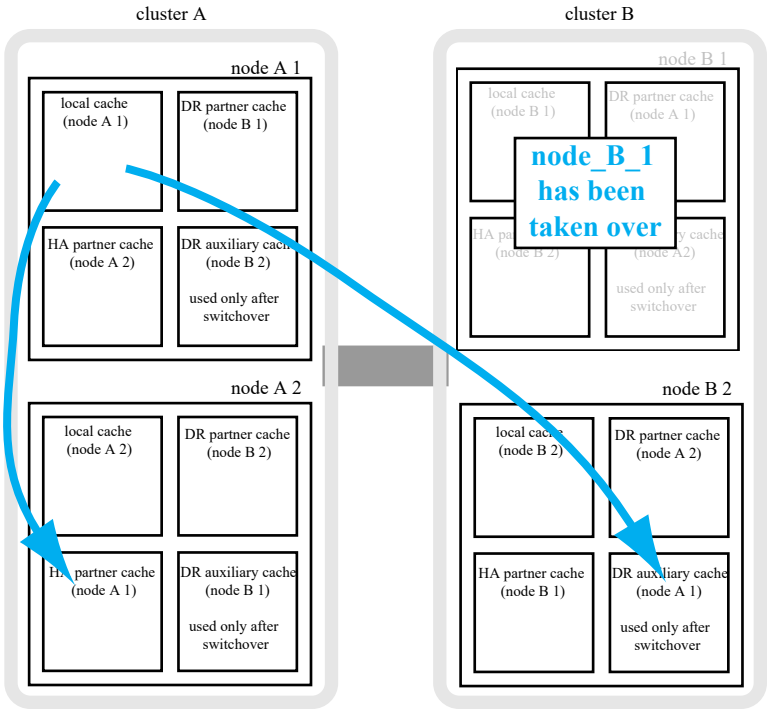
For example, the local cache of a node (node_A_1) is mirrored both locally and remotely at the MetroCluster sites. The following illustration shows that the local cache of node_A_1 is mirrored to the HA partner (node_A_2) and DR partner (node_B_1):



Dynamic mirroring in event of a local HA takeover

If a local HA takeover occurs in a four-node MetroCluster configuration, the taken-over node can no longer act as a mirror for its DR partner. To allow DR mirroring to continue, the mirroring automatically switches to the DR auxiliary partner. After a successful giveback, mirroring automatically returns to the DR partner.

For example, node_B_1 fails and is taken over by node_B_2. The local cache of node_A_1 can no longer be mirrored to node_B_1. The mirroring switches to the DR auxiliary partner, node_B_2.



Types of disasters and recovery methods

You need to be familiar with different types of failures and disasters so that you can use the MetroCluster configuration to respond appropriately.

- Single-node failure

  A single component in the local HA pair fails.

  In a four-node MetroCluster configuration, this failure might lead to an automatic or a negotiated takeover of the impaired node, depending on the component that failed. Data recovery is described in the *High Availability Configuration Guide*.

  In a two-node MetroCluster configuration, this failure leads to an automatic unplanned switchover (AUSO).
- Site-wide controller failure

  All controller modules fail at a site because of loss of power, replacement of equipment, or disaster. Typically, MetroCluster configurations cannot differentiate between failures and disasters. However, witness software, such as the MetroCluster Tiebreaker software, can differentiate between them. A site-wide controller failure condition can lead to an automatic switchover if Inter-Switch Link (ISL) links and switches are up and the storage is accessible.

  The *High-Availability Configuration Guide* has more information about how to recover from site-wide controller failures that do not include controller failures, as well as failures that include of one or more controllers.
- ISL failure

  The links between the sites fail. The MetroCluster configuration takes no action. Each node continues to serve data normally, but the mirrors are not written to the respective disaster recovery sites because access to them is lost.
- Multiple sequential failures

  Multiple components fail in a sequence. For example, a controller module, a switch fabric, and a shelf fail in a sequence and result in a storage failover, fabric redundancy, and SyncMirror sequentially protecting against downtime and data loss.

The following table shows failure types, and the corresponding disaster recovery (DR) mechanism and recovery method:

Note: AUSO (automatic unscheduled switchover) is not supported on MetroCluster IP configurations.

Failure type	DR mechanism		Summary of recovery method	
	Four-node configuration	**Two-node configuration**	**Four-node configuration**	**Two-node configuration**
Single-node failure	Local HA failover	AUSO	Not required if automatic failover and giveback is enabled.	After the node is restored, manual healing and switchback using the `metrocluster heal -phase aggregates`, `metrocluster heal -phase root-aggregates`, and `metrocluster switchback` commands is required. **Note:** The `metrocluster heal` commands are not required on MetroCluster IP configurations running ONTAP 9.5.
Site failure	MetroCluster switchover		After the node is restored, manual healing and switchback using the `metrocluster healing` and `metrocluster switchback` commands is required. **Note:** The `metrocluster heal` commands are not required on MetroCluster IP configurations running ONTAP 9.5.	
Site-wide controller failure	AUSO Only if the storage at the disaster site is accessible.	AUSO (same as single-node failure)		
Multiple sequential failures	Local HA failover followed by MetroCluster forced switchover using the `metrocluster switchover -forced-on-disaster` command. **Note:** Depending on the component that failed, a forced switchover might not be required.	MetroCluster forced switchover using the `metrocluster switchover -forced-on-disaster` command.		
ISL failure	No MetroCluster switchover; the two clusters independently serve their data		Not required for this type of failure. After you restore connectivity, the storage resynchronizes automatically.	

Related tasks

Performing a forced switchover after a disaster on page 38

An administrator, or the MetroCluster Tiebreaker software if it is configured, must determine that a disaster has occurred and perform the MetroCluster switchover. In either case, there are steps

you must perform on both the disaster cluster and the surviving cluster after the switchover to ensure safe and continued data service.

Related information

High-availability configuration

How an eight-node or four-node MetroCluster configuration provides nondisruptive operations

In the case of an issue limited to a single node, a failover and giveback within the local HA pair provides continued nondisruptive operation. In this case, the MetroCluster configuration does not require a switchover to the remote site.

Because the eight-node or four-node MetroCluster configuration consists of one or more HA pair at each site, each site can withstand local failures and perform nondisruptive operations without requiring a switchover to the partner site. The operation of the HA pair is the same as HA pairs in non-MetroCluster configurations.

For four-node and eight-node MetroCluster configurations, node failures due to panic or power loss can cause an automatic switchover.

High-availability configuration

If a second failure occurs after a local failover, the MetroCluster switchover event provides continued nondisruptive operations. Similarly, after a switchover operation, in the event of a second failure in one of the surviving nodes, a local failover event provides continued nondisruptive operations. In this case, the single surviving node serves data for the other three nodes in the DR group.

Switchover and switchback during MetroCluster transition

MetroCluster FC-to-IP transition involves adding MetroCluster IP nodes and IP switches to an existing MetroCluster FC configuration, and then retiring the MetroCluster FC nodes. Depending on the stage of the transition process, the MetroCluster switchover, healing, and switchback operations use different workflows.

See *Switchover, healing, and switchback operations during transition*.

Consequences of local failover after switchover

If a MetroCluster switchover occurs, and then an issue arises at the surviving site, a local failover can provide continued, nondisruptive operation. However, the system is at risk because it is no longer in a redundant configuration.

If a local failover occurs after a switchover has occurred, a single controller serves data for all storage systems in the MetroCluster configuration, leading to possible resource issues, and is vulnerable to additional failures.

How a two-node MetroCluster configuration provides nondisruptive operations

If one of the two sites has an issue due to panic, the MetroCluster switchover provides continued nondisruptive operation. If the power loss impacts both the node and the storage, then the switchover is not automatic and there is a disruption until the `metrocluster switchover` command is issued.

Because all storage is mirrored, a switchover operation can be used to provide nondisruptive resiliency in case of a site failure similar to that found in a storage failover in an HA pair for a node failure.

For two-node configurations, the same events that trigger an automatic storage failover in an HA pair trigger an automatic unplanned switchover (AUSO). This means that a two-node MetroCluster configuration has the same level of protection as an HA pair.

Related concepts

Automatic unplanned switchover in MetroCluster FC configurations on page 19
In MetroCluster FC configurations, certain scenarios can trigger an automatic unplanned switchover (AUSO) in the event of a site-wide controller failure to provide nondisruptive operations. AUSO can be disabled if desired.

Overview of the switchover process

The MetroCluster switchover operation enables immediate resumption of services following a disaster by moving storage and client access from the source cluster to the remote site. You must be aware of what changes to expect and which actions you need to perform if a switchover occurs.

During a switchover operation, the system takes the following actions:

- Ownership of the disks that belong to the disaster site is changed to the disaster recovery (DR) partner.

  This is similar to the case of a local failover in a high-availability (HA) pair, in which ownership of the disks belonging to the partner that is down is changed to the healthy partner.
- The surviving plexes that are located on the surviving site but belong to the nodes in the disaster cluster are brought online on the cluster at the surviving site.
- The sync-source storage virtual machine (SVM) that belongs to the disaster site is brought down only during a negotiated switchover.

  Note: This is applicable only to a negotiated switchover.
- The sync-destination SVM belonging to the disaster site is brought up.

While being switched over, the root aggregates of the DR partner are not brought online.

The `metrocluster switchover` command switches over the nodes in all DR groups in the MetroCluster configuration. For example, in an eight-node MetroCluster configuration, it switches over the nodes in both DR groups.

If you are switching over only services to the remote site, you should perform a negotiated switchover without fencing the site. If storage or equipment is unreliable, you should fence the disaster site, and then perform an unplanned switchover. Fencing prevents RAID reconstructions when the disks power up in a staggered manner.

Note: This procedure should be only used if the other site is stable and not intended to be taken offline.

Availability of commands during switchover

The following table shows the availability of commands during switchover:

Command	Availability
`storage aggregate create`	You can create an aggregate: • If it is owned by a node that is part of the surviving cluster You cannot create an aggregate: • For a node at the disaster site • For a node that is part of the surviving cluster
`storage aggregate delete`	You can delete a data aggregate.

Command	Availability
`storage aggregate mirror`	You can create a plex for a non-mirrored aggregate.
`storage aggregate plex delete`	You can delete a plex for a mirrored aggregate.
`vserver create`	You can create an SVM: • If its root volume resides in a data aggregate owned by the surviving cluster You cannot create an SVM: • If its root volume resides in a data aggregate owned by the disaster-site cluster
`vserver delete`	You can delete both sync-source and sync-destination SVMs.
`network interface create -lif`	You can create a data SVM LIF for both sync-source and sync-destination SVMs.
`network interface delete -lif`	You can delete a data SVM LIF for both sync-source and sync-destination SVMs.
`lif create`	You can create LIFs.
`lif delete`	You can delete LIFs.
`volume create`	You can create a volume for both sync-source and sync-destination SVMs. • For a sync-source SVM, the volume must reside in a data aggregate owned by the surviving cluster • For a sync-destination SVM, the volume must reside in a data aggregate owned by the disaster-site cluster
`volume delete`	You can delete a volume for both sync-source and sync-destination SVMs.
`volume move`	You can move a volume for both sync-source and sync-destination SVMs. • For a sync-source SVM, the surviving cluster must own the destination aggregate • For a sync-destination SVM, the disaster-site cluster must own the destination aggregate
`snapmirror break`	You can break a SnapMirror relationship between a source and destination endpoint of a data protection mirror.

Differences in switchover between MetroCluster FC and IP configurations

In MetroCluster IP configurations, because the remote disks are accessed through the remote DR partner nodes acting as iSCSI targets, the remote disks are not accessible when the remote nodes are taken down in a switchover operation. This results in differences with MetroCluster FC configurations:

- Mirrored aggregates that are owned by the local cluster become degraded.
- Mirrored aggregates that were switched over from the remote cluster become degraded.

Note: When unmirrored aggregates are supported on a MetroCluster IP configuration, the unmirrored aggregates that are not switched over from the remote cluster are not accessible.

Related tasks

Fencing off the disaster site on page 38

After the disaster, if the disaster site nodes must be replaced, you must halt them to prevent the site from resuming service. Otherwise, you risk the possibility of data corruption if clients start accessing the nodes before the replacement procedure is completed.

Disk ownership changes during HA takeover and MetroCluster switchover in a four-node MetroCluster configuration

The ownership of disks temporarily changes automatically during high availability and MetroCluster operations. It is helpful to know how the system tracks which node owns which disks.

In ONTAP, a controller module's unique system ID (obtained from a node's NVRAM card or NVMEM board) is used to identify which node owns a specific disk. Depending on the HA or DR state of the system, the ownership of the disk might temporarily change. If the ownership changes because of an HA takeover or a DR switchover, the system records which node is the original (called "home") owner of the disk, so that it can return the ownership after HA giveback or DR switchback. The system uses the following fields to track disk ownership:

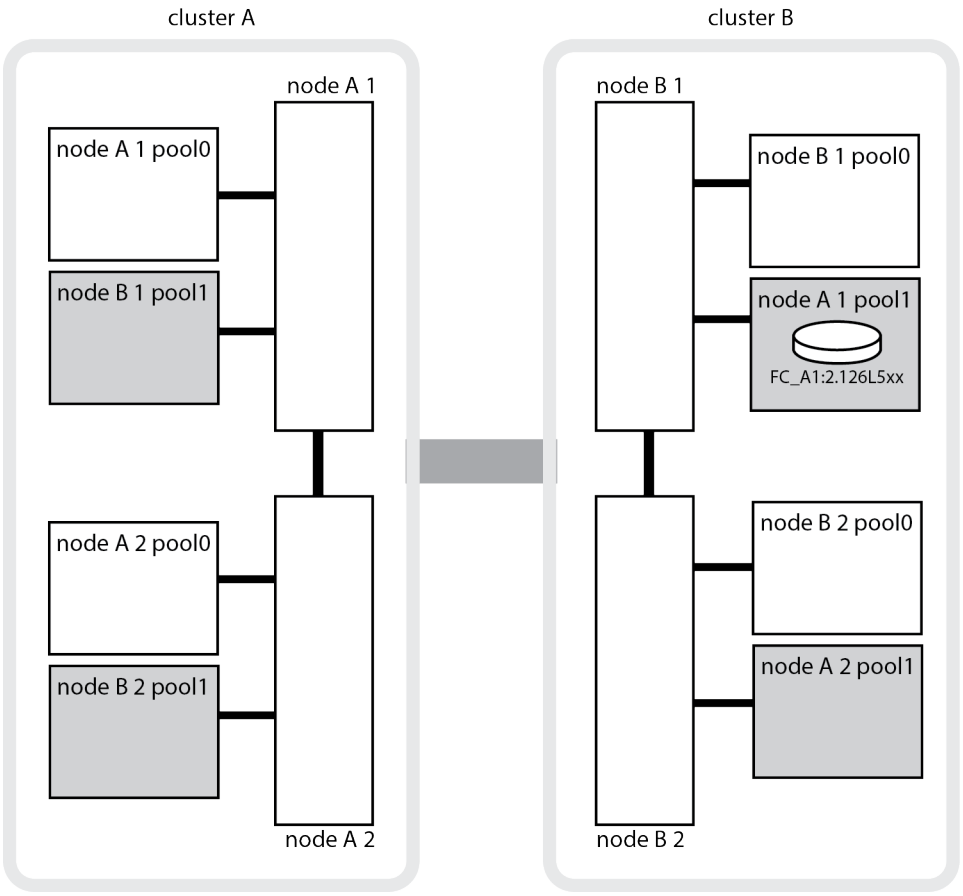
- `Owner`
- `Home owner`
- `DR Home owner`

In the MetroCluster configuration, in the event of a switchover, a node can take ownership of an aggregate originally owned by nodes in the partner cluster. Such aggregates are referred to as *cluster-foreign aggregates*. The distinguishing feature of a cluster-foreign aggregate is that it is an aggregate not currently known to the cluster, and so the `DR Home owner` field is used to show that it is owned by a node from the partner cluster. A traditional foreign aggregate within an HA pair is identified by `Owner` and `Home owner` values being different, but the `Owner` and `Home owner` values are the same for a cluster-foreign aggregate; thus, you can identify a cluster-foreign aggregate by the `DR Home owner` value.

As the state of the system changes, the values of the fields change, as shown in the following table:

	Value during...			
Field	**Normal operation**	**Local HA takeover**	**MetroCluster switchover**	**Takeover during switchover**
`Owner`	ID of the node that has access to the disk.	ID of the HA partner, which temporarily has access to the disk.	ID of the DR partner, which temporarily has access to the disk.	ID of the DR auxiliary partner, which temporarily has access to the disk.
`Home owner`	ID of the original owner of the disk within the HA pair.	ID of the original owner of the disk within the HA pair.	ID of the DR partner, which is the `Home owner` in the HA pair during the switchover.	ID of the DR partner, which is the `Home owner` in the HA pair during the switchover.
`DR Home owner`	Empty	Empty	ID of the original owner of the disk within the MetroCluster configuration.	ID of the original owner of the disk within the MetroCluster configuration.

The following illustration and table provide an example of how ownership changes, for a disk in node_A_1's disk pool1, physically located in cluster_B.



MetroCluster state	Owner	Home owner	DR Home owner	Notes
Normal with all nodes fully operational.	node_A_1	node_A_1	not applicable	
Local HA takeover, node_A_2 has taken over disks belonging to its HA partner node_A_1.	node_A_2	node_A_1	not applicable	
DR switchover, node_B_1 has taken over disks belong to its DR partner, node_A_1.	node_B_1	node_B_1	node_A_1	The original home node ID is moved to the `DR Home owner` field. After aggregate switchback or healing, ownership goes back to node_A_1.
In DR switchover and local HA takeover (double failure), node_B_2 has taken over disks belonging to its HA node_B_1.	node_B_2	node_B_1	node_A_1	After giveback, ownership goes back to node_B_1. After switchback or healing, ownership goes back to node_A_1.
After HA giveback and DR switchback, all nodes fully operational.	node_A_1	node_A_1	not applicable	

Considerations when using unmirrored aggregates

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

Considerations for unmirrored aggregates when doing maintenance requiring power shutdown

If you are performing negotiated switchover for maintenance reasons requiring site-wide power shutdown, you should first manually take offline any unmirrored aggregates owned by the disaster site.

If you do not, nodes at the surviving site might go down due to multi-disk panics. This could occur if switched-over unmirrored aggregates go offline or are missing because of the loss of connectivity to storage at the disaster site due to the power shutdown or a loss of ISLs.

Considerations for unmirrored aggregates and hierarchical namespaces

If you are using hierarchical namespaces, you should configure the junction path so that all of the volumes in that path are either on mirrored aggregates only or on unmirrored aggregates only. Configuring a mix of unmirrored and mirrored aggregates in the junction path might prevent access to the unmirrored aggregates after the switchover operation.

Considerations for unmirrored aggregates and CRS metadata volume and data SVM root volumes

The configuration replication service (CRS) metadata volume and data SVM root volumes must be on a mirrored aggregate. You cannot move these volumes to unmirrored aggregate. If they are on unmirrored aggregate, negotiated switchover and switchback operations are vetoed. The `metrocluster check` command provides a warning if this is the case.

Considerations for unmirrored aggregates and SVMs

SVMs should be configured on mirrored aggregates only or on unmirrored aggregates only. Configuring a mix of unmirrored and mirrored aggregates can result in a switchover operation that exceeds 120 seconds and result in a data outage if the unmirrored aggregates do not come online.

Considerations for unmirrored aggregates and SAN

A LUN should not be located on an unmirrored aggregate. Configuring a LUN on an unmirrored aggregate can result in a switchover operation that exceeds 120 seconds and a data outage.

Automatic unplanned switchover in MetroCluster FC configurations

In MetroCluster FC configurations, certain scenarios can trigger an automatic unplanned switchover (AUSO) in the event of a site-wide controller failure to provide nondisruptive operations. AUSO can be disabled if desired.

Note: Automatic unplanned switchover is not supported in MetroCluster IP configurations.

In a MetroCluster FC configuration, an AUSO can be triggered if all nodes at a site are failed because of the following reasons:

- Power down
- Power loss
- Power panic

Note: In an eight-node MetroCluster FC configuration, you can set an option to trigger an AUSO if both nodes in an HA pair fail.

Because there is no local HA failover available in a two-node MetroCluster configuration, the system performs an AUSO to provide continued operation after a controller failure. This functionality is similar to the HA takeover capability in an HA pair. In a two-node MetroCluster configuration, an AUSO can be triggered in the following scenarios:

- Node power down
- Node power loss
- Node panic
- Node reboot

If an AUSO occurs, disk ownership for the impaired node's pool0 and pool1 disks is changed to the disaster recovery (DR) partner. This ownership change prevents the aggregates from going into a degraded state after the switchover.

After the automatic switchover, you must manually proceed through the healing and switchback operations to return the controller to normal operation.

Hardware-assisted AUSO in two-node MetroCluster configurations

In a two-node MetroCluster configuration, the controller module's service processor (SP) monitors the configuration. In some scenarios, the SP can detect a failure faster than the ONTAP software. In this case, the SP triggers AUSO. This feature is automatically enabled.

The SP sends and receives SNMP traffic to and from its DR partner to monitor its health.

Changing the AUSO setting

AUSO is set to **`auso-on-cluster-disaster`** by default. Its status can be viewed in the `metrocluster show` command.

You can disable AUSO with the `metrocluster modify -auto-switchover-failure-domain auto-disabled` command. This command prevents triggering AUSO in DR site-wide controller failure. It should be run on both the sites if you want to disable AUSO on both the sites.

AUSO can be reenabled with the `metrocluster modify -auto-switchover-failure-domain auso-on-cluster-disaster` command.

AUSO can also be set to **`auso-on-dr-group-disaster`**. This advance level command triggers AUSO on HA failover at one site. It should be run on both the sites with the `metrocluster modify -auto-switchover-failure-domain auso-on-dr-group-disaster` command.

The AUSO setting during switchover

When switchover occurs, the AUSO setting is disabled internally because if a site is in switchover, it cannot automatically switch over.

Recovering from AUSO

To recover from an AUSO, you perform the same steps as for a planned switchover.

Performing switchover for tests or maintenance on page 28

Mediator-assisted automatic unplanned switchover in MetroCluster IP configurations

In MetroCluster IP configurations, the system can use the ONTAP Mediator to detect failures and perform a Mediator-assisted automatic unplanned switchover (MAUSO).

Note: MAUSO is not supported in MetroCluster FC configurations.

The ONTAP Mediator provides mailbox LUNs for the MetroCluster IP nodes. These LUNs are colocated with the ONTAP Mediator, which runs on a Linux host physically separate from the MetroCluster sites.

The MetroCluster nodes use the mailbox information to determine if a MAUSO is required. MAUSO will not be initiated if the nonvolatile memory (NVRAM or NVMEM, depending on the platform model) in the storage controllers is not mirrored to the remote disaster recovery (DR) partner on the partner site

What happens during healing (MetroCluster FC configurations)

During healing in MetroCluster FC configurations, the resynchronization of mirrored aggregates occurs in a phased process that prepares the nodes at the repaired disaster site for switchback. It is a planned event, thereby giving you full control of each step to minimize downtime. Healing is a two-step process that occurs on the storage and controller components.

Data aggregate healing

After the problem at the disaster site is resolved, you start the storage healing phase:

1. Checks that all nodes are up and running at the surviving site.
2. Changes ownership of all the pool 0 disks at the disaster site, including root aggregates.

During this phase of healing, the RAID subsystem resynchronizes mirrored aggregates, and the WAFL subsystem replays the `nvsave` files of mirrored aggregates that had a failed pool 1 plex at the time of switchover.

If some source storage components failed, the command reports the errors at applicable levels: `Storage`, `Sanown`, or `RAID`.

If no errors are reported, the aggregates are successfully resynchronized. This process can sometimes take hours to complete.

Healing the data aggregates

Root aggregate healing

After the aggregates are synchronized, you start the controller healing phase by giving back the CFO aggregates and root aggregates to their respective DR partners.

Healing the root aggregates

What happens during healing (MetroCluster IP configurations)

During healing in MetroCluster IP configurations, the resynchronization of mirrored aggregates occurs in a phased process that prepares the nodes at the repaired disaster site for switchback. It is a planned event, thereby giving you full control of each step to minimize downtime. Healing is a two-step process that occurs on the storage and controller components.

Differences with MetroCluster FC configurations

In MetroCluster IP configurations, you must boot the nodes in the disaster site cluster before the healing operation is performed.

The nodes in the disaster site cluster must be running so that the remote iSCSI disks can be accessed when aggregates are resynchronized.

If the disaster site nodes are not running, the healing operation fails because the disaster node cannot perform the disk ownership changes needed.

Data aggregate healing

After the problem at the disaster site is resolved, you start the storage healing phase:

1. Checks that all nodes are up and running at the surviving site.
2. Changes ownership of all the pool 0 disks at the disaster site, including root aggregates.

During this phase of healing, the RAID subsystem resynchronizes mirrored aggregates, and the WAFL subsystem replays the `nvsave` files of mirrored aggregates that had a failed pool 1 plex at the time of switchover.

If some source storage components failed, the command reports the errors at applicable levels: `Storage`, `Sanown`, or `RAID`.

If no errors are reported, the aggregates are successfully resynchronized. This process can sometimes take hours to complete.

Healing the data aggregates

Root aggregate healing

After the aggregates are synchronized, you perform the root aggregate healing phase. In MetroCluster IP configurations, this phase confirms that aggregates have been healed.

Healing the root aggregates

Automatic healing of aggregates on MetroCluster IP configurations after switchover

Starting with ONTAP 9.5, healing is automated during negotiated switchover operations on MetroCluster IP configurations. Starting with ONTAP 9.6, automated healing after unscheduled switchover is supported. This removes the requirement to issue the `metrocluster heal` commands.

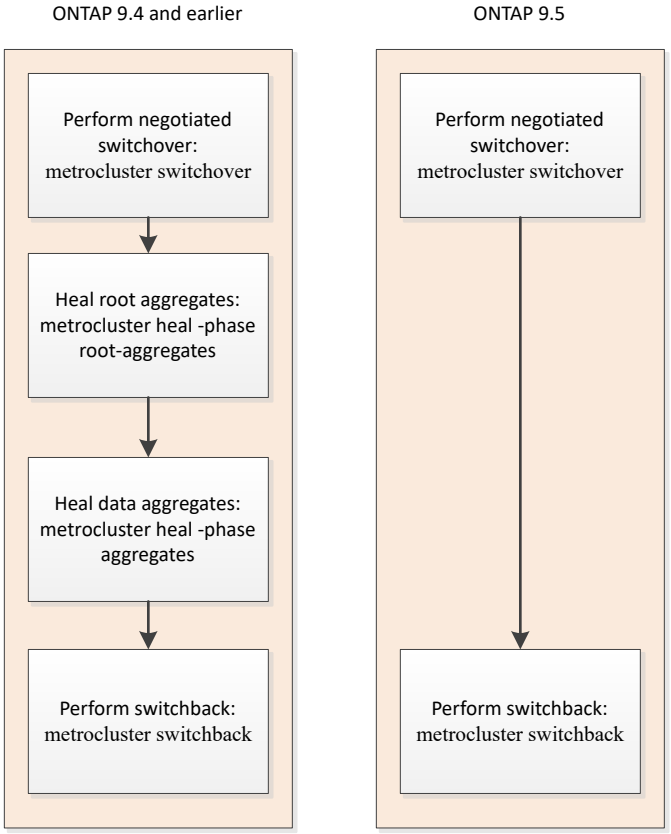
Automatic healing after negotiated switchover (starting with ONTAP 9.5)

After performing a negotiated switchover (a switchover command issued without the `-forced-on-disaster true` option), the automatic healing functionality simplifies the steps required to return the system to normal operation. On systems with automatic healing, the following occurs after the switchover:

- The disaster site nodes remain up.
  Because they are in switchover state, they are not serving data from their local mirrored plexes.
- The disaster site nodes are moved to the `Waiting for switchback` state.
  You can confirm the status of the disaster site nodes by using the `metrocluster operation show` command.
- You can perform the switchback operation without issuing the healing commands.

This feature applies to MetroCluster IP configurations running ONTAP 9.5 and later. It does not apply to MetroCluster FC configurations.

The manual healing commands are still required on MetroCluster IP configurations running ONTAP 9.4 and earlier.



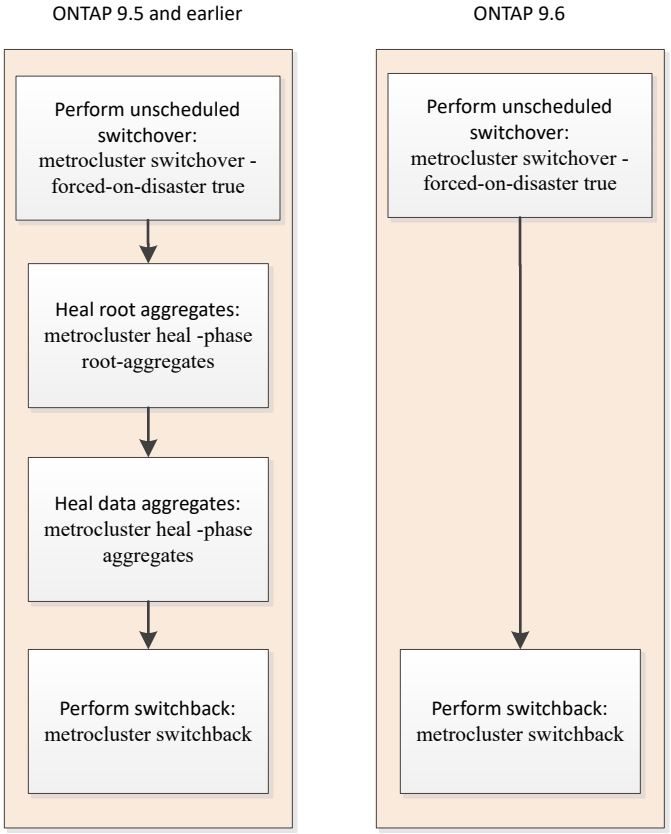
Automatic healing after unscheduled switchover (starting with ONTAP 9.6)

Automatic healing after an unscheduled switchover is supported on MetroCluster IP configurations starting with ONTAP 9.6. An unscheduled switchover is one in which in you issue the switchover command with the `-forced-on-disaster true` option.

Automatic healing after an unscheduled switchover is not supported on MetroCluster FC configurations, and the manual healing commands are still required after unscheduled switchover on MetroCluster IP configurations running ONTAP 9.5 and earlier.

On systems running ONTAP 9.6 and later, the following occurs after the unscheduled switchover:

- Depending on the extent of the disaster, the disaster site nodes can be down.
  Because they are in switchover state, they are not serving data from their local mirrored plexes, even if they are powered up.
- If the disaster sites were down, when booted up, the disaster site nodes are moved to the `Waiting for switchback` state.
  If the disaster sites remained up, they are immediately moved to the `Waiting for switchback` state.
- The healing operations are performed automatically.
  You can confirm the status of the disaster site nodes, and that healing operations succeeded, by using the `metrocluster operation show` command.



If automatic healing fails

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 9.6. You can use the `metrocluster operation show` and `metrocluster operation history show -instance` commands to monitor the status of healing and determine the cause of a failure.

Creating SVMs for a MetroCluster configuration

You can create SVMs for a MetroCluster configuration to provide synchronous disaster recovery and high availability of data on clusters that are set up for a MetroCluster configuration.

Before you begin

- The two clusters must be in a MetroCluster configuration.
- Aggregates must be available and online in both clusters.
- If required, IPspaces with the same names must be created on both clusters.
- If one of the clusters forming the MetroCluster configuration is rebooted without utilizing a switchover, then the sync-source SVMs might come online as `stopped` rather than `started`.

About this task

When you create an SVM on one of the clusters in a MetroCluster configuration, the SVM is created as the source SVM, and the partner SVM is automatically created with the same name but with the "-mc" suffix on the partner cluster. If the SVM name contains a period, the "-mc" suffix is applied prior to the first period, for example, SVM-MC.DNS.NAME.

In a MetroCluster configuration, you can create 64 SVMs on a cluster. A MetroCluster configuration supports 128 SVMs.

Steps

1. Use the `vserver create` command.

   The following example shows the SVM with the subtype **`sync-source`** on the local site and the SVM with the subtype **`sync-destination`** on the partner site:

   ```
   cluster_A::>vserver create -vserver vs4 -rootvolume vs4_root -aggregate aggr1
   -rootvolume-security-style mixed
   [Job 196] Job succeeded:
   Vserver creation completed
   ```

   The SVM vs4 is created on the local site and the SVM vs4-mc is created on the partner site.
2. View the newly created SVMs.
   - On the local cluster, verify the configuration state of SVMs:

     `metrocluster vserver show`

     The following example shows the partner SVMs and their configuration state:

     ```
     cluster_A::> metrocluster vserver show

                          Partner    Configuration
     Cluster     Vserver  Vserver    State
     ---------  --------  --------- -----------------
     cluster_A   vs4       vs4-mc     healthy
     cluster_B   vs1       vs1-mc     healthy
     ```

   - From the local and partner clusters, verify the state of the newly configured SVMs:

     `vserver show command`

     The following example displays the administrative and operational states of the SVMs:

     ```
     cluster_A::> vserver show

                                  Admin   Operational Root
     Vserver Type  Subtype        State   State       Volume     Aggregate
     ------- ----- -------       ------- --------    ---------- ----------
     vs4     data  sync-source   running   running    vs4_root   aggr1

     cluster_B::> vserver show

                                     Admin   Operational  Root
     Vserver Type  Subtype          State   State        Volume      Aggregate
     ------- ----- -------          ------  ---------    ----------- ----------
     vs4-mc  data  sync-destination running stopped      vs4_root    aggr1
     ```

   SVM creation might fail if any intermediate operations, such as root volume creation, fail and the SVM is in the `initializing` state. You must delete the SVM and re-create it.

Result

The SVMs for the MetroCluster configuration are created with a root volume size of 1 GB. The sync-source SVM is in the `running` state, and the sync-destination SVM is in the `stopped` state.

What happens during a switchback

After the disaster site has recovered and aggregates have healed, the MetroCluster switchback process returns storage and client access from the disaster recovery site to the home cluster.

The `metrocluster switchback` command returns the primary site to full, normal MetroCluster operation. Any configuration changes are propagated to the original SVMs. Data server operation is then returned to the sync-source SVMs on the disaster site and the sync-dest SVMs that had been operating on the surviving site are deactivated.

If SVMs were deleted on the surviving site while the MetroCluster configuration was in switchover state, the switchback process does the following:

- Deletes the corresponding SVMs on the partner site (the former disaster site).
- Deletes any peering relationships of the deleted SVMs.

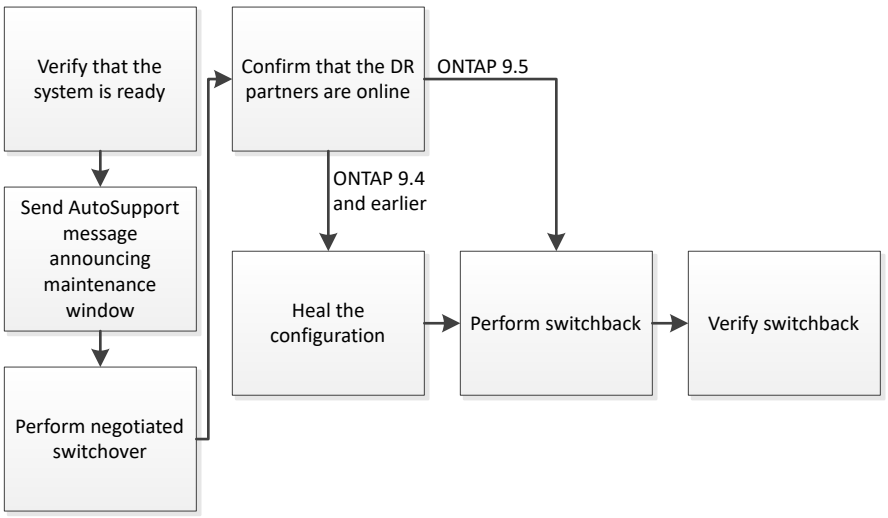
Performing switchover and switchback operations In MetroCluster IP configurations with ONTAP System Manager

Starting with ONTAP 9.6, switchover and switchback operations can be performed with ONTAP System Manager.

Performing switchover for tests or maintenance

If you want to test the MetroCluster functionality or to perform planned maintenance, you can perform a negotiated switchover in which one cluster is cleanly switched over to the partner cluster. You can then heal and switch back the configuration.

About this task



Steps

1. *Verifying that your system is ready for a switchover* on page 28
2. *Sending a custom AutoSupport message prior to negotiated switchover* on page 29
3. *Performing a negotiated switchover* on page 30
4. *Output for the storage aggregate plex show command is indeterminate after a MetroCluster switchover* on page 31
5. *Confirming that the DR partners have come online* on page 31
6. *Healing the configuration* on page 32
7. *Performing a switchback* on page 35
8. *Verifying a successful switchback* on page 37

Verifying that your system is ready for a switchover

You can use the `-simulate` option to preview the results of a switchover operation. A verification check gives you a way to verify that most of the preconditions for a successful run are met before you start the operation.

Steps

1. Set the privilege level to advanced:

   `set -privilege advanced`
2. Simulate a switchover operation:

   `metrocluster switchover -simulate`
3. Review the output that is returned.

The output shows whether any vetoes would prevent a switchover operation. Every time you perform a MetroCluster operation, you must verify a set of criteria for the success of the operation. A "veto" is a mechanism to prohibit the operation if one or more of the criteria are not fulfilled. There are two types of veto: a "soft" veto and a "hard" veto. You can override a soft veto, but not a hard veto. For example, to perform a negotiated switchover in a four-node MetroCluster configuration, one criterion is that all of the nodes are up and healthy. Suppose one node is down and was taken over by its HA partner. The switchover operation will be hard vetoed because it is a hard criterion that all of the nodes must be up and healthy. Because this is a hard veto, you cannot override the veto.

Attention: It is best not to override any veto.

Example: Verification results

The following example shows the errors that are encountered in a simulation of a switchover operation:

```
cluster4::*> metrocluster switchover -simulate

[Job 126] Preparing the cluster for the switchover operation...
[Job 126] Job failed: Failed to prepare the cluster for the switchover
operation. Use the "metrocluster operation show" command to view detailed error
information. Resolve the errors, then try the command again.
```

Note: Negotiated switchover and switchback will fail until you replace all of the failed disks. You can perform disaster recovery after you replace the failed disks. If you want to ignore the warning for failed disks, you can add a soft veto for the negotiated switchover and switchback.

Sending a custom AutoSupport message prior to negotiated switchover

Before performing a negotiated switchover, you should issue an AutoSupport message to notify NetApp technical support that maintenance is underway. The negotiated switchover might result in plex or MetroCluster operation failures that trigger AutoSupport messages. Informing technical support that maintenance is underway prevents them from opening a case on the assumption that a disruption has occurred.

About this task

This task must be performed on each MetroCluster site.

Steps

1. Log in to the cluster at Site_A.
2. Invoke an AutoSupport message indicating the start of the maintenance:

   `system node autosupport invoke -node * -type all -message MAINT=`*`maintenance-window-in-hours`*

   `maintenance-window-in-hours` specifies the length of the maintenance window and can be a maximum of 72 hours. If the maintenance is completed before the time has elapsed, you can issue a command to indicating that the maintenance period has ended:

   `system node autosupport invoke -node * -type all -message MAINT=end`
3. Repeat this step on the partner site.

Performing a negotiated switchover

A negotiated switchover cleanly shuts down processes on the partner site, and then switches over operations from the partner site. You can use a negotiated switchover to perform maintenance on a MetroCluster site or to test the switchover functionality.

Before you begin

- All previous configuration changes must be completed before performing a switchback operation.
  This is to avoid competition with the negotiated switchover or switchback operation.
- Any nodes that were previously down must be booted and in cluster quorum.
  The *System Administration Reference* has more information about cluster quorum in the "Understanding quorum and epsilon" section.
  System administration
- The cluster peering network must be available from both sites.
- All of the nodes in the MetroCluster configuration must be running the same version of ONTAP software.
- The option `replication.create_data_protection_rels.enable` must be set to **`ON`** on both of the sites in a MetroCluster configuration before creating a new SnapMirror relationship.
- For a two-node MetroCluster configuration, a new SnapMirror relationship should not be created during an upgrade when there are mismatched versions of ONTAP between the sites.
- For a four-node MetroCluster configuration, the mismatched versions of ONTAP between the sites are not supported.

About this task

The recovering site can take a few hours to be able to perform the switchback operation.

The `metrocluster switchover` command switches over the nodes in all DR groups in the MetroCluster configuration. For example, in an eight-node MetroCluster configuration, it switches over the nodes in both DR groups.

While preparing for and executing a negotiated switchover, you must not make configuration changes to either cluster or perform any takeover or giveback operations.

For MetroCluster FC configurations:

- Mirrored aggregates will remain in normal state if the remote storage is accessible.
- Mirrored aggregates will become degraded after the negotiated switchover if access to the remote storage is lost.
- Unmirrored aggregates that are located at the disaster site will become unavailable if access to the remote storage is lost. This might lead to a controller outage.

For MetroCluster IP configurations:

- For ONTAP 9.4 and earlier:
  - Mirrored aggregates will become degraded after the negotiated switchover.
- For ONTAP 9.5 and later:
  - Mirrored aggregates will remain in normal state if the remote storage is accessible.
  - Mirrored aggregates will become degraded after the negotiated switchover if access to the remote storage is lost.
- For ONTAP 9.8 and later:
  - Unmirrored aggregates that are located at the disaster site will become unavailable if access to the remote storage is lost. This might lead to a controller outage.

Steps

1. Use the `metrocluster check run`, `metrocluster check show` and `metrocluster check config-replication show` commands to make sure no configuration updates are in progress or pending.
2. Enter the following command to implement the switchover:

   `metrocluster switchover`

   The operation can take several minutes to complete.

   Attention:
3. Monitor the completion of the switchover:

   `metrocluster operation show`

```
cluster_A::*> metrocluster operation show
  Operation: Switchover
 Start time: 10/4/2012 19:04:13
      State: in-progress
   End time: -
     Errors:

cluster_A::*> metrocluster operation show
  Operation: Switchover
 Start time: 10/4/2012 19:04:13
      State: successful
   End time: 10/4/2012 19:04:22
     Errors: -
```

4. Reestablish any SnapMirror or SnapVault configurations.

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.

Confirming that the DR partners have come online

After the switchover is complete, you should verify that the DR partners have taken ownership of the disks and the partner SVMs have come online.

Steps

1. Confirm that the aggregate disks have switched over to the disaster site:

   `storage disk show -fields owner,dr-home`

   In this example, the output shows that the switched over disks have the `dr-home` field set:

```
cluster_A::> storage disk show -fields owner,dr-home
disk                owner           dr-home
------------------- --------------- -------
1.11.0              node_A_1        node_B_1
1.11.1              node_A_1        node_B_1
1.11.2              node_A_1        node_B_1
1.11.3              node_A_1        node_B_1
1.11.4              node_A_1        node_B_1
1.11.5              node_A_1        node_B_1
1.11.6              node_A_1        node_B_1
```

```
1.11.7                node_A_1        node_B_1
1.11.8                node_A_1        node_B_1
```

2. Check that the aggregates were switched over by using the `storage aggregate show` command.
   In this example, the aggregates were switched over. The root aggregate (aggr0_b2) is in a degraded state. The data aggregate (b2_aggr2) is in a mirrored, normal state:

```
cluster_A::*> storage aggregate show

.
.
.
mcc1-b Switched Over Aggregates:
Aggregate     Size Available Used% State   #Vols  Nodes            RAID Status
--------- -------- --------- ----- ------- ------ ---------------- ------------
aggr0_b2    227.1GB   45.1GB    80% online       0 node_A_1         raid_dp,
                                                                    mirror
                                                                    degraded
b2_aggr1    227.1GB   200.3GB    20% online       0 node_A_1       raid_dp,
                                                                    mirrored
                                                                    normal
```

3. Confirm that the secondary SVMs have come online by using the `vserver show` command.
   In this example, the previously dormant sync-destination SVMs on the secondary site have been activated and have an `Admin State` of `running`:

```
cluster_A::*> vserver show
                                    Admin      Operational  Root                Name    Name
Vserver           Type  Subtype     State      State        Volume    Aggregate Service Mapping
-----------       ----- ----------  ---------- -----------  --------- --------- ------- -------
...
cluster_B-vs1b-mc data   sync-destination  running   running      vs1b_vol   aggr_b1   file    file
```

Healing the configuration

After a negotiated switchover operation, you must perform the healing operations in specific order to restore MetroCluster functionality. The procedure you use depends on the type of MetroCluster configuration you have.

About this task

On MetroCluster IP systems running ONTAP 9.5, healing is performed automatically, and you can skip these tasks.

Choices

- *Healing the configuration in a MetroCluster FC configuration* on page 32
- *Healing the configuration in a MetroCluster IP configuration (ONTAP 9.4 and earlier)* on page 34

Healing the configuration in a MetroCluster FC configuration

Following a switchover, you must perform the healing operations in specific order to restore MetroCluster functionality.

Before you begin

- Switchover must have been performed and the surviving site must be serving data.
- Nodes on the disaster site must be halted or remain powered off.
  They must not be fully booted during the healing process.
- Storage at the disaster site must be accessible (shelves are powered up, functional, and accessible).
- In fabric-attached MetroCluster configurations, inter-switch links (ISLs) must be up and operating.
- In four-node MetroCluster configurations, nodes in the surviving site must not be in HA failover state (all nodes must be up and running for each HA pair).

About this task

The healing operation must first be performed on the data aggregates, and then on the root aggregates.

Steps

1. *Healing the data aggregates after negotiated switchover* on page 33
2. *Healing the root aggregates after negotiated switchover* on page 34

Healing the data aggregates after negotiated switchover

You must heal the data aggregates after completing any maintenance or testing. This process resynchronizes the data aggregates and prepares the disaster site for normal operation. You must heal the data aggregates prior to healing the root aggregates.

About this task

All configuration updates in the remote cluster successfully replicate to the local cluster. You power up the storage on the disaster site as part of this procedure, but you do not and must not power up the controller modules on the disaster site.

Steps

1. Ensure that switchover has been completed by running the `metrocluster operation show` command.

```
controller_A_1::> metrocluster operation show
  Operation: switchover
      State: successful
 Start Time: 7/25/2014 20:01:48
   End Time: 7/25/2014 20:02:14
     Errors: -
```

2. Resynchronize the data aggregates by running the `metrocluster heal -phase aggregates` command from the surviving cluster.

```
controller_A_1::> metrocluster heal -phase aggregates
[Job 130] Job succeeded: Heal Aggregates is successful.
```

   If the healing is vetoed, you have the option of reissuing the `metrocluster heal` command with the `–override-vetoes` parameter. If you use this optional parameter, the system overrides any soft vetoes that prevent the healing operation.

3. Verify that the operation has been completed by running the `metrocluster operation show` command.

```
controller_A_1::> metrocluster operation show
    Operation: heal-aggregates
      State: successful
Start Time: 7/25/2014 18:45:55
   End Time: 7/25/2014 18:45:56
     Errors: -
```

4. Check the state of the aggregates by running the `storage aggregate show` command.

```
controller_A_1::> storage aggregate show
Aggregate     Size Available Used% State   #Vols  Nodes            RAID Status
--------- -------- --------- ----- ------- ------ ---------------- ------------
...
aggr_b2    227.1GB   227.1GB    0% online       0 mcc1-a2          raid_dp, mirrored,
normal...
```

5. If storage has been replaced at the disaster site, you might need to remirror the aggregates.

Healing the root aggregates after negotiated switchover

After the data aggregates have been healed, you must heal the root aggregates in preparation for the switchback operation.

Before you begin

The data aggregates phase of the MetroCluster healing process must have been completed successfully.

Steps

1. Switch back the mirrored aggregates by running the `metrocluster heal -phase root-aggregates` command.

```
cluster_A::> metrocluster heal -phase root-aggregates
[Job 137] Job succeeded: Heal Root Aggregates is successful
```

   If the healing is vetoed, you have the option of reissuing the `metrocluster heal` command with the `–override-vetoes` parameter. If you use this optional parameter, the system overrides any soft vetoes that prevent the healing operation.

2. Confirm the heal operation is complete by running the `metrocluster operation show` command on the healthy cluster:

```
cluster_A::> metrocluster operation show
  Operation: heal-root-aggregates
      State: successful
 Start Time: 7/29/2014 20:54:41
   End Time: 7/29/2014 20:54:42
     Errors: -
```

3. Check for and remove any failed disks belonging to the disaster site by issuing the following command on the healthy site:

   `disk show -broken`

4. Power up or boot each controller module on the disaster site.

   If the system displays the LOADER prompt, run the **`boot_ontap`** command.

5. After nodes are booted, verify that the root aggregates are mirrored.

   If both plexes are present, resynchronization will occur automatically if the plexes are not synchronized. If one plex has failed, that plex must be destroyed and the mirror must be recreated using the `storage aggregate mirror -aggregate` *`aggregate-name`* command to reestablish the mirror relationship.

Healing the configuration in a MetroCluster IP configuration (ONTAP 9.4 and earlier)

You must heal the aggregates in preparation for the switchback operation.

Before you begin

The following conditions must exist before performing the healing procedure:

- Switchover must have been performed and the surviving site must be serving data.
- Storage shelves at the disaster site must be powered up, functional, and accessible.
- ISLs must be up and operating.
- Nodes in the surviving site must not be in HA failover state (both nodes must be up and running).

About this task

This task applies to MetroCluster IP configurations running ONTAP versions prior to 9.5 only.

This procedure differs from the healing procedure for MetroCluster FC configurations.

Steps

1. Power up each controller module on the site that was switched over and let them fully boot.

   If the system displays the LOADER prompt, run the **`boot_ontap`** command.

2. Perform the root aggregate healing phase:

   `metrocluster heal root-aggregates`

   ```
   cluster_A::> metrocluster heal root-aggregates
   [Job 137] Job succeeded: Heal Root-Aggregates is successful
   ```

   If the healing is vetoed, you have the option of reissuing the `metrocluster heal root-aggregates` command with the `–override-vetoes` parameter. If you use this optional parameter, the system overrides any soft vetoes that prevent the healing operation.

3. Resynchronize the aggregates:

   `metrocluster heal aggregates`

   ```
   cluster_A::> metrocluster heal aggregates
   [Job 137] Job succeeded: Heal Aggregates is successful
   ```

   If the healing is vetoed, you have the option of reissuing the `metrocluster heal` command with the `–override-vetoes` parameter. If you use this optional parameter, the system overrides any soft vetoes that prevent the healing operation.

4. Confirm the heal operation is complete by running the `metrocluster operation show` command on the healthy cluster:

   ```
   cluster_A::> metrocluster operation show
     Operation: heal-aggregates
         State: successful
    Start Time: 7/29/2017 20:54:41
      End Time: 7/29/2017 20:54:42
        Errors: -
   ```

Performing a switchback

After you heal the MetroCluster configuration, you can perform the MetroCluster switchback operation. The MetroCluster switchback operation returns the configuration to its normal operating state, with the sync-source storage virtual machines (SVMs) on the disaster site active and serving data from the local disk pools.

Before you begin

- The disaster cluster must have successfully switched over to the surviving cluster.
- Healing must have been performed on the data and root aggregates.
- The surviving cluster nodes must not be in the HA failover state (all nodes must be up and running for each HA pair).
- The disaster site controller modules must be completely booted and not in the HA takeover mode.
- The root aggregate must be mirrored.
- The Inter-Switch Links (ISLs) must be online.
- Any required licenses must be installed on the system.

Steps

1. Confirm that all nodes are in the enabled state:

   `metrocluster node show`

   The following example displays the nodes that are in the enabled state:

```
cluster_B::>  metrocluster node show

DR                        Configuration  DR
Group Cluster Node        State          Mirroring Mode
----- ------- ----------- -------------- --------- --------------------
1     cluster_A
              node_A_1    configured     enabled   heal roots completed
              node_A_2    configured     enabled   heal roots completed
      cluster_B
              node_B_1    configured     enabled   waiting for switchback recovery
              node_B_2    configured     enabled   waiting for switchback recovery
4 entries were displayed.
```

2. Confirm that resynchronization is complete on all SVMs:

   `metrocluster vserver show`

3. Verify that any automatic LIF migrations being performed by the healing operations have been successfully completed: `metrocluster check lif show`
4. Perform the switchback by running the `metrocluster switchback` command from any node in the surviving cluster.
5. Check the progress of the switchback operation:

   `metrocluster show`

   The switchback operation is still in progress when the output displays `waiting-for-switchback`:

```
cluster_B::> metrocluster show
Cluster                   Entry Name          State
------------------------- ------------------- -----------
 Local: cluster_B         Configuration state configured
                          Mode                switchover
                          AUSO Failure Domain -
Remote: cluster_A         Configuration state configured
                          Mode                waiting-for-switchback
                          AUSO Failure Domain -
```

   The switchback operation is complete when the output displays `normal`:

```
cluster_B::> metrocluster show
Cluster                   Entry Name          State
------------------------- ------------------- -----------
 Local: cluster_B         Configuration state configured
                          Mode                normal
                          AUSO Failure Domain -
Remote: cluster_A         Configuration state configured
                          Mode                normal
                          AUSO Failure Domain -
```

   If a switchback takes a long time to finish, you can check on the status of in-progress baselines by using the `metrocluster config-replication resync-status show` command. This command is at the advanced privilege level.
6. Reestablish any SnapMirror or SnapVault configurations.

   In ONTAP 8.3, you need to manually reestablish a lost SnapMirror configuration after a MetroCluster switchback operation. In ONTAP 9.0 and later, the relationship is reestablished automatically.

Verifying a successful switchback

After performing the switchback, you want to confirm that all aggregates and storage virtual machines (SVMs) are switched back and online.

Steps

1. Verify that the switched-over data aggregates are switched back:

   `storage aggregate show`
   In the following example, aggr_b2 on node B2 has switched back:

```
node_B_1::> storage aggregate show
Aggregate     Size Available Used% State   #Vols  Nodes            RAID Status
--------- -------- --------- ----- ------- ------ ---------------- ------------
...
aggr_b2    227.1GB   227.1GB    0% online       0 node_B_2   raid_dp,
                                                                  mirrored,
                                                                  normal

node_A_1::> aggr show
Aggregate     Size Available Used% State   #Vols  Nodes            RAID Status
--------- -------- --------- ----- ------- ------ ---------------- ------------
...
aggr_b2          -         -     - unknown      - node_A_1
```

   If the disaster site included unmirrored aggregates and the unmirrored aggregates are no longer present, the aggregate may show up with a `State` of `unknown` in the output of the `storage aggregate show` command. Contact technical support to remove the out-of-date entries for the unmirrored aggregates.

2. Verify that all sync-destination SVMs on the surviving cluster are dormant (showing an `Admin State` of `stopped`) and the sync-source SVMs on the disaster cluster are up and running:

   `vserver show -subtype sync-source`

```
node_B_1::> vserver show -subtype sync-source
                              Admin      Root                        Name    Name
Vserver     Type    Subtype    State      Volume     Aggregate       Service Mapping
----------- ------- ---------- ---------- ---------- ----------      ------- -------
...
vs1a        data    sync-source
                              running    vs1a_vol   node_B_2        file    file
                                                                    aggr_b2

node_A_1::> vserver show -subtype sync-destination
                              Admin      Root                          Name    Name
Vserver            Type    Subtype    State      Volume     Aggregate  Service Mapping
-----------        ------- ---------- ---------- ---------- ---------- ------- -------
...
cluster_A-vs1a-mc  data    sync-destination
                                    stopped    vs1a_vol   sosb_      file    file
                                                                      aggr_b2
```

   Sync-destination aggregates in the MetroCluster configuration have the suffix "-mc" automatically appended to their name to help identify them.

3. Confirm that the switchback operations succeeded by using the `metrocluster operation show` command.

If the command output shows...	Then...
That the switchback operation state is successful.	The switchback process is complete and you can proceed with operation of the system.
That the switchback operation or switchback-continuation-agent operation is partially successful.	Perform the suggested fix provided in the output of the `metrocluster operation show` command.

After you finish

You must repeat the previous sections to perform the switchback in the opposite direction. If site_A did a switchover of site_B, have site_B do a switchover of site_A.

Performing a forced switchover after a disaster

An administrator, or the MetroCluster Tiebreaker software if it is configured, must determine that a disaster has occurred and perform the MetroCluster switchover. In either case, there are steps you must perform on both the disaster cluster and the surviving cluster after the switchover to ensure safe and continued data service.

Steps

1. *Fencing off the disaster site* on page 38
2. *Performing a forced switchover* on page 38
3. *Output for the storage aggregate plex show command is indeterminate after a MetroCluster switchover* on page 39
4. *Accessing volumes in NVFAIL state after a switchover* on page 39

Fencing off the disaster site

After the disaster, if the disaster site nodes must be replaced, you must halt them to prevent the site from resuming service. Otherwise, you risk the possibility of data corruption if clients start accessing the nodes before the replacement procedure is completed.

Step

Halt the nodes at the disaster site and keep them powered down or at the LOADER prompt until directed to boot ONTAP:

`system node halt -node` ***`disaster-site-node-name`***

If the disaster site nodes have been destroyed or cannot be halted, turn off power to the nodes and do not boot the replacement nodes until directed to in the recovery procedure.

Performing a forced switchover

The switchover process, in addition to providing nondisruptive operations during testing and maintenance, enables you to recover from a site failure with a single command.

Before you begin

- At least one of the surviving site nodes must up and running before you perform the switchover.
- All previous configuration changes must be complete before performing a switchback operation.
  This is to avoid competition with the negotiated switchover or switchback operation.

Note: SnapMirror and SnapVault configurations are deleted automatically.

About this task

The `metrocluster switchover` command switches over the nodes in all DR groups in the MetroCluster configuration. For example, in an eight-node MetroCluster configuration, it switches over the nodes in both DR groups.

Steps

1. Implement the switchover by running the `metrocluster switchover -forced-on-disaster true` command.
   The operation can take a period of minutes to complete.
2. Answer **y** when prompted to continue with the switchover.

3. Verify that the switchover was completed successfully by running the `metrocluster operation show` command.

```
mcc1A::> metrocluster operation show
  Operation: switchover
 Start time: 10/4/2012 19:04:13
      State: in-progress
   End time: -
     Errors:

mcc1A::> metrocluster operation show
  Operation: switchover
 Start time: 10/4/2012 19:04:13
      State: successful
   End time: 10/4/2012 19:04:22
     Errors: -
```

If the switchover is vetoed, you have the option of reissuing the `metrocluster switchover -forced-on-disaster true` command with the `–override-vetoes` option. If you use this optional parameter, the system overrides any soft vetoes that prevented the switchover.

After you finish

SnapMirror relationships need to be reestablished after switchover.

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.

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

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.

Related concepts

Monitoring and protecting the file system consistency using NVFAIL on page 147

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.

Choosing the correct recovery procedure

After a failure in a MetroCluster configuration, you must select the correct recovery procedure, depending on the extent of the failure.

Evaluate the disaster, using the following table and examples to select the appropriate recovery procedure.

Scope of failures at disaster site	Procedure
• No controller module failure • Other hardware has failed	*Recovering from a non-controller failure* on page 135
• Single controller module failure or failure of FRU components within the controller module • Drives have not failed	If a failure is limited to a single controller module, you must use the controller module FRU replacement procedure for the platform model. In a four or eight-node MetroCluster configuration, such a failure is isolated to the local HA pair. **Note:** The controller module FRU replacement procedure can be used in a two-node MetroCluster configuration if there are no drive or other hardware failures. *AFF and FAS Documentation Center*
• Single controller module failure or failure of FRU components within the controller module • Drives have failed	*Recovering from a multi-controller or storage failure* on page 47
• Single controller module failure or failure of FRU components within the controller module • Drives have not failed • Additional hardware outside the controller module has failed	*Recovering from a multi-controller or storage failure* on page 47 You should skip all steps for drive assignment.
• Multiple controller module failure (with or without additional failures) within a DR group	*Recovering from a multi-controller or storage failure* on page 47

Controller module failure scenarios during MetroCluster FC-to-IP Transition

The recovery procedure can be used if a site failure occurs during transition. However, it can only be used if the configuration is a stable mixed configuration, with the FC DR group and IP DR group both fully configured. The output of the `metrocluster node show` command should show both DR groups with all eight nodes.

Attention: If the failure occurred during transition when the nodes are in the process of being added or removed, you must contact technical support.

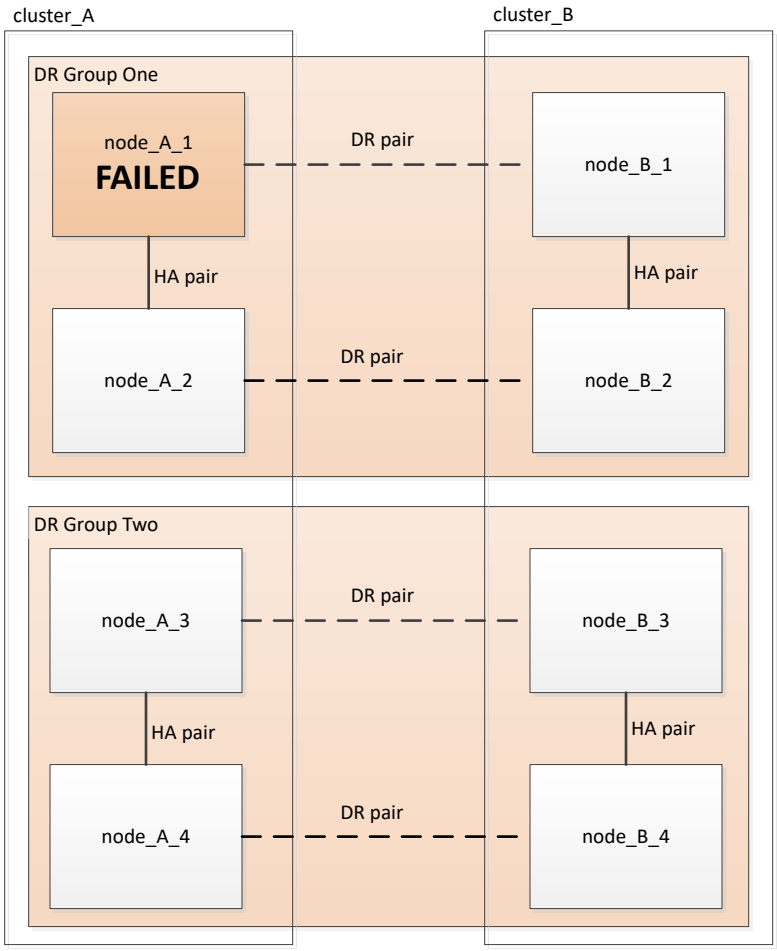
Controller module failure scenarios in eight-node MetroCluster configurations

Single controller module failures in a single DR group

In this case the failure is limited to an HA pair.

- If no storage requires replacement, you can use the controller module FRU replacement procedure for the platform model.
  AFF and FAS Documentation Center
- If storage requires replacement, you can use the multi-controller module recovery procedure.
  Recovering from a multi-controller or storage failure on page 47

This scenario applies to four-node MetroCluster configurations also.

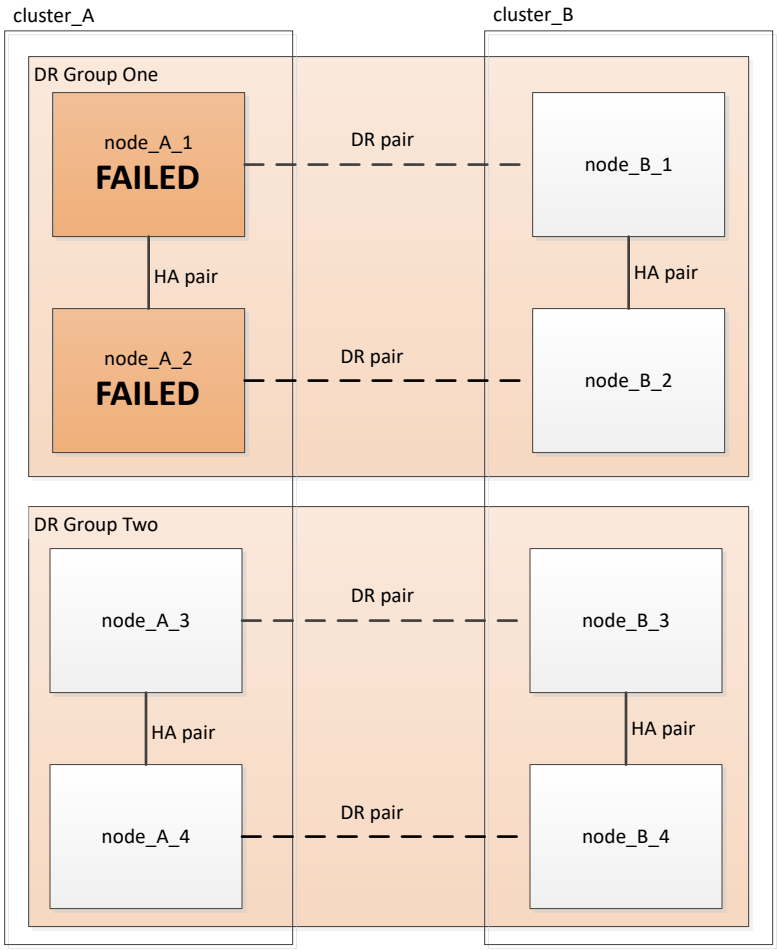


Two controller module failures in a single DR group

In this case the failure requires a switchover. You can use the multi-controller module failure recovery procedure.

Recovering from a multi-controller or storage failure on page 47

This scenario applies to four-node MetroCluster configurations also.



Single controller module failures in separate DR groups.

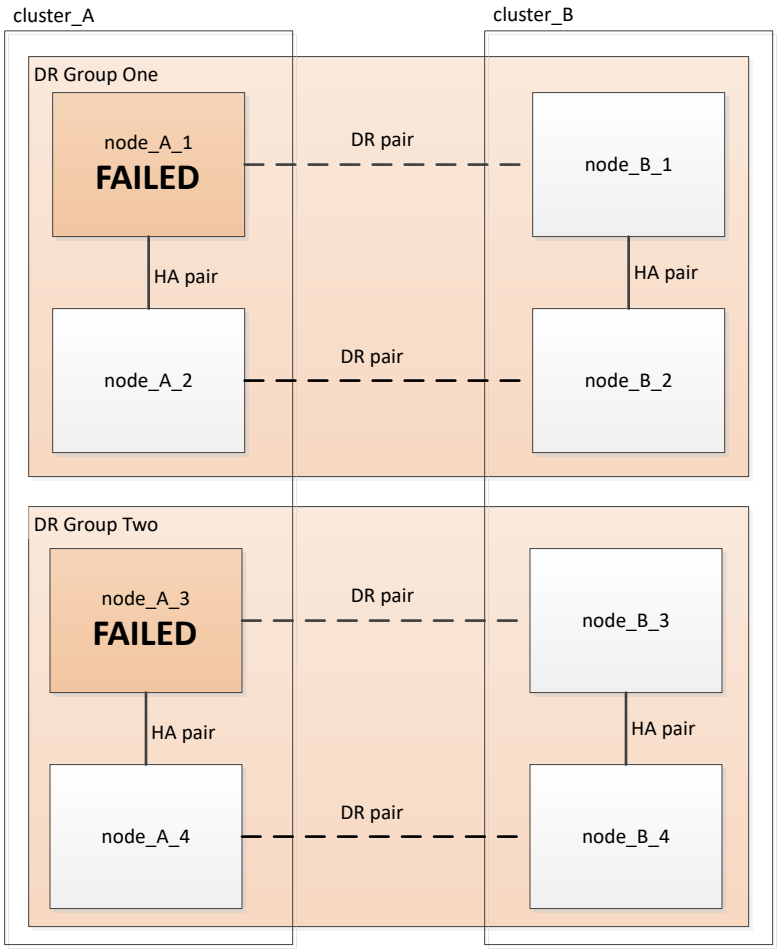
In this case the failure is limited to separate HA pairs.

- If no storage requires replacement, you can use the controller module FRU replacement procedure for the platform model.

  The FRU replacement procedure is performed twice, once for each failed controller module.

  AFF and FAS Documentation Center
- If storage requires replacement, you can use the multi-controller module recovery procedure.

  Recovering from a multi-controller or storage failure on page 47



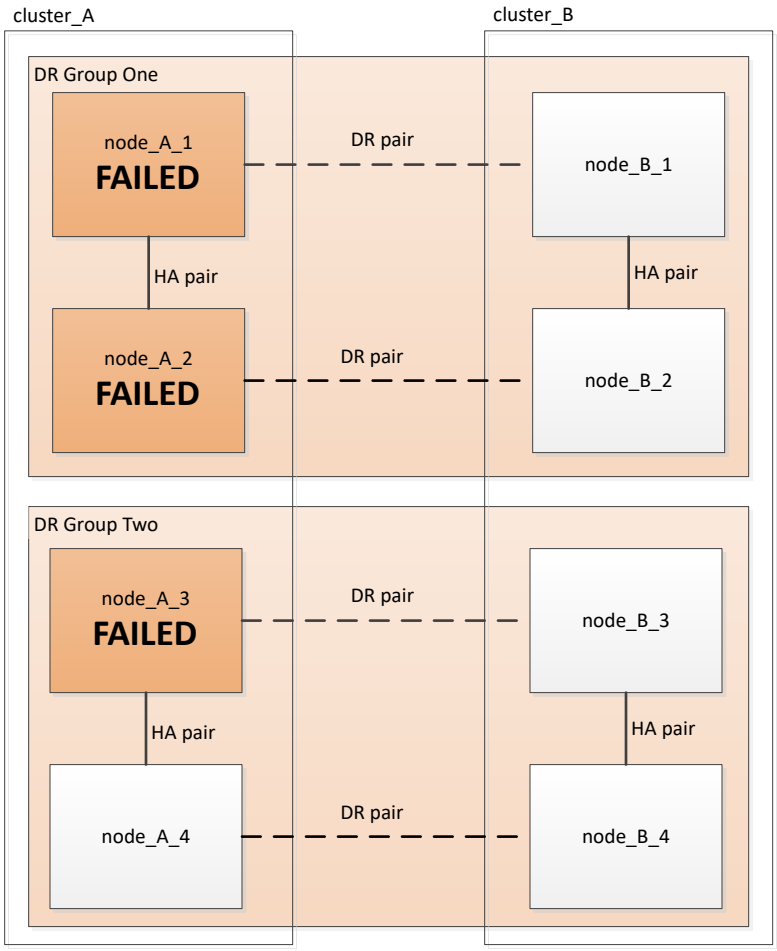
Three controller module failures spread across the DR groups

In this case the failure requires a switchover. You can use the multi-controller module failure recovery procedure for DR Group One.

Recovering from a multi-controller or storage failure on page 47

You can use the platform-specific controller module FRU replacement procedure for DR Group Two.

AFF and FAS Documentation Center

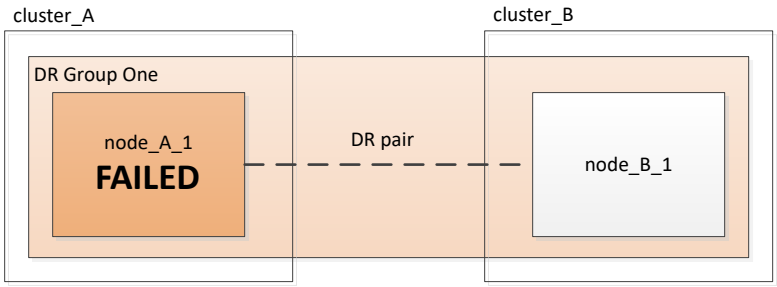


Controller module failure scenarios in two-node MetroCluster configurations

The procedure you use depends on the extent of the failure.

- If no storage requires replacement, you can use the controller module FRU replacement procedure for the platform model.
  AFF and FAS Documentation Center
- If storage requires replacement, you can use the multi-controller module recovery procedure.

Recovering from a multi-controller or storage failure on page 47


cluster_A
cluster_B
DR Group One
node_A_1
FAILED
DR pair
node_B_1

Recovering from a multi-controller or storage failure

If the controller failure extends to all controller modules on one side of a DR group in a four or eight-node MetroCluster configuration, or storage has been replaced, you must replace the equipment and reassign ownership of drives to recover from the disaster.

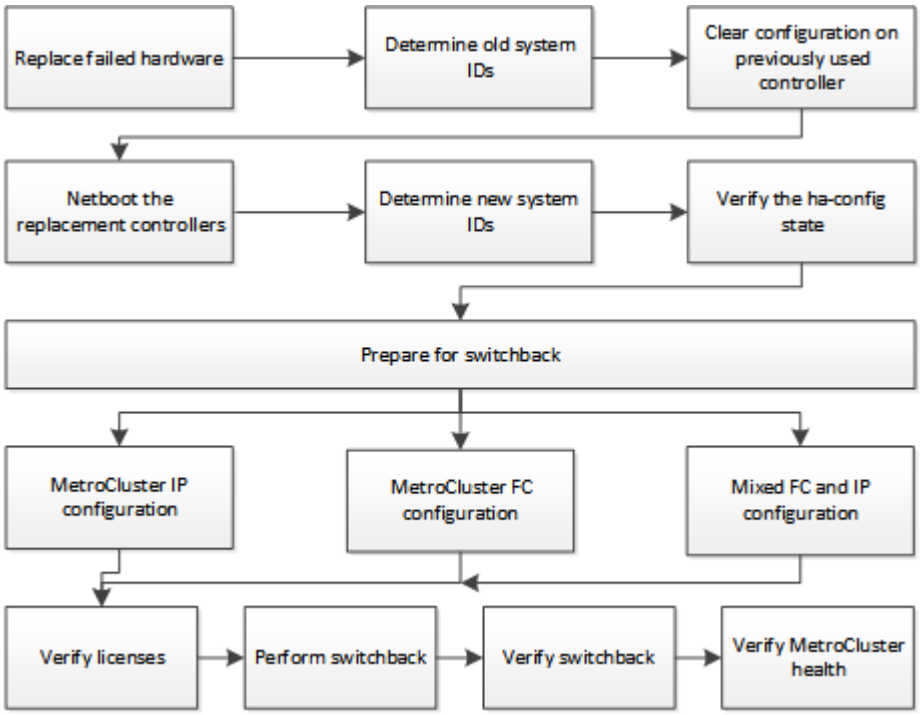
Before you begin

- You should review the available recovery procedures before deciding to use this procedure.
  Choosing the correct recovery procedure on page 41
- The disaster site must be fenced off.
  Fencing off the disaster site.
- Switchover must have been performed.
  Performing a forced switchover.
- Replacement drives and the controller modules must be new and must not have been assigned ownership previously.

About this task

The examples in this procedure show two or four-node configurations. If you have an eight-node configuration (two DR groups), you must take into account any failures and perform the required recovery task on the additional controller modules.

This procedure uses the following workflow:



This procedure can be used when performing recovery on a system that was in mid-transition when the failure occurred. In that case, you must perform the appropriate steps when preparing for switchback, as indicated in the procedure.

Steps

1. *Replacing hardware at the disaster site* on page 48
2. *Determining the system IDs and VLAN IDs of the old controller modules* on page 50

3. *Isolating replacement drives from the surviving site (MetroCluster IP configurations)* on page 52
4. *Clearing the configuration on a controller module* on page 52
5. *Netbooting the new controller modules* on page 53
6. *Determining the system IDs of the replacement controller modules* on page 55
7. *Verifying the ha-config state of components* on page 56
8. *Preparing the disaster site for switchback* on page 57
9. *Reestablishing object stores for FabricPool configurations* on page 128
10. *Verifying licenses on the replaced nodes* on page 128
11. *Performing a switchback* on page 129
12. *Verifying a successful switchback* on page 130
13. *Mirroring the root aggregates of the replacement nodes* on page 131
14. *Reconfiguring the ONTAP Mediator service (MetroCluster IP configurations)* on page 133
15. *Verifying the health of the MetroCluster configuration* on page 133

Replacing hardware at the disaster site

If hardware components have to be replaced, you must replace them using their individual hardware replacement and installation guides.

Before you begin

The storage controllers must be powered off or remain halted (showing the LOADER prompt).

Steps

1. Replace the components as necessary.

   Note: In this step, you replace and cable the components exactly as they were cabled prior to the disaster. You must not power up the components.

If you are replacing...	Perform these steps...	Using these guides...
FC switches in a MetroCluster FC configuration	**a.** Install the new switches. **b.** Cable the ISL links. Do not power on the FC switches at this time.	*MetroCluster Service and Expansion Guide*
IP switches in a MetroCluster IP configuration	**a.** Install the new switches. **b.** Cable the ISL links. Do not power on the IP switches at this time.	*MetroCluster IP installation and configuration*

If you are replacing...	Perform these steps...	Using these guides...
Disk shelves	**a.** Install the disk shelves and disks. • Disk shelf stacks should be the same configuration as at the surviving site. • Disks can be the same size or larger, but must be of the same type (SAS or SATA). **b.** Cable the disk shelves to adjacent shelves within the stack and to the FC-to-SAS bridge. Do not power on the disk shelves at this time.	*AFF and FAS Documentation Center*
SAS cables	**a.** Install the new cables. Do not power on the disk shelves at this time.	*AFF and FAS Documentation Center*
FC-to-SAS bridges in a MetroCluster FC configuration	**a.** Install the FC-to-SAS bridges. **b.** Cable the FC-to-SAS bridges. Cable them to the FC switches or to the controller modules, depending on your MetroCluster configuration type. Do not power on the FC-to-SAS bridges at this time.	*Fabric-attached MetroCluster installation and configuration* *Stretch MetroCluster installation and configuration*

If you are replacing...	Perform these steps...	Using these guides...
Controller modules	**a.** Install the new controller modules: • The controller modules must be the same model as those being replaced. For example, 8080 controller modules must be replaced with 8080 controller modules. • The controller modules must not have previously been part of either cluster within the MetroCluster configuration or any previously existing cluster configuration. If they were, you must set defaults and perform a "wipeconfig" process. • Ensure that all network interface cards (such as Ethernet or FC) are in the same slots used on the old controller modules. **b.** Cable the new controller modules exactly the same as the old ones. The ports connecting the controller module to the storage (either by connections to the IP or FC switches, FC-to-SAS bridges, or directly) should be the same as those used prior to the disaster. Do not power on the controller modules at this time.	*AFF and FAS Documentation Center*

2. Verify that all components are cabled correctly according the *MetroCluster Installation and Configuration Guide* for your configuration.

Determining the system IDs and VLAN IDs of the old controller modules

After you have replaced all hardware at the disaster site, you must determine the system IDs of the replaced controller modules. You need the old system IDs when you reassign disks to the new controller modules. If the systems are AFF A220, AFF A250, AFF A400, AFF A800, FAS2750, FAS500f, FAS8300, or FAS8700 models, you must also determine the VLAN IDs used by the MCC IP interfaces.

Before you begin

All equipment at the disaster site must be powered off.

About this task

This discussion provides examples for two and four-node configurations. For eight-node configurations, you must account for any failures in the additional nodes on the second DR group.

For a two-node MetroCluster configuration, you can ignore references to the second controller module at each site.

The examples in this procedure are based on the following assumptions:

- Site A is the disaster site.

- node_A_1 has failed and is being completely replaced.
- node_A_2 has failed and is being completely replaced.
  node _A_2 is present in a four-node MetroCluster configuration only.
- Site B is the surviving site.
- node_B_1 is healthy.
- node_B_2 is healthy.
  node_B_2 is present in a four-node MetroCluster configuration only.

The controller modules have the following original system IDs:

Number of nodes in MetroCluster configuration	Node	Original system ID
Four	node_A_1	4068741258
	node_A_2	4068741260
	node_B_1	4068741254
	node_B_2	4068741256
Two	node_A_1	4068741258
	node_B_1	4068741254

Steps

1. From the surviving site, display the system IDs of the nodes in the MetroCluster configuration.

Number of nodes in MetroCluster configuration	Use this command
Four or eight	**metrocluster node show -fields node-systemid,ha-partner-systemid,dr-partner-systemid,dr-auxiliary-systemid**
Two	**metrocluster node show -fields node-systemid,dr-partner-systemid**

In this example for a four-node MetroCluster configuration, the following old system IDs are retrieved:

- Node_A_1: 4068741258
- Node_A_2: 4068741260

Disks owned by the old controller modules are still owned these system IDs.

```
metrocluster node show -fields node-systemid,ha-partner-systemid,dr-partner-systemid,dr-auxiliary-systemid

dr-group-id cluster    node      node-systemid ha-partner-systemid dr-partner-systemid dr-auxiliary-systemid
----------- ---------- --------  ------------- ------------------- ------------------- ---------------------
1           Cluster_A  Node_A_1  4068741258    4068741260          4068741254          4068741256
1           Cluster_A  Node_A_2  4068741260    4068741258          4068741256          4068741254
1           Cluster_B  Node_B_1  -             -                   -                   -
1           Cluster_B  Node_B_2  -             -                   -                   -
4 entries were displayed.
```

In this example for a two-node MetroCluster configuration, the following old system ID is retrieved:

- Node_A_1: 4068741258

Disks owned by the old controller module are still owned this system ID.

```
metrocluster node show -fields node-systemid,dr-partner-systemid

dr-group-id cluster    node      node-systemid dr-partner-systemid
----------- ---------- --------  ------------- -------------------
1           Cluster_A  Node_A_1  4068741258    4068741254
1           Cluster_B  Node_B_1  -             -
2 entries were displayed.
```

2. For MetroCluster IP configurations using the ONTAP Mediator service, get the IP address of the ONTAP Mediator service:

   storage iscsi-initiator show -node * -label mediator

3. If the systems are AFF A220, AFF A400, FAS2750, FAS8300, or FAS8700 models, determine the VLAN IDs:

   metrocluster interconnect show

   The VLAN IDs are included in the adapter name shown in the Adapter column of the output.

   In this example the VLAN IDs are 120 and 130:

```
metrocluster interconnect show
                    Mirror   Mirror
                Partner Admin    Oper
Node Partner Name Type    Status   Status  Adapter Type   Status
---- ------------ ------- -------- ------- ------- ------ ------
Node_A_1 Node_A_2 HA      enabled  online
                                           e0a-120 iWARP  Up
                                           e0b-130 iWARP  Up
         Node_B_1 DR      enabled  online
                                           e0a-120 iWARP  Up
                                           e0b-130 iWARP  Up
         Node_B_2 AUX     enabled  offline
                                           e0a-120 iWARP  Up
                                           e0b-130 iWARP  Up
Node_A_2 Node_A_1 HA      enabled  online
                                           e0a-120 iWARP  Up
                                           e0b-130 iWARP  Up
         Node_B_2 DR      enabled  online
                                           e0a-120 iWARP  Up
                                           e0b-130 iWARP  Up
         Node_B_1 AUX     enabled  offline
                                           e0a-120 iWARP  Up
                                           e0b-130 iWARP  Up
12 entries were displayed.
```

Isolating replacement drives from the surviving site (MetroCluster IP configurations)

You must isolate any replacement drives by taking down the MetroCluster iSCSI initiator connections from the surviving nodes.

About this task

This procedure is only required on MetroCluster IP configurations.

Steps

1. From either surviving node's prompt, change to the advanced privilege level:

   set -privilege advanced

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

2. Disconnect the iSCSI initiators on both surviving nodes in the DR group:

   storage iscsi-initiator disconnect -node *surviving-node* -label *

   This command must be issued twice, once for each of the surviving nodes.

   The following example shows the commands for disconnecting the initiators on site B:

```
site_B::*> storage iscsi-initiator disconnect -node node_B_1 -label *
site_B::*> storage iscsi-initiator disconnect -node node_B_2 -label *
```

3. Return to the admin privilege level:

   set -privilege admin

Clearing the configuration on a controller module

Before using a new controller module in the MetroCluster configuration, you must clear the configuration.

Steps

1. If necessary, halt the node to display the LOADER prompt:

   `halt`
2. At the LOADER prompt, set the environmental variables to default values:

   `set-defaults`
3. Save the environment:

   `saveenv`

   `bye`
4. At the LOADER prompt, launch the boot menu:

   `boot_ontap menu`
5. At the boot menu prompt, clear the configuration:

   `wipeconfig`

   Respond **`yes`** to the confirmation prompt.

   The node reboots and the boot menu is displayed again.
6. At the boot menu, select option **5** to boot the system into Maintenance mode.

   Respond **`yes`** to the confirmation prompt.

Netbooting the new controller modules

If the new controller modules have a different version of ONTAP from the version on the surviving controller modules, you must netboot the new controller modules.

Before you begin

- You must have access to an HTTP server.
- You must have access to the NetApp Support Site to download the necessary system files for your platform and version of ONTAP software that is running on it.
  NetApp Support

Steps

1. Netboot the new controllers:
   a. Access the *NetApp Support Site* to download the files used for performing the netboot of the system.
   b. Download the appropriate ONTAP software from the software download section of the NetApp Support Site and store the *`ontap-version`*`_image.tgz` file on a web-accessible directory.
   c. Change to the web-accessible directory and verify that the files you need are available.

If the platform model is...	Then...
FAS/AFF8000 series systems	Extract the contents of the *ontap-version*_image.tgz file to the target directory: `tar -zxvf ontap-version_image.tgz` **Note:** If you are extracting the contents on Windows, use 7-Zip or WinRAR to extract the netboot image.Your directory listing should contain a netboot folder with a kernel file:netboot/kernel Your directory listing should contain a netboot folder with a kernel file: `netboot/kernel`
All other systems	Your directory listing should contain a netboot folder with a kernel file: *ontap-version*_image.tgz You do not need to extract the *ontap-version*_image.tgz file.

d. At the LOADER prompt, configure the netboot connection for a management LIF:

If IP addressing is...	Then...
DHCP	Configure the automatic connection: **ifconfig e0M -auto**
Static	Configure the manual connection: **ifconfig e0M -addr=*ip_addr* -mask=*netmask* -gw=*gateway***

e. Perform the netboot.

If the platform model is...	Then...
FAS/AFF8000 series systems	**netboot http://*web_server_ip/path_to_web-accessible_directory/*netboot/kernel**
All other systems	**netboot http://*web_server_ip/path_to_web-accessible_directory/ontap-version*_image.tgz**

f. From the boot menu, select option **(7) Install new software first** to download and install the new software image to the boot device.

Disregard the following message: `"This procedure is not supported for Non-Disruptive Upgrade on an HA pair"`. It applies to nondisruptive upgrades of software, not to upgrades of controllers.

g. If you are prompted to continue the procedure, enter **y**, and when prompted for the package, enter the URL of the image file:

http://*web_server_ip/path_to_web-accessible_directory/ontap-version*_image.tgz

Enter username/password if applicable, or press Enter to continue.

h. Be sure to enter **n** to skip the backup recovery when you see a prompt similar to the following:

```
Do you want to restore the backup configuration now? {y|n} n
```

i. Reboot by entering **y** when you see a prompt similar to the following:

```
The node must be rebooted to start using the newly installed software. Do you want to reboot
now? {y|n}
```

2. From the Boot menu, select **option 5** to enter Maintenance mode.
3. If you have a four-node MetroCluster configuration, repeat this procedure on the other new controller module.

Determining the system IDs of the replacement controller modules

After you have replaced all hardware at the disaster site, you must determine the system ID of the newly installed storage controller module or modules.

About this task

You must perform this procedure with the replacement controller modules in Maintenance mode.

This section provides examples for two and four-node configurations. For two-node configurations, you can ignore references to the second node at each site. For eight-node configurations, you must account for the additional nodes on the second DR group. The examples make the following assumptions:

- Site A is the disaster site.
- node_A_1 has been replaced.
- node_A_2 has been replaced.
  Present only in four-node MetroCluster configurations.
- Site B is the surviving site.
- node_B_1 is healthy.
- node_B_2 is healthy.
  Present only in four-node MetroCluster configurations.

The examples in this procedure use controllers with the following system IDs:

Number of nodes in MetroCluster configuration	Node	Original system ID	New system ID	Will pair with this node as DR partner
Four	node_A_1	4068741258	1574774970	node_B_1
	node_A_2	4068741260	1574774991	node_B_2
	node_B_1	4068741254	unchanged	node_A_1
	node_B_2	4068741256	unchanged	node_A_2
Two	node_A_1	4068741258	1574774970	node_B_1
	node_B_1	4068741254	unchanged	node_A_1

Note: In a four-node MetroCluster configuration, the system determines DR partnerships by pairing the node with the lowest system ID at site_A and the node with the lowest system ID at site_B. Because the system IDs change, the DR pairs might be different after the controller replacements are completed than they were prior to the disaster.

In the preceding example:

- node_A_1 (1574774970) will be paired with node_B_1 (4068741254)
- node_A_2 (1574774991) will be paired with node_B_2 (4068741256)

Steps

1. With the node in Maintenance mode, display the local system ID of the node from each node:

   `disk show`
   In the following example, the new local system ID is 1574774970:

   ```
   *> disk show
    Local System ID: 1574774970
    ...
   ```

2. On the second node, repeat the previous step.

   Note: This step is not required in a two-node MetroCluster configuration.

   In the following example, the new local system ID is 1574774991:

   ```
   *> disk show
    Local System ID: 1574774991
    ...
   ```

Verifying the ha-config state of components

In a MetroCluster configuration, the ha-config state of the controller module and chassis components must be set to `mcc` or `mcc-2n` so they boot up properly.

Before you begin

The system must be in Maintenance mode.

About this task

This task must be performed on each new controller module.

Steps

1. In Maintenance mode, display the HA state of the controller module and chassis:

   `ha-config show`

   The correct HA state depends on your MetroCluster configuration.

Number of controllers in the MetroCluster configuration	HA state for all components should be...
Eight- or four-node MetroCluster FC configuration	mcc
Two-node MetroCluster FC configuration	mcc-2n
MetroCluster IP configuration	mccip

2. If the displayed system state of the controller is not correct, set the HA state for the controller module:

Number of controllers in the MetroCluster configuration	Command
Eight- or four-node MetroCluster FC configuration	`ha-config modify controller mcc`
Two-node MetroCluster FC configuration	`ha-config modify controller mcc-2n`

Number of controllers in the MetroCluster configuration	Command
MetroCluster IP configuration	`ha-config modify controller mccip`

3. If the displayed system state of the chassis is not correct, set the HA state for the chassis:

Number of controllers in the MetroCluster configuration	Command
Eight- or four-node MetroCluster FC configuration	`ha-config modify chassis mcc`
Two-node MetroCluster FC configuration	`ha-config modify chassis mcc-2n`
MetroCluster IP configuration	`ha-config modify chassis mccip`

4. Repeat these steps on the other replacement node.

Preparing the disaster site for switchback

You must take steps to prepare for the switchback operation. The steps you take depend on your configuration.

Choices

- *Preparing for switchback in a MetroCluster FC configuration* on page 57
- *Preparing for switchback in a MetroCluster IP configuration* on page 106
- *Preparing the nodes for switchback in a mixed configuration (recovery during transition)* on page 126

Preparing for switchback in a MetroCluster FC configuration

You must perform certain tasks in order to prepare the MetroCluster FC configuration for the switchback operation.

Steps

1. *Verifying port configuration (MetroCluster FC configurations only)* on page 57
2. *Configuring the FC-to-SAS bridges (MetroCluster FC configurations only)* on page 58
3. *Configuring the FC switches (MetroCluster FC configurations only)* on page 60
4. *Verifying the storage configuration* on page 97
5. *Powering on the equipment at the disaster site* on page 97
6. *Assigning ownership for replaced drives* on page 99
7. *Performing aggregate healing and restoring mirrors (MetroCluster FC configurations)* on page 102
8. *Reassigning disk ownership for root aggregates to replacement controller modules (MetroCluster FC configurations)* on page 103
9. *Booting the new controller modules (MetroCluster FC configurations)* on page 105

Verifying port configuration (MetroCluster FC configurations only)

You must set the environmental variables on the node and then power it off to prepare it for MetroCluster configuration.

About this task

This procedure is performed with the replacement controller modules in Maintenance mode.

The steps to check configuration of ports is needed only on systems in which FC or CNA ports are used in initiator mode.

Steps

1. In Maintenance mode, enter the following command to restore the FC port configuration:

   ucadmin modify -m fc -t *initiator adapter_name*

   If you only want to use one of a port pair in the initiator configuration, enter a precise *adapter_name*.

2. Take one of the following actions, depending on your configuration:

If the FC port configuration is...	Then...
The same for both ports	Answer **y** when prompted by the system since modifying one port in a port pair modifies the other port as well.
Different	**a.** Answer **n** when prompted by the system. **b.** Enter the following command to restore the FC port configuration: **ucadmin modify -m fc -t *initiator*|*target adapter_name***

3. Exit Maintenance mode by entering the following command:

   halt

   After you issue the command, wait until the system stops at the LOADER prompt.

4. Boot the node back into Maintenance mode for the configuration changes to take effect:

   boot_ontap maint

5. Verify the values of the variables by entering the following command:

   ucadmin show

6. Exit Maintenance mode and display the LOADER prompt:

   halt

Configuring the FC-to-SAS bridges (MetroCluster FC configurations only)

If you replaced the FC-to-SAS bridges, you must configure them when restoring the MetroCluster configuration. The procedure is identical to the initial configuration of an FC-to-SAS bridge.

Steps

1. Power on the FC-to-SAS bridges.
2. Set the IP address on the Ethernet ports by using the `set IPAddress` *`port ipaddress`* command.

   `port` can be either **`MP1`** or **`MP2`**.

   `ipaddress` can be an IP address in the format `xxx.xxx.xxx.xxx`.

   In the following example, the IP address is 10.10.10.55 on Ethernet port 1:

   ```
   Ready.
   set IPAddress MP1 10.10.10.55

   Ready. *
   ```

3. Set the IP subnet mask on the Ethernet ports by using the `set IPSubnetMask` *`port mask`* command.

   `port` can be **`MP1`** or **`MP2`**.

mask can be a subnet mask in the format `xxx.xxx.xxx.xxx`.

In the following example, the IP subnet mask is 255.255.255.0 on Ethernet port 1:

```
Ready.
set IPSubnetMask MP1 255.255.255.0

Ready. *
```

4. Set the speed on the Ethernet ports by using the `set EthernetSpeed` *port speed* command.

   port can be **MP1** or **MP2**.

   speed can be **100**, **1000**, or **auto**.

   In the following example, the Ethernet speed is set to 1000 on Ethernet port 1.

```
Ready.
set EthernetSpeed MP1 1000

Ready. *
```

5. Save the configuration by using the `saveConfiguration` command, and restart the bridge when prompted to do so.

   Saving the configuration after configuring the Ethernet ports enables you to proceed with the bridge configuration using Telnet and enables you to access the bridge using FTP to perform firmware updates.

   The following example shows the `saveConfiguration` command and the prompt to restart the bridge.

```
Ready.
SaveConfiguration
  Restart is necessary....
  Do you wish to restart (y/n) ?
Confirm with 'y'. The bridge will save and restart with the new settings.
```

6. After the FC-to-SAS bridge reboots, log in again.
7. Set the speed on the FC ports by using the `set fcdatarate` *port speed* command.

   port can be **1** or **2**.

   speed can be **2 Gb**, **4 Gb**, **8 Gb**, or **16 Gb**, depending on your model bridge.

   In the following example, the port FC1 speed is set to 8 Gb.

```
Ready.
set fcdatarate 1 8Gb

Ready. *
```

8. Set the topology on the FC ports by using the `set FCConnMode` *port mode* command.

   port can be **1** or **2**.

   mode can be **ptp**, **loop**, **ptp-loop**, or **auto**.

   In the following example, the port FC1 topology is set to ptp.

```
Ready.
set FCConnMode 1 ptp
```

```
Ready. *
```

9. Save the configuration by using the `saveConfiguration` command, and restart the bridge when prompted to do so.
   The following example shows the `saveConfiguration` command and the prompt to restart the bridge.

```
Ready.
SaveConfiguration
   Restart is necessary....
   Do you wish to restart (y/n) ?
Confirm with 'y'. The bridge will save and restart with the new settings.
```

10. After the FC-to-SAS bridge reboots, log in again.
11. If the FC-to-SAS bridge is running firmware 1.60 or later, enable SNMP.

```
Ready.
set snmp enabled

Ready. *
saveconfiguration

Restart is necessary....
Do you wish to restart (y/n) ?

Verify with 'y' to restart the FibreBridge.
```

12. Power off the FC-to-SAS bridges.

Configuring the FC switches (MetroCluster FC configurations only)

If you have replaced the FC switches in the disaster site, you must configure them using the vendor-specific procedures. You must configure one switch, verify that storage access on the surviving site is not impacted, and then configure the second switch.

Related concepts

Port assignments for FC switches when using ONTAP 9.0 on page 65
You need to verify that you are using the specified port assignments when you cable the FC switches. The port assignments are different between ONTAP 9.0 and later versions of ONTAP.

Port assignments for FC switches when using ONTAP 9.1 and later on page 79
You need to verify that you are using the specified port assignments when you cable the FC switches when using ONTAP 9.1 and later.

Related tasks

Configuring a Brocade FC switch after site disaster on page 60
You must use this Brocade-specific procedure to configure the replacement switch and enable the ISL ports.

Configuring a Cisco FC switch after site disaster on page 62
You must use the Cisco-specific procedure to configure the replacement switch and enable the ISL ports.

Configuring a Brocade FC switch after site disaster

You must use this Brocade-specific procedure to configure the replacement switch and enable the ISL ports.

About this task

The examples in this procedure are based on the following assumptions:

- Site A is the disaster site.

- FC_switch_A_1 has been replaced.
- FC_switch_A_2 has been replaced.
- Site B is the surviving site.
- FC_switch_B_1 is healthy.
- FC_switch_B_2 is healthy.

You must verify that you are using the specified port assignments when you cable the FC switches:

- *Port assignments for FC switches when using ONTAP 9.0* on page 65
- *Port assignments for FC switches when using ONTAP 9.1 and later* on page 79

The examples show two FC-to-SAS bridges. If you have more bridges, you must disable and subsequently enable the additional ports.

Steps

1. Boot and pre-configure the new switch:
   a. Power up the new switch and let it boot up.
   b. Check the firmware version on the switch to confirm it matches the version of the other FC switches:

      `firmwareShow`
   c. Configure the new switch as described in the *MetroCluster Installation and Configuration Guide*, skipping the steps for configuring zoning on the switch.

      Fabric-attached MetroCluster installation and configuration

      Stretch MetroCluster installation and configuration
   d. Disable the switch persistently:

      `switchcfgpersistentdisable`

      The switch will remain disabled after a reboot or fastboot. If this command is not available, you should use the `switchdisable` command.

      The following example shows the command on BrocadeSwitchA:

      ```
      BrocadeSwitchA:admin> switchcfgpersistentdisable
      ```

      The following example shows the command on BrocadeSwitchB:

      ```
      BrocadeSwitchA:admin> switchcfgpersistentdisable
      ```
2. Complete configuration of the new switch:
   a. Enable the ISLs on the surviving site:

      `portcfgpersistentenable` ***`port-number`***

      ```
      FC_switch_B_1:admin> portcfgpersistentenable 10
      FC_switch_B_1:admin> portcfgpersistentenable 11
      ```
   b. Enable the ISLs on the replacement switches:

      `portcfgpersistentenable` ***`port-number`***

      ```
      FC_switch_A_1:admin> portcfgpersistentenable 10
      FC_switch_A_1:admin> portcfgpersistentenable 11
      ```
   c. On the replacement switch (FC_switch_A_1 in our example) verify that the ISL's are online:

      `switchshow`

      ```
      FC_switch_A_1:admin> switchshow
      switchName: FC_switch_A_1
      switchType: 71.2
      ```

```
switchState:Online
switchMode: Native
switchRole: Principal
switchDomain:       4
switchId:   fffc03
switchWwn:  10:00:00:05:33:8c:2e:9a
zoning:             OFF
switchBeacon:       OFF

Index Port Address Media Speed State  Proto
==============================================
...
10   10    030A00 id   16G     Online  FC E-Port 10:00:00:05:33:86:89:cb
"FC_switch_A_1"
11   11    030B00 id   16G     Online  FC E-Port 10:00:00:05:33:86:89:cb
"FC_switch_A_1" (downstream)
...
```

3. Persistently enable the switch:

   `switchcfgpersistentenable`

4. Verify that the ports are online:

   `switchshow`

Configuring a Cisco FC switch after site disaster

You must use the Cisco-specific procedure to configure the replacement switch and enable the ISL ports.

About this task

The examples in this procedure are based on the following assumptions:

- Site A is the disaster site.
- FC_switch_A_1 has been replaced.
- FC_switch_A_2 has been replaced.
- Site B is the surviving site.
- FC_switch_B_1 is healthy.
- FC_switch_B_2 is healthy.

Steps

1. Configure the switch:
   a. Download the Fabric-attached MetroCluster Installation and Configuration Guide.

      Fabric-attached MetroCluster installation and configuration

   b. Follow the steps for configuring the switch in the "Configuring the Cisco FC switches" section, *except* for the "Configuring zoning on a Cisco FC switch" section.

   Zoning is configured later in this procedure.
2. On the healthy switch (in this example, FC_switch_B_1), enable the ISL ports.
   The following example shows the commands to enable the ports:

```
FC_switch_B_1# conf t
FC_switch_B_1(config)# int fc1/14-15
FC_switch_B_1(config)# no shut
FC_switch_B_1(config)# end
FC_switch_B_1# copy running-config startup-config
FC_switch_B_1#
```

3. Verify that the ISL ports are up by using the `show interface brief` command.
4. Retrieve the zoning information from the fabric.
   The following example shows the commands to distribute the zoning configuration:

```
FC_switch_B_1(config-zone)# zoneset distribute full vsan 10
FC_switch_B_1(config-zone)# zoneset distribute full vsan 20
FC_switch_B_1(config-zone)# end
```

FC_switch_B_1 is distributed to all other switches in the fabric for vsan 10 and vsan 20, and the zoning information is retrieved from FC_switch_A_1.

5. On the healthy switch, verify that the zoning information is properly retrieved from the partner switch:

   `show zone`

```
FC_switch_B_1# show zone
zone name FC-VI_Zone_1_10 vsan 10
  interface fc1/1 swwn 20:00:54:7f:ee:e3:86:50
  interface fc1/2 swwn 20:00:54:7f:ee:e3:86:50
  interface fc1/1 swwn 20:00:54:7f:ee:b8:24:c0
  interface fc1/2 swwn 20:00:54:7f:ee:b8:24:c0

zone name STOR_Zone_1_20_25A vsan 20
  interface fc1/5 swwn 20:00:54:7f:ee:e3:86:50
  interface fc1/8 swwn 20:00:54:7f:ee:e3:86:50
  interface fc1/9 swwn 20:00:54:7f:ee:e3:86:50
  interface fc1/10 swwn 20:00:54:7f:ee:e3:86:50
  interface fc1/11 swwn 20:00:54:7f:ee:e3:86:50
  interface fc1/8 swwn 20:00:54:7f:ee:b8:24:c0
  interface fc1/9 swwn 20:00:54:7f:ee:b8:24:c0
  interface fc1/10 swwn 20:00:54:7f:ee:b8:24:c0
  interface fc1/11 swwn 20:00:54:7f:ee:b8:24:c0

zone name STOR_Zone_1_20_25B vsan 20
  interface fc1/8 swwn 20:00:54:7f:ee:e3:86:50
  interface fc1/9 swwn 20:00:54:7f:ee:e3:86:50
  interface fc1/10 swwn 20:00:54:7f:ee:e3:86:50
  interface fc1/11 swwn 20:00:54:7f:ee:e3:86:50
  interface fc1/5 swwn 20:00:54:7f:ee:b8:24:c0
  interface fc1/8 swwn 20:00:54:7f:ee:b8:24:c0
  interface fc1/9 swwn 20:00:54:7f:ee:b8:24:c0
  interface fc1/10 swwn 20:00:54:7f:ee:b8:24:c0
  interface fc1/11 swwn 20:00:54:7f:ee:b8:24:c0
FC_switch_B_1#
```

6. Determine the worldwide names (WWNs) of the switches in the switch fabric.
   In this example, the two switch WWNs are as follows:

   - FC_switch_A_1: 20:00:54:7f:ee:b8:24:c0
   - FC_switch_B_1: 20:00:54:7f:ee:c6:80:78

```
FC_switch_B_1# show wwn switch
Switch WWN is 20:00:54:7f:ee:c6:80:78
FC_switch_B_1#

FC_switch_A_1# show wwn switch
Switch WWN is 20:00:54:7f:ee:b8:24:c0
FC_switch_A_1#
```

7. Enter configuration mode for the zone and remove zone members that do not belong to the switch WWNs of the two switches:

   `no member interface` ***`interface-ide`*** **`swwn`** ***`wwn`***

   In this example, the following members are not associated with the WWN of either of the switches in the fabric and must be removed:

   - Zone name FC-VI_Zone_1_10 vsan 10
     - Interface fc1/1 swwn 20:00:54:7f:ee:e3:86:50
     - Interface fc1/2 swwn 20:00:54:7f:ee:e3:86:50

Note: AFF A700 and FAS9000 systems support four FC-VI ports. You must remove all four ports from the FC-VI zone.

- Zone name STOR_Zone_1_20_25A vsan 20
  - Interface fc1/5 swwn 20:00:54:7f:ee:e3:86:50
  - Interface fc1/8 swwn 20:00:54:7f:ee:e3:86:50
  - Interface fc1/9 swwn 20:00:54:7f:ee:e3:86:50
  - Interface fc1/10 swwn 20:00:54:7f:ee:e3:86:50
  - Interface fc1/11 swwn 20:00:54:7f:ee:e3:86:50
- Zone name STOR_Zone_1_20_25B vsan 20
  - Interface fc1/8 swwn 20:00:54:7f:ee:e3:86:50
  - Interface fc1/9 swwn 20:00:54:7f:ee:e3:86:50
  - Interface fc1/10 swwn 20:00:54:7f:ee:e3:86:50
  - Interface fc1/11 swwn 20:00:54:7f:ee:e3:86:50

The following example shows the removal of these interfaces:

```
FC_switch_B_1# conf t
FC_switch_B_1(config)# zone name FC-VI_Zone_1_10 vsan 10
FC_switch_B_1(config-zone)# no member interface fc1/1 swwn 20:00:54:7f:ee:e3:86:50
FC_switch_B_1(config-zone)# no member interface fc1/2 swwn 20:00:54:7f:ee:e3:86:50
FC_switch_B_1(config-zone)# zone name STOR_Zone_1_20_25A vsan 20
FC_switch_B_1(config-zone)# no member interface fc1/5 swwn 20:00:54:7f:ee:e3:86:50
FC_switch_B_1(config-zone)# no member interface fc1/8 swwn 20:00:54:7f:ee:e3:86:50
FC_switch_B_1(config-zone)# no member interface fc1/9 swwn 20:00:54:7f:ee:e3:86:50
FC_switch_B_1(config-zone)# no member interface fc1/10 swwn 20:00:54:7f:ee:e3:86:50
FC_switch_B_1(config-zone)# no member interface fc1/11 swwn 20:00:54:7f:ee:e3:86:50
FC_switch_B_1(config-zone)# zone name STOR_Zone_1_20_25B vsan 20
FC_switch_B_1(config-zone)# no member interface fc1/8 swwn 20:00:54:7f:ee:e3:86:50
FC_switch_B_1(config-zone)# no member interface fc1/9 swwn 20:00:54:7f:ee:e3:86:50
FC_switch_B_1(config-zone)# no member interface fc1/10 swwn 20:00:54:7f:ee:e3:86:50
FC_switch_B_1(config-zone)# no member interface fc1/11 swwn 20:00:54:7f:ee:e3:86:50
FC_switch_B_1(config-zone)# save running-config startup-config
FC_switch_B_1(config-zone)# zoneset distribute full 10
FC_switch_B_1(config-zone)# zoneset distribute full 20
FC_switch_B_1(config-zone)# end
FC_switch_B_1# copy running-config startup-config
```

8. Add the ports of the new switch to the zones.
   The following example assumes that the cabling on the replacement switch is the same as on the old switch:

```
FC_switch_B_1# conf t
FC_switch_B_1(config)# zone name FC-VI_Zone_1_10 vsan 10
FC_switch_B_1(config-zone)# member interface fc1/1 swwn 20:00:54:7f:ee:c6:80:78
FC_switch_B_1(config-zone)# member interface fc1/2 swwn 20:00:54:7f:ee:c6:80:78
FC_switch_B_1(config-zone)# zone name STOR_Zone_1_20_25A vsan 20
FC_switch_B_1(config-zone)# member interface fc1/5 swwn 20:00:54:7f:ee:c6:80:78
FC_switch_B_1(config-zone)# member interface fc1/8 swwn 20:00:54:7f:ee:c6:80:78
FC_switch_B_1(config-zone)# member interface fc1/9 swwn 20:00:54:7f:ee:c6:80:78
FC_switch_B_1(config-zone)# member interface fc1/10 swwn 20:00:54:7f:ee:c6:80:78
FC_switch_B_1(config-zone)# member interface fc1/11 swwn 20:00:54:7f:ee:c6:80:78
FC_switch_B_1(config-zone)# zone name STOR_Zone_1_20_25B vsan 20
FC_switch_B_1(config-zone)# member interface fc1/8 swwn 20:00:54:7f:ee:c6:80:78
FC_switch_B_1(config-zone)# member interface fc1/9 swwn 20:00:54:7f:ee:c6:80:78
FC_switch_B_1(config-zone)# member interface fc1/10 swwn 20:00:54:7f:ee:c6:80:78
FC_switch_B_1(config-zone)# member interface fc1/11 swwn 20:00:54:7f:ee:c6:80:78
FC_switch_B_1(config-zone)# save running-config startup-config
FC_switch_B_1(config-zone)# zoneset distribute full 10
FC_switch_B_1(config-zone)# zoneset distribute full 20
```

```
FC_switch_B_1(config-zone)# end
FC_switch_B_1# copy running-config startup-config
```

9. Verify that the zoning is properly configured:

   show zone

   The following example output shows the three zones:

```
FC_switch_B_1# show zone
  zone name FC-VI_Zone_1_10 vsan 10
    interface fc1/1 swwn 20:00:54:7f:ee:c6:80:78
    interface fc1/2 swwn 20:00:54:7f:ee:c6:80:78
    interface fc1/1 swwn 20:00:54:7f:ee:b8:24:c0
    interface fc1/2 swwn 20:00:54:7f:ee:b8:24:c0

  zone name STOR_Zone_1_20_25A vsan 20
    interface fc1/5 swwn 20:00:54:7f:ee:c6:80:78
    interface fc1/8 swwn 20:00:54:7f:ee:c6:80:78
    interface fc1/9 swwn 20:00:54:7f:ee:c6:80:78
    interface fc1/10 swwn 20:00:54:7f:ee:c6:80:78
    interface fc1/11 swwn 20:00:54:7f:ee:c6:80:78
    interface fc1/8 swwn 20:00:54:7f:ee:b8:24:c0
    interface fc1/9 swwn 20:00:54:7f:ee:b8:24:c0
    interface fc1/10 swwn 20:00:54:7f:ee:b8:24:c0
    interface fc1/11 swwn 20:00:54:7f:ee:b8:24:c0

  zone name STOR_Zone_1_20_25B vsan 20
    interface fc1/8 swwn 20:00:54:7f:ee:c6:80:78
    interface fc1/9 swwn 20:00:54:7f:ee:c6:80:78
    interface fc1/10 swwn 20:00:54:7f:ee:c6:80:78
    interface fc1/11 swwn 20:00:54:7f:ee:c6:80:78
    interface fc1/5 swwn 20:00:54:7f:ee:b8:24:c0
    interface fc1/8 swwn 20:00:54:7f:ee:b8:24:c0
    interface fc1/9 swwn 20:00:54:7f:ee:b8:24:c0
    interface fc1/10 swwn 20:00:54:7f:ee:b8:24:c0
    interface fc1/11 swwn 20:00:54:7f:ee:b8:24:c0
FC_switch_B_1#
```

Port assignments for FC switches when using ONTAP 9.0

You need to verify that you are using the specified port assignments when you cable the FC switches. The port assignments are different between ONTAP 9.0 and later versions of ONTAP.

Ports that are not used for attaching initiator ports, FC-VI ports, or ISLs can be reconfigured to act as storage ports. However, if the supported RCFs are being used, the zoning must be changed accordingly.

If the supported RCF files are used, ISL ports may not connect to the same ports shown here and may need to be reconfigured manually.

Overall cabling guidelines

You should be aware of the following guidelines when using the cabling tables:

- The Brocade and Cisco switches use different port numbering:
  - On Brocade switches, the first port is numbered 0.
  - On Cisco switches, the first port is numbered 1.
- The cabling is the same for each FC switch in the switch fabric.
- AFF A300 and FAS8200 storage systems can be ordered with one of two options for FC-VI connectivity:
  - Onboard ports 0e and 0f configured in FC-VI mode.
  - Ports 1a and 1b on an FC-VI card in slot 1.

Brocade port usage for controller connections in an eight-node MetroCluster configuration running ONTAP 9.0

The cabling is the same for each FC switch in the switch fabric.

The following table shows controller port usage on Brocade switches:

MetroCluster eight-node configuration			
Component	**Port**	**Brocade 6505, 6510, or DCX 8510-8**	
		FC_switch_x_1	**FC_switch_x_2**
controller_x_1	FC-VI port a	0	-
	FC-VI port b	-	0
	HBA port a	1	-
	HBA port b	-	1
	HBA port c	2	-
	HBA port d	-	2
controller_x_2	FC-VI port a	3	-
	FC-VI port b	-	3
	HBA port a	4	-
	HBA port b	-	4
	HBA port c	5	-
	HBA port d	-	5
controller_x_3	FC-VI port a	6	
	FC-VI port b	-	6
	HBA port a	7	-
	HBA port b	-	7
	HBA port c	8	-
	HBA port d	-	8
controller_x_4	FC-VI port a	9	-
	FC-VI port b	-	9
	HBA port a	10	-
	HBA port b	-	10
	HBA port c	11	-
	HBA port d	-	11

Brocade port usage for FC-to-SAS bridge connections in an eight-node MetroCluster configuration running ONTAP 9.0

The following table shows bridge port usage when using FibreBridge 7500 bridges:

MetroCluster eight-node configuration			
FibreBridge 7500 bridge	**Port**	**Brocade 6505, 6510, or DCX 8510-8**	
		FC_switch_x_1	**FC_switch_x_2**
bridge_x_1a	FC1	12	-
	FC2	-	12
bridge_x_1b	FC1	13	-
	FC2	-	13
bridge_x_2a	FC1	14	-
	FC2	-	14
bridge_x_2b	FC1	15	-
	FC2	-	15
bridge_x_3a	FC1	16	-
	FC2	-	16
bridge_x_3b	FC1	17	-
	FC2	-	17
bridge_x_4a	FC1	18	-
	FC2	-	18
bridge_x_4b	FC1	19	-
	FC2	-	19

The following table shows bridge port usage when using FibreBridge 6500 bridges:

MetroCluster eight-node configuration			
FibreBridge 6500 bridge	**Port**	**Brocade 6505, 6510, or DCX 8510-8**	
		FC_switch_x_1	**FC_switch_x_2**
bridge_x_1a	FC1	12	-
bridge_x_1b	FC1	-	12
bridge_x_2a	FC1	13	-
bridge_x_2b	FC1	-	13
bridge_x_3a	FC1	14	-
bridge_x_3b	FC1	-	14
bridge_x_4a	FC1	15	-
bridge_x_4b	FC1	-	15
bridge_x_5a	FC1	16	-
bridge_x_5b	FC1	-	16
bridge_x_6a	FC1	17	-

MetroCluster eight-node configuration			
FibreBridge 6500 bridge	**Port**	**Brocade 6505, 6510, or DCX 8510-8**	
		FC_switch_x_1	**FC_switch_x_2**
bridge_x_6b	FC1	-	17
bridge_x_7a	FC1	18	-
bridge_x_7b	FC1	-	18
bridge_x_8a	FC1	19	-
bridge_x_8b	FC1	-	19

Brocade port usage for ISLs in an eight-node MetroCluster configuration running ONTAP 9.0

The following table shows ISL port usage:

MetroCluster eight-node configuration		
ISL port	**Brocade 6505, 6510, or DCX 8510-8**	
	FC_switch_x_1	**FC_switch_x_2**
ISL port 1	20	20
ISL port 2	21	21
ISL port 3	22	22
ISL port 4	23	23

Brocade port usage for controllers in a four-node MetroCluster configuration running ONTAP 9.0

The cabling is the same for each FC switch in the switch fabric.

MetroCluster four-node configuration			
Component	**Port**	**Brocade 6505, 6510, or DCX 8510-8**	
		FC_switch_x_1	**FC_switch_x_2**
controller_x_1	FC-VI port a	0	-
	FC-VI port b	-	0
	HBA port a	1	-
	HBA port b	-	1
	HBA port c	2	-
	HBA port d	-	2
controller_x_2	FC-VI port a	3	-
	FC-VI port b	-	3
	HBA port a	4	-
	HBA port b	-	4
	HBA port c	5	-
	HBA port d	-	5

Brocade port usage for bridges in a four-node MetroCluster configuration running ONTAP 9.0

The cabling is the same for each FC switch in the switch fabric.

The following table shows bridge port usage up to port 17 when using FibreBridge 7500 bridges. Additional bridges can be cabled to ports 18 through 23.

MetroCluster four-node configuration					
FibreBridge 7500 bridge	**Port**	**Brocade 6510 or DCX 8510-8**		**Brocade 6505**	
		FC_switch_x_1	**FC_switch_x_2**	**FC_switch_x_1**	**FC_switch_x_2**
bridge_x_1a	FC1	6	-	6	-
	FC2	-	6	-	6
bridge_x_1b	FC1	7	-	7	-
	FC2	-	7	-	7
bridge_x_2a	FC1	8	-	12	-
	FC2	-	8	-	12
bridge_x_2b	FC1	9	-	13	-
	FC2	-	9	-	13
bridge_x_3a	FC1	10	-	14	-
	FC2	-	10	-	14
bridge_x_3b	FC1	11	-	15	-
	FC2	-	11	-	15
bridge_x_4a	FC1	12	-	16	-
	FC2	-	12	-	16
bridge_x_4b	FC1	13	-	17	-
	FC2	-	13	-	17
		additional bridges can be cabled through port 19, then ports 24 through 47		additional bridges can be cabled through port 23	

The following table shows bridge port usage when using FibreBridge 6500 bridges:

FibreBridge 6500 bridge	**Port**	**Brocade 6510, DCX 8510-8**		**Brocade 6505**	
		FC_switch_x_1	**FC_switch_x_2**	**FC_switch_x_1**	**FC_switch_x_2**
bridge_x_1a	FC1	6	-	6	-
bridge_x_1b	FC1	-	6	-	6
bridge_x_2a	FC1	7	-	7	-
bridge_x_2b	FC1	-	7	-	7
bridge_x_3a	FC1	8	-	12	-
bridge_x_3b	FC1	-	8	-	12
bridge_x_4a	FC1	9	-	13	-

FibreBridge 6500 bridge	Port	Brocade 6510, DCX 8510-8		Brocade 6505	
		FC_switch_x_1	FC_switch_x_2	FC_switch_x_1	FC_switch_x_2
bridge_x_4b	FC1	-	9	-	13
bridge_x_5a	FC1	10	-	14	-
bridge_x_5b	FC1	-	10	-	14
bridge_x_6a	FC1	11	-	15	-
bridge_x_6b	FC1	-	11	-	15
bridge_x_7a	FC1	12	-	16	-
bridge_x_7b	FC1	-	12	-	16
bridge_x_8a	FC1	13	-	17	-
bridge_x_8b	FC1	-	13	-	17
		additional bridges can be cabled through port 19, then ports 24 through 47		additional bridges can be cabled through port 23	

Brocade port usage for ISLs in a four-node MetroCluster configuration running ONTAP 9.0

The following table shows ISL port usage:

MetroCluster four-node configuration				
ISL port	Brocade 6510, DCX 8510-8		Brocade 6505	
	FC_switch_x_1	FC_switch_x_2	FC_switch_x_1	FC_switch_x_2
ISL port 1	20	20	8	8
ISL port 2	21	21	9	9
ISL port 3	22	22	10	10
ISL port 4	23	23	11	11

Brocade port usage for controllers in a two-node MetroCluster configuration running ONTAP 9.0

The cabling is the same for each FC switch in the switch fabric.

MetroCluster two-node configuration			
Component	Port	Brocade 6505, 6510, or DCX 8510-8	
		FC_switch_x_1	FC_switch_x_2
controller_x_1	FC-VI port a	0	-
	FC-VI port b	-	0
	HBA port a	1	-
	HBA port b	-	1
	HBA port c	2	-
	HBA port d	-	2

Brocade port usage for bridges in a two-node MetroCluster configuration running ONTAP 9.0

The cabling is the same for each FC switch in the switch fabric.

The following table shows bridge port usage up to port 17 when using FibreBridge 7500 bridges. Additional bridges can be cabled to ports 18 through 23.

MetroCluster two-node configuration					
FibreBridge 7500 bridge	**Port**	**Brocade 6510, DCX 8510-8**		**Brocade 6505**	
		FC_switch_x_1	**FC_switch_x_2**	**FC_switch_x_1**	**FC_switch_x_2**
bridge_x_1a	FC1	6	-	6	-
	FC2	-	6	-	6
bridge_x_1b	FC1	7	-	7	-
	FC2	-	7	-	7
bridge_x_2a	FC1	8	-	12	-
	FC2	-	8	-	12
bridge_x_2b	FC1	9	-	13	-
	FC2	-	9	-	13
bridge_x_3a	FC1	10	-	14	-
	FC2	-	10	-	14
bridge_x_3b	FC1	11	-	15	-
	FC2	-	11	-	15
bridge_x_4a	FC1	12	-	16	-
	FC2	-	12	-	16
bridge_x_4b	FC1	13	-	17	-
	FC2	-	13	-	17
		additional bridges can be cabled through port 19, then ports 24 through 47		additional bridges can be cabled through port 23	

The following table shows bridge port usage when using FibreBridge 6500 bridges:

MetroCluster two-node configuration					
FibreBridge 6500 bridge	**Port**	**Brocade 6510, DCX 8510-8**		**Brocade 6505**	
		FC_switch_x_1	**FC_switch_x_2**	**FC_switch_x_1**	**FC_switch_x_2**
bridge_x_1a	FC1	6	-	6	-
bridge_x_1b	FC1	-	6	-	6
bridge_x_2a	FC1	7	-	7	-
bridge_x_2b	FC1	-	7	-	7
bridge_x_3a	FC1	8	-	12	-
bridge_x_3b	FC1	-	8	-	12
bridge_x_4a	FC1	9	-	13	-
bridge_x_4b	FC1	-	9	-	13

MetroCluster two-node configuration					
FibreBridge 6500 bridge	**Port**	**Brocade 6510, DCX 8510-8**		**Brocade 6505**	
		FC_switch_x_1	**FC_switch_x_2**	**FC_switch_x_1**	**FC_switch_x_2**
bridge_x_5a	FC1	10	-	14	-
bridge_x_5b	FC1	-	10	-	14
bridge_x_6a	FC1	11	-	15	-
bridge_x_6b	FC1	-	11	-	15
bridge_x_7a	FC1	12	-	16	-
bridge_x_7b	FC1	-	12	-	16
bridge_x_8a	FC1	13	-	17	-
bridge_x_8b	FC1	-	13	-	17
		additional bridges can be cabled through port 19, then ports 24 through 47		additional bridges can be cabled through port 23	

Brocade port usage for ISLs in a two-node MetroCluster configuration running ONTAP 9.0

The following table shows ISL port usage:

MetroCluster two-node configuration				
ISL port	**Brocade 6510, DCX 8510-8**		**Brocade 6505**	
	FC_switch_x_1	**FC_switch_x_2**	**FC_switch_x_1**	**FC_switch_x_2**
ISL port 1	20	20	8	8
ISL port 2	21	21	9	9
ISL port 3	22	22	10	10
ISL port 4	23	23	11	11

Cisco port usage for controllers in an eight-node MetroCluster configuration running ONTAP 9.0

The following table shows controller port usage on Cisco switches:

MetroCluster eight-node configuration			
Component	**Port**	**Cisco 9148 or 9148S**	
		FC_switch_x_1	**FC_switch_x_2**
controller_x_1	FC-VI port a	1	-
	FC-VI port b	-	1
	HBA port a	2	-
	HBA port b	-	2
	HBA port c	3	-
	HBA port d	-	3

MetroCluster eight-node configuration			
Component	**Port**	**Cisco 9148 or 9148S**	
		FC_switch_x_1	**FC_switch_x_2**
controller_x_2	FC-VI port a	4	-
	FC-VI port b	-	4
	HBA port a	5	-
	HBA port b	-	5
	HBA port c	6	-
	HBA port d	-	6
controller_x_3	FC-VI port a	7	
	FC-VI port b	-	7
	HBA port a	8	-
	HBA port b	-	8
	HBA port c	9	-
	HBA port d	-	9
controller_x_4	FC-VI port a	10	-
	FC-VI port b	-	10
	HBA port a	11	-
	HBA port b	-	11
	HBA port c	13	-
	HBA port d	-	13

Cisco port usage for FC-to-SAS bridges in an eight-node MetroCluster configuration running ONTAP 9.0

The following table shows bridge port usage up to port 23 when using FibreBridge 7500 bridges.
Additional bridges can be attached using ports 25 through 48.

MetroCluster eight-node configuration			
FibreBridge 7500 bridge	**Port**	**Cisco 9148 or 9148S**	
		FC_switch_x_1	**FC_switch_x_2**
bridge_x_1a	FC1	14	14
	FC2	-	-
bridge_x_1b	FC1	15	15
	FC2	-	-
bridge_x_2a	FC1	17	17
	FC2	-	-
bridge_x_2b	FC1	18	18
	FC2	-	-

MetroCluster eight-node configuration			
FibreBridge 7500 bridge	**Port**	**Cisco 9148 or 9148S**	
		FC_switch_x_1	**FC_switch_x_2**
bridge_x_3a	FC1	19	19
	FC2	-	-
bridge_x_3b	FC1	21	21
	FC2	-	-
bridge_x_4a	FC1	22	22
	FC2	-	-
bridge_x_4b	FC1	23	23
	FC2	-	-
Additional bridges can be attached using ports 25 through 48 following the same pattern.			

The following table shows bridge port usage up to port 23 when using FibreBridge 6500 bridges. Additional bridges can be attached using ports 25-48.

FibreBridge 6500 bridge	**Port**	**Cisco 9148 or 9148S**	
		FC_switch_x_1	**FC_switch_x_2**
bridge_x_1a	FC1	14	-
bridge_x_1b	FC1	-	14
bridge_x_2a	FC1	15	-
bridge_x_2b	FC1	-	15
bridge_x_3a	FC1	17	-
bridge_x_3b	FC1	-	17
bridge_x_4a	FC1	18	-
bridge_x_4b	FC1	-	18
bridge_x_5a	FC1	19	-
bridge_x_5b	FC1	-	19
bridge_x_6a	FC1	21	-
bridge_x_6b	FC1	-	21
bridge_x_7a	FC1	22	-
bridge_x_7b	FC1	-	22
bridge_x_8a	FC1	23	-
bridge_x_8b	FC1	-	23
Additional bridges can be attached using ports 25 through 48 following the same pattern.			

Cisco port usage for ISLs in an eight-node MetroCluster configuration running ONTAP 9.0

The following table shows ISL port usage:

MetroCluster eight-node configuration		
ISL port	**Cisco 9148 or 9148S**	
	FC_switch_x_1	**FC_switch_x_2**
ISL port 1	12	12
ISL port 2	16	16
ISL port 3	20	20
ISL port 4	24	24

Cisco port usage for controllers in a four-node MetroCluster configuration

The cabling is the same for each FC switch in the switch fabric.

The following table shows controller port usage on Cisco switches:

MetroCluster four-node configuration			
Component	**Port**	**Cisco 9148, 9148S, or 9250i**	
		FC_switch_x_1	**FC_switch_x_2**
controller_x_1	FC-VI port a	1	-
	FC-VI port b	-	1
	HBA port a	2	-
	HBA port b	-	2
	HBA port c	3	-
	HBA port d	-	3
controller_x_2	FC-VI port a	4	-
	FC-VI port b	-	4
	HBA port a	5	-
	HBA port b	-	5
	HBA port c	6	-
	HBA port d	-	6

Cisco port usage for FC-to-SAS bridges in a four-node MetroCluster configuration running ONTAP 9.0

The following table shows bridge port usage up to port 14 when using FibreBridge 7500 bridges.
Additional bridges can be attached to ports 15 through 32 following the same pattern.

MetroCluster four-node configuration			
FibreBridge 7500 bridge	**Port**	**Cisco 9148, 9148S, or 9250i**	
		FC_switch_x_1	**FC_switch_x_2**
bridge_x_1a	FC1	7	-
	FC2	-	7

MetroCluster four-node configuration			
FibreBridge 7500 bridge	**Port**	**Cisco 9148, 9148S, or 9250i**	
		FC_switch_x_1	**FC_switch_x_2**
bridge_x_1b	FC1	8	-
	FC2	-	8
bridge_x_2a	FC1	9	-
	FC2	-	9
bridge_x_2b	FC1	10	-
	FC2	-	10
bridge_x_3a	FC1	11	-
	FC2	-	11
bridge_x_3b	FC1	12	-
	FC2	-	12
bridge_x_4a	FC1	13	-
	FC2	-	13
bridge_x_4b	FC1	14	-
	FC2	-	14

The following table shows bridge port usage when using FibreBridge 6500 bridges up to port 14. Additional bridges can be attached to ports 15 through 32 following the same pattern.

FibreBridge 6500 bridge	**Port**	**Cisco 9148, 9148S, or 9250i**	
		FC_switch_x_1	**FC_switch_x_2**
bridge_x_1a	FC1	7	-
bridge_x_1b	FC1	-	7
bridge_x_2a	FC1	8	-
bridge_x_2b	FC1	-	8
bridge_x_3a	FC1	9	-
bridge_x_3b	FC1	-	9
bridge_x_4a	FC1	10	-
bridge_x_4b	FC1	-	10
bridge_x_5a	FC1	11	-
bridge_x_5b	FC1	-	11
bridge_x_6a	FC1	12	-
bridge_x_6b	FC1	-	12
bridge_x_7a	FC1	13	-
bridge_x_7b	FC1	-	13

FibreBridge 6500 bridge	Port	Cisco 9148, 9148S, or 9250i	
		FC_switch_x_1	FC_switch_x_2
bridge_x_8a	FC1	14	-
bridge_x_8b	FC1	-	14
Additional bridges can be attached to ports 15 through 32 following the same pattern.			

Cisco 9148 and 9148S port usage for ISLs on a four-node MetroCluster configuration running ONTAP 9.0

The cabling is the same for each FC switch in the switch fabric.

The following table shows ISL port usage:

MetroCluster four-node configuration		
ISL port	Cisco 9148 or 9148S	
	FC_switch_x_1	FC_switch_x_2
ISL port 1	36	36
ISL port 2	40	40
ISL port 3	44	44
ISL port 4	48	48

Cisco 9250i port usage for ISLs on a four-node MetroCluster configuration running ONTAP 9.0

The Cisco 9250i switch uses the FCIP ports for the ISL.

Ports 40 through 48 are 10 GbE ports and are not used in the MetroCluster configuration.

Cisco port usage for controllers in a two-node MetroCluster configuration

The cabling is the same for each FC switch in the switch fabric.

The following table shows controller port usage on Cisco switches:

MetroCluster two-node configuration			
Component	Port	Cisco 9148, 9148S, or 9250i	
		FC_switch_x_1	FC_switch_x_2
controller_x_1	FC-VI port a	1	-
	FC-VI port b	-	1
	HBA port a	2	-
	HBA port b	-	2
	HBA port c	3	-
	HBA port d	-	3

Cisco port usage for FC-to-SAS bridges in a two-node MetroCluster configuration running ONTAP 9.0

The following table shows bridge port usage up to port 14 when using FibreBridge 7500 bridges.
Additional bridges can be attached to ports 15 through 32 following the same pattern.

MetroCluster two-node configuration			
FibreBridge 7500 bridge	**Port**	**Cisco 9148, 9148S, or 9250i**	
		FC_switch_x_1	**FC_switch_x_2**
bridge_x_1a	FC1	7	-
	FC2	-	7
bridge_x_1b	FC1	8	-
	FC2	-	8
bridge_x_2a	FC1	9	-
	FC2	-	9
bridge_x_2b	FC1	10	-
	FC2	-	10
bridge_x_3a	FC1	11	-
	FC2	-	11
bridge_x_3b	FC1	12	-
	FC2	-	12
bridge_x_4a	FC1	13	-
	FC2	-	13
bridge_x_4b	FC1	14	-
	FC2	-	14

The following table shows bridge port usage when using FibreBridge 6500 bridges up to port 14. Additional bridges can be attached to ports 15 through 32 following the same pattern.

MetroCluster two-node configuration			
FibreBridge 6500 bridge	**Port**	**Cisco 9148, 9148S, or 9250i**	
		FC_switch_x_1	**FC_switch_x_2**
bridge_x_1a	FC1	7	-
bridge_x_1b	FC1	-	7
bridge_x_2a	FC1	8	-
bridge_x_2b	FC1	-	8
bridge_x_3a	FC1	9	-
bridge_x_3b	FC1	-	9
bridge_x_4a	FC1	10	-
bridge_x_4b	FC1	-	10
bridge_x_5a	FC1	11	-
bridge_x_5b	FC1	-	11
bridge_x_6a	FC1	12	-

MetroCluster two-node configuration			
FibreBridge 6500 bridge	**Port**	**Cisco 9148, 9148S, or 9250i**	
		FC_switch_x_1	**FC_switch_x_2**
bridge_x_6b	FC1	-	12
bridge_x_7a	FC1	13	-
bridge_x_7b	FC1	-	13
bridge_x_8a	FC1	14	-
bridge_x_8b	FC1	-	14
Additional bridges can be attached to ports 15 through 32 following the same pattern.			

Cisco 9148 or 9148S port usage for ISLs on a two-node MetroCluster configuration running ONTAP 9.0

The cabling is the same for each FC switch in the switch fabric.

The following table shows ISL port usage:

MetroCluster two-node configuration		
ISL port	**Cisco 9148 or 9148S**	
	FC_switch_x_1	**FC_switch_x_2**
ISL port 1	36	36
ISL port 2	40	40
ISL port 3	44	44
ISL port 4	48	48

Cisco 9250i port usage for ISLs on a two-node MetroCluster configuration running ONTAP 9.0

The Cisco 9250i switch uses the FCIP ports for the ISL.

Ports 40 through 48 are 10 GbE ports and are not used in the MetroCluster configuration.

Port assignments for FC switches when using ONTAP 9.1 and later

You need to verify that you are using the specified port assignments when you cable the FC switches when using ONTAP 9.1 and later.

Ports that are not used for attaching initiator ports, FC-VI ports, or ISLs can be reconfigured to act as storage ports. However, if the supported RCFs are being used, the zoning must be changed accordingly.

If the supported RCFs are used, ISL ports might not connect to the same ports shown here and might need to be reconfigured manually.

If you configured your switches using the port assignments for ONTAP 9, you can continue to use the older assignments. However, new configurations running ONTAP 9.1 or later releases should use the port assignments shown here.

Overall cabling guidelines

You should be aware of the following guidelines when using the cabling tables:

- The Brocade and Cisco switches use different port numbering:
  - On Brocade switches, the first port is numbered 0.
  - On Cisco switches, the first port is numbered 1.

- The cabling is the same for each FC switch in the switch fabric.
- AFF A300 and FAS8200 storage systems can be ordered with one of two options for FC-VI connectivity:
  - Onboard ports 0e and 0f configured in FC-VI mode.
  - Ports 1a and 1b on an FC-VI card in slot 1.
- AFF A700 and FAS9000 storage systems require four FC-VI ports. The following tables show cabling for the FC switches with four FC-VI ports on each controller except for the Cisco 9250i switch.

  For other storage systems, use the cabling shown in the tables but ignore the cabling for FC-VI ports c and d.

  You can leave those ports empty.
- AFF A400 and FAS8300 storage systems use ports 2a and 2b for FC-VI connectivity.
- If you have two MetroCluster configurations sharing ISLs, use the same port assignments as that for an eight-node MetroCluster cabling.

  The number of ISLs you cable may vary depending on your site's requirements.

Brocade port usage for controllers in a MetroCluster configuration running ONTAP 9.1 or later

The following tables show port usage on Brocade switches. The tables show the maximum supported configuration, with eight controller modules in two DR groups. For smaller configurations, ignore the rows for the additional controller modules. Note that eight ISLs are supported only on the Brocade 6510, Brocade DCX 8510-8, G620, G630, G620-1, G630-1, and G720 switches.

Note: Port usage for the Brocade 6505 and Brocade G610 switches in an eight-node MetroCluster configuration is not shown. Due to the limited number of ports, port assignments must be made on a site-by-site basis depending on the controller module model and the number of ISLs and bridge pairs in use.

Note: The Brocade DCX 8510-8 switch can use the same port layout as the 6510 switch **or** the 7840 switch.

Configurations using FibreBridge 6500N bridges or FibreBridge 7500N or 7600N using one FC port (FC1 or FC2) only				
MetroCluster 1 or DR Group 1				
Component	**Port**	**Brocade switch models 6505, 6510, 6520, 7810, 7840, G610, G620, G620-1, G630, G630-1 and DCX 8510-8**		**Brocade switch model G720**
		Connects to FC switch...	**Connects to switch port...**	**Connects to switch port...**
controller_x_1	FC-VI port a	1	0	0
	FC-VI port b	2	0	0
	FC-VI port c	1	1	1
	FC-VI port d	2	1	1
	HBA port a	1	2	8
	HBA port b	2	2	8
	HBA port c	1	3	9
	HBA port d	2	3	9

Configurations using FibreBridge 6500N bridges or FibreBridge 7500N or 7600N using one FC port (FC1 or FC2) only				
MetroCluster 1 or DR Group 1				
Component	**Port**	**Brocade switch models 6505, 6510, 6520, 7810, 7840, G610, G620, G620-1, G630, G630-1 and DCX 8510-8**		**Brocade switch model G720**
		Connects to FC switch...	**Connects to switch port...**	**Connects to switch port...**
controller_x_2	FC-VI port a	1	4	4
	FC-VI port b	2	4	4
	FC-VI port c	1	5	5
	FC-VI port d	2	5	5
	HBA port a	1	6	12
	HBA port b	2	6	12
	HBA port c	1	7	13
	HBA port d	2	7	13
Stack 1	bridge_x_1a	1	8	10
	bridge_x_1b	2	8	10
Stack 2	bridge_x_2a	1	9	11
	bridge_x_2b	2	9	11
Stack 3	bridge_x_3a	1	10	14
	bridge_x_4b	2	10	14
Stack y	bridge_x_*y*a	1	11	15
	bridge_x_*y*b	2	11	15

Note:

- On G620, G630, G620-1 and G630-1 switches, additional bridges can be cabled to ports 12 - 17, 20 and 21.
- On G610 switches, additional bridges can be cabled to ports 12 - 19.
- On G720 switches, additional bridges can be cabled to ports 16 - 17, 20 and 21.

Configurations using FibreBridge 6500N bridges or FibreBridge 7500N or 7600N using one FC port (FC1 or FC2) only							
MetroCluster 2 or DR Group 2							
			Brocade switch model				
Component	**Port**	**Connects to FC_switch...**	**6510, DCX 8510-8**	**6520**	**7840, DCX 8510-8**	**G620, G620-1, G630, G630-1**	**G720**
controller_x_3	FC-VI port a	1	24	48	12	18	18
	FC-VI port b	2	24	48	12	18	18
	FC-VI port c	1	25	49	13	19	19
	FC-VI port d	2	25	49	13	19	19
	HBA port a	1	26	50	14	24	26
	HBA port b	2	26	50	14	24	26
	HBA port c	1	27	51	15	25	27
	HBA port d	2	27	51	15	25	27
controller_x_4	FC-VI port a	1	28	52	16	22	22
	FC-VI port b	2	28	52	16	22	22
	FC-VI port c	1	29	53	17	23	23
	FC-VI port d	2	29	53	17	23	23
	HBA port a	1	30	54	18	28	30
	HBA port b	2	30	54	18	28	30
	HBA port c	1	31	55	19	29	31
	HBA port d	2	32	55	19	29	31
Stack 1	bridge_x_51a	1	32	56	20	26	32
	bridge_x_51b	2	32	56	20	26	32
Stack 2	bridge_x_52a	1	33	57	21	27	33
	bridge_x_52b	2	33	57	21	27	33
Stack 3	bridge_x_53a	1	34	58	22	30	34
	bridge_x_54b	2	34	58	22	30	34
Stack *y*	bridge_x_*y*a	1	35	59	23	31	35
	bridge_x_*y*b	2	35	59	23	31	35

Note:

- On G720 switches, additional bridges can be cabled to ports 36-39.

Configurations using FibreBridge 7500N or 7600N using both FC ports (FC1 and FC2)					
MetroCluster 1 or DR Group 1					
Component		**Port**	**Brocade switch models 6505, 6510, 6520, 7810, 7840, G610, G620, G620-1, G630, G630-1, and DCX 8510-8**		**Brocade switch G720**
			Connects to FC_switch...	**Connects to switch port...**	**Connects to switch port...**
controller_x_1		FC-VI port a	1	0	0
		FC-VI port b	2	0	0
		FC-VI port c	1	1	1
		FC-VI port d	2	1	1
		HBA port a	1	2	8
		HBA port b	2	2	8
		HBA port c	1	3	9
		HBA port d	2	3	9
controller_x_2		FC-VI port a	1	4	4
		FC-VI port b	2	4	4
		FC-VI port c	1	5	5
		FC-VI port d	2	5	5
		HBA port a	1	6	12
		HBA port b	2	6	12
		HBA port c	1	7	13
		HBA port d	2	7	13
Stack 1	bridge_x_1a	FC1	1	8	10
		FC2	2	8	10
	bridge_x_1B	FC1	1	9	11
		FC2	2	9	11
Stack 2	bridge_x_2a	FC1	1	10	14
		FC2	2	10	14
	bridge_x_2B	FC1	1	11	15
		FC2	2	11	15
Stack 3	bridge_x_3a	FC1	1	12*	16
		FC2	2	12*	16
	bridge_x_3B	FC1	1	13*	17
		FC2	2	13*	17

Configurations using FibreBridge 7500N or 7600N using both FC ports (FC1 and FC2)					
MetroCluster 1 or DR Group 1					
Component		**Port**	**Brocade switch models 6505, 6510, 6520, 7810, 7840, G610, G620, G620-1, G630, G630-1, and DCX 8510-8**		**Brocade switch G720**
			Connects to FC_switch...	**Connects to switch port...**	**Connects to switch port...**
Stack *y*	bridge_x_*y*a	FC1	1	14*	20
		FC2	2	14*	20
	bridge_x_*y*b	FC1	1	15*	21
		FC2	2	15*	21
* - Ports 12 through 15 are reserved for the second MetroCluster or DR group on the Brocade 7840 switch. **Note:** Additional bridges can be cabled to ports 16, 17, 20 and 21 in G620, G630, G620-1 and G630-1 switches.					

Configurations using FibreBridge 7500N or 7600N using both FC ports (FC1 and FC2)							
MetroCluster 2 or DR Group 2							
Component	**Port**	**Brocade switch model**					
		Connects to FC_switch...	**6510, DCX 8510-8**	**6520**	**7840, DCX 8510-8**	**G620, G620-1, G630, G630-1**	**G720**
controller_x_3	FC-VI port a	1	24	48	12	18	18
	FC-VI port b	2	24	48	12	18	18
	FC-VI port c	1	25	49	13	19	19
	FC-VI port d	2	25	49	13	19	19
	HBA port a	1	26	50	14	24	26
	HBA port b	2	26	50	14	24	26
	HBA port c	1	27	51	15	25	27
	HBA port d	2	27	51	15	25	27

Configurations using FibreBridge 7500N or 7600N using both FC ports (FC1 and FC2)								
MetroCluster 2 or DR Group 2								
Component		**Port**	**Brocade switch model**					
			Connects to FC_switch...	**6510, DCX 8510-8**	**6520**	**7840, DCX 8510-8**	**G620, G620-1, G630, G630-1**	**G720**
controller_x_4		FC-VI port a	1	28	52	16	22	22
		FC-VI port b	2	28	52	16	22	22
		FC-VI port c	1	29	53	17	23	23
		FC-VI port d	2	29	53	17	23	23
		HBA port a	1	30	54	18	28	30
		HBA port b	2	30	54	18	28	30
		HBA port c	1	31	55	19	29	31
		HBA port d	2	31	55	19	29	31
Stack 1	bridge_x_51a	FC1	1	32	56	20	26	32
		FC2	2	32	56	20	26	32
	bridge_x_51b	FC1	1	33	57	21	27	33
		FC2	2	33	57	21	27	33
Stack 2	bridge_x_52a	FC1	1	34	58	22	30	34
		FC2	2	34	58	22	30	34
	bridge_x_52b	FC1	1	35	59	23	31	35
		FC2	2	35	59	23	31	35
Stack 3	bridge_x_53a	FC1	1	36	60	-	32	36
		FC2	2	36	60	-	32	36
	bridge_x_53b	FC1	1	37	61	-	33	37
		FC2	2	37	61	-	33	37
Stack *y*	bridge_x_5*y*a	FC1	1	38	62	-	34	38
		FC2	2	38	62	-	34	38
	bridge_x_5*y*b	FC1	1	39	63	-	35	39
		FC2	2	39	63	-	35	39

Configurations using FibreBridge 7500N or 7600N using both FC ports (FC1 and FC2)							
MetroCluster 2 or DR Group 2							
Component	**Port**	**Brocade switch model**					
		Connects to FC_switch...	**6510, DCX 8510-8**	**6520**	**7840, DCX 8510-8**	**G620, G620-1, G630, G630-1**	**G720**
Note: Additional bridges can be cabled to ports 36 to 39 in G620, G630, G620-1, and G630-1 switches.							

Brocade port usage for ISLs in a MetroCluster configuration running ONTAP 9.1 or later

The following table shows ISL port usage for the Brocade switches.

Note: AFF A700 or FAS9000 systems support up to eight ISLs for improved performance. Eight ISLs are supported on the Brocade 6510 and G620 switches.

Switch model	ISL port	Switch port
Brocade 6520	ISL port 1	23
	ISL port 2	47
	ISL port 3	71
	ISL port 4	95
Brocade 6505	ISL port 1	20
	ISL port 2	21
	ISL port 3	22
	ISL port 4	23
Brocade 6510 and Brocade DCX 8510-8	ISL port 1	40
	ISL port 2	41
	ISL port 3	42
	ISL port 4	43
	ISL port 5	44
	ISL port 6	45
	ISL port 7	46
	ISL port 8	47
Brocade 7810	ISL port 1	ge2 (10-Gbps)
	ISL port 2	ge3(10-Gbps)
	ISL port 3	ge4 (10-Gbps)
	ISL port 4	ge5 (10-Gbps)
	ISL port 5	ge6 (10-Gbps)
	ISL port 6	ge7 (10-Gbps)

Switch model	ISL port	Switch port
Brocade 7840 **Note:** The Brocade 7840 switch supports either two 40 Gbps VE-ports or up to four 10 Gbps VE-ports per switch for the creation of FCIP ISLs.	ISL port 1	ge0 (40-Gbps) or ge2 (10-Gbps)
	ISL port 2	ge1 (40-Gbps) or ge3 (10-Gbps)
	ISL port 3	ge10 (10-Gbps)
	ISL port 4	ge11 (10-Gbps)
Brocade G610	ISL port 1	20
	ISL port 2	21
	ISL port 3	22
	ISL port 4	23
Brocade G620, G620-1, G630, G630-1, G720	ISL port 1	40
	ISL port 2	41
	ISL port 3	42
	ISL port 4	43
	ISL port 5	44
	ISL port 6	45
	ISL port 7	46
	ISL port 8	47

Cisco port usage for controllers in a MetroCluster configuration running ONTAP 9.4 or later

The tables show the maximum supported configuration, with eight controller modules in two DR groups. For smaller configurations, ignore the rows for the additional controller modules.

Cisco 9396S			
Component	**Port**	**Switch 1**	**Switch 2**
controller_x_1	FC-VI port a	1	-
	FC-VI port b	-	1
	FC-VI port c	2	-
	FC-VI port d	-	2
	HBA port a	3	-
	HBA port b	-	3
	HBA port c	4	-
	HBA port d	-	4

Cisco 9396S			
Component	**Port**	**Switch 1**	**Switch 2**
controller_x_2	FC-VI port a	5	-
	FC-VI port b	-	5
	FC-VI port c	6	-
	FC-VI port d	-	6
	HBA port a	7	-
	HBA port b	-	7
	HBA port c	8	-
	HBA port d	-	8
controller_x_3	FC-VI port a	49	
	FC-VI port b	-	49
	FC-VI port c	50	
	FC-VI port d	-	50
	HBA port a	51	
	HBA port b	-	51
	HBA port c	52	
	HBA port d	-	52
controller_x_4	FC-VI port a	53	-
	FC-VI port b	-	53
	FC-VI port c	54	-
	FC-VI port d	-	54
	HBA port a	55	-
	HBA port b	-	55
	HBA port c	56	-
	HBA port d	-	56

Cisco 9148S			
Component	**Port**	**Switch 1**	**Switch 2**
controller_x_1	FC-VI port a	1	-
	FC-VI port b	-	1
	FC-VI port c	2	-
	FC-VI port d	-	2
	HBA port a	3	-
	HBA port b	-	3
	HBA port c	4	-
	HBA port d	-	4
controller_x_2	FC-VI port a	5	-
	FC-VI port b	-	5
	FC-VI port c	6	-
	FC-VI port d	-	6
	HBA port a	7	-
	HBA port b	-	7
	HBA port c	8	-
	HBA port d	-	8
controller_x_3	FC-VI port a	25	
	FC-VI port b	-	25
	FC-VI port c	26	-
	FC-VI port d	-	26
	HBA port a	27	-
	HBA port b	-	27
	HBA port c	28	-
	HBA port d	-	28
controller_x_4	FC-VI port a	29	-
	FC-VI port b	-	29
	FC-VI port c	30	-
	FC-VI port d	-	30
	HBA port a	31	-
	HBA port b	-	31
	HBA port c	32	-
	HBA port d	-	32

Cisco 9132T			
MDS module 1			
Component	**Port**	**Switch 1**	**Switch 2**
controller_x_1	FC-VI port a	1	-
	FC-VI port b	-	1
	FC-VI port c	2	-
	FC-VI port d	-	2
	HBA port a	3	-
	HBA port b	-	3
	HBA port c	4	-
	HBA port d	-	4
controller_x_2	FC-VI port a	5	-
	FC-VI port b	-	5
	FC-VI port c	6	-
	FC-VI port d	-	6
	HBA port a	7	-
	HBA port b	-	7
	HBA port c	8	-
	HBA port d	-	8
MDS module 2			
Component	Port	Switch 1	Switch 2
controller_x_3	FC-VI port a	1	-
	FC-VI port b	-	1
	FC-VI port c	2	-
	FC-VI port d	-	2
	HBA port a	3	-
	HBA port b	-	3
	HBA port c	4	-
	HBA port d	-	4

Cisco 9132T			
MDS module 1			
Component	**Port**	**Switch 1**	**Switch 2**
controller_x_4	FC-VI port a	5	-
	FC-VI port b	-	5
	FC-VI port c	6	-
	FC-VI port d	-	6
	HBA port a	7	-
	HBA port b	-	7
	HBA port c	8	-
	HBA port d	-	8

Note: The following table shows systems with two FC-VI ports. AFF A700 and FAS9000 systems have four FC-VI ports (a, b, c, and d). If using an AFF A700 or FAS9000 system, the port assignments move along by one position. For example, FC-VI ports c and d go to switch port 2 and HBA ports a and b go to switch port 3.

Cisco 9250i*			
Component	**Port**	**Switch 1**	**Switch 2**
controller_x_1	FC-VI port a	1	-
	FC-VI port b	-	1
	HBA port a	2	-
	HBA port b	-	2
	HBA port c	3	-
	HBA port d	-	3
controller_x_2	FC-VI port a	4	-
	FC-VI port b	-	4
	HBA port a	5	-
	HBA port b	-	5
	HBA port c	6	-
	HBA port d	-	6
controller_x_3	FC-VI port a	7	-
	FC-VI port b	-	7
	HBA port a	8	-
	HBA port b	-	8
	HBA port c	9	-
	HBA port d	-	9

Cisco 9250i*			
Component	**Port**	**Switch 1**	**Switch 2**
controller_x_4	FC-VI port a	10	-
	FC-VI port b	-	10
	HBA port a	11	-
	HBA port b	-	11
	HBA port c	13	-
	HBA port d	-	13

* - The Cisco 9250i switch is not supported for eight-node MetroCluster configurations.

Cisco port usage for FC-to-SAS bridges in a MetroCluster configuration running ONTAP 9.1 or later

Cisco 9396S			
FibreBridge 7500 using two FC ports	**Port**	**Switch 1**	**Switch 2**
bridge_x_1a	FC1	9	-
	FC2	-	9
bridge_x_1b	FC1	10	-
	FC2	-	10
bridge_x_2a	FC1	11	-
	FC2	-	11
bridge_x_2b	FC1	12	-
	FC2	-	12
bridge_x_3a	FC1	13	-
	FC2	-	13
bridge_x_3b	FC1	14	-
	FC2	-	14
bridge_x_4a	FC1	15	-
	FC2	-	15
bridge_x_4b	FC1	16	-
	FC2	-	16
		Additional bridges can be attached using ports 17 through 40 and 57 through 88 following the same pattern.	

Cisco 9148S			
FibreBridge 7500 using two FC ports	**Port**	**Switch 1**	**Switch 2**
bridge_x_1a	FC1	9	-
	FC2	-	9

Cisco 9148S			
FibreBridge 7500 using two FC ports	**Port**	**Switch 1**	**Switch 2**
bridge_x_1b	FC1	10	-
	FC2	-	10
bridge_x_2a	FC1	11	-
	FC2	-	11
bridge_x_2b	FC1	12	-
	FC2	-	12
bridge_x_3a	FC1	13	-
	FC2	-	13
bridge_x_3b	FC1	14	-
	FC2	-	14
bridge_x_4a	FC1	15	-
	FC2	-	15
bridge_x_4b	FC1	16	-
	FC2	-	16
		Additional bridges for a second DR group or second MetroCluster configuration can be attached using ports 33 through 40 following the same pattern.	

Cisco 9132T			
FibreBridge 7500 using two FC ports	**Port**	**Switch 1**	**Switch 2**
bridge_x_1a	FC1	9	-
	FC2	-	9
bridge_x_1b	FC1	10	-
	FC2	-	10
bridge_x_2a	FC1	11	-
	FC2	-	11
bridge_x_2b	FC1	12	-
	FC2	-	12
		Additional bridges for a second DR group or second MetroCluster configuration can be attached using the same port numbers on the second MDS module.	

Cisco 9250i			
FibreBridge 7500 using two FC ports	**Port**	**Switch 1**	**Switch 2**
bridge_x_1a	FC1	14	-
	FC2	-	14
bridge_x_1b	FC1	15	-
	FC2	-	15
bridge_x_2a	FC1	17	-
	FC2	-	17
bridge_x_2b	FC1	18	-
	FC2	-	18
bridge_x_3a	FC1	19	-
	FC2	-	19
bridge_x_3b	FC1	21	-
	FC2	-	21
bridge_x_4a	FC1	22	-
	FC2	-	22
bridge_x_4b	FC1	23	-
	FC2	-	23
		Additional bridges for a second DR group or second MetroCluster configuration can be attached using ports 25 through 48 following the same pattern.	

The following tables show bridge port usage when using FibreBridge 6500 bridges or FibreBridge 7500 bridges using one FC port (FC1 or FC2) only. For FibreBridge 7500 bridges using one FC port, either FC1 or FC2 can be cabled to the port indicated as FC1. Additional bridges can be attached using ports 25-48.

FibreBridge 6500 bridge or FibreBridge 7500 using one FC port	**Port**	**Cisco 9396S**	
		Switch 1	**Switch 2**
bridge_x_1a	FC1	9	-
bridge_x_1b	FC1	-	9
bridge_x_2a	FC1	10	-
bridge_x_2b	FC1	-	10
bridge_x_3a	FC1	11	-
bridge_x_3b	FC1	-	11
bridge_x_4a	FC1	12	-
bridge_x_4b	FC1	-	12
bridge_x_5a	FC1	13	-

FibreBridge 6500 bridge or FibreBridge 7500 using one FC port	Port	Cisco 9396S	
		Switch 1	Switch 2
bridge_x_5b	FC1	-	13
bridge_x_6a	FC1	14	-
bridge_x_6b	FC1	-	14
bridge_x_7a	FC1	15	-
bridge_x_7b	FC1	-	15
bridge_x_8a	FC1	16	-
bridge_x_8b	FC1	-	16
		Additional bridges can be attached using ports 17 through 40 and 57 through 88 following the same pattern.	

FibreBridge 6500 bridge or FibreBridge 7500 using one FC port	Port	Cisco 9148S	
		Switch 1	Switch 2
bridge_x_1a	FC1	9	-
bridge_x_1b	FC1	-	9
bridge_x_2a	FC1	10	-
bridge_x_2b	FC1	-	10
bridge_x_3a	FC1	11	-
bridge_x_3b	FC1	-	11
bridge_x_4a	FC1	12	-
bridge_x_4b	FC1	-	12
bridge_x_5a	FC1	13	-
bridge_x_5b	FC1	-	13
bridge_x_6a	FC1	14	-
bridge_x_6b	FC1	-	14
bridge_x_7a	FC1	15	-
bridge_x_7b	FC1	-	15
bridge_x_8a	FC1	16	-
bridge_x_8b	FC1	-	16
		Additional bridges for a second DR group or second MetroCluster configuration can be attached using ports 25 through 48 following the same pattern.	

Cisco 9250i			
FibreBridge 6500 bridge or FibreBridge 7500 using one FC port	**Port**	**Switch 1**	**Switch 2**
bridge_x_1a	FC1	14	-
bridge_x_1b	FC1	-	14
bridge_x_2a	FC1	15	-
bridge_x_2b	FC1	-	15
bridge_x_3a	FC1	17	-
bridge_x_3b	FC1	-	17
bridge_x_4a	FC1	18	-
bridge_x_4b	FC1	-	18
bridge_x_5a	FC1	19	-
bridge_x_5b	FC1	-	19
bridge_x_6a	FC1	21	-
bridge_x_6b	FC1	-	21
bridge_x_7a	FC1	22	-
bridge_x_7b	FC1	-	22
bridge_x_8a	FC1	23	-
bridge_x_8b	FC1	-	23
		Additional bridges can be attached using ports 25 through 48 following the same pattern.	

Cisco port usage for ISLs in an eight-node configuration in a MetroCluster configuration running ONTAP 9.1 or later

The following table shows ISL port usage. ISL port usage is the same on all switches in the configuration.

Switch model	ISL port	Switch port
Cisco 9396S	ISL 1	44
	ISL 2	48
	ISL 3	92
	ISL 4	96
Cisco 9250i with 24 port license	ISL 1	12
	ISL 2	16
	ISL 3	20
	ISL 4	24

Switch model	ISL port	Switch port
Cisco 9148S	ISL 1	20
	ISL 2	24
	ISL 3	44
	ISL 4	48
Cisco 9132T	ISL 1	MDS module 1 port 13
	ISL 2	MDS module 1 port 14
	ISL 3	MDS module 1 port 15
	ISL 4	MDS module 1 port 16

Verifying the storage configuration

You must confirm that all storage is visible from the surviving nodes.

Steps

1. Confirm that all storage components at the disaster site are the same in quantity and type at the surviving site.

   The surviving site and disaster site should have the same number of disk shelf stacks, disk shelves, and disks. In a bridge-attached or fabric-attached MetroCluster configuration, the sites should have the same number of FC-to-SAS bridges.
2. Confirm that all disks that have been replaced at the disaster site are unowned: `run local disk show -n`

   Disks should appear as being `unowned`.
3. If no disks were replaced, confirm that all disks are present: `disk show`

Powering on the equipment at the disaster site

You must power on the MetroCluster components at the disaster site when you are ready to prepare for switchback. In addition, you must also recable the SAS storage connections in direct-attached MetroCluster configurations and enable non-Inter-Switch Link ports in fabric-attached MetroCluster configurations.

Before you begin

You must have already replaced and cabled the MetroCluster components exactly as the old ones.

Fabric-attached MetroCluster installation and configuration

Stretch MetroCluster installation and configuration

About this task

The examples in this procedure assume the following:

- Site A is the disaster site.
- FC_switch_A_1 has been replaced.
- FC_switch_A_2 has been replaced.
- Site B is the surviving site.
- FC_switch_B_1 is healthy.
- FC_switch_B_2 is healthy.

The FC switches are present only in fabric-attached MetroCluster configurations.

Steps

1. In a stretch MetroCluster configuration using SAS cabling (and no FC switch fabric or FC-to-SAS bridges), connect all the storage including the remote storage across both sites.

   The controller at the disaster site must remain powered off or at the LOADER prompt.

2. On the surviving site, disable disk autoassignment:

   storage disk option modify -autoassign off *

   ```
   cluster_B::> storage disk option modify -autoassign off *
   2 entries were modified.
   ```

3. On the surviving site, confirm that disk autoassignment is off:

   storage disk option show

   ```
    cluster_B::> storage disk option show
    Node     BKg. FW. Upd.  Auto Copy   Auto Assign  Auto Assign Policy
   --------  -------------  -----------  -----------  ------------------
   node_B_1       on            on          off             default
   node_B_2       on            on          off             default
   2 entries were displayed.

    cluster_B::>
   ```

4. Turn on the disk shelves at the disaster site and make sure that all disks are running.
5. In a bridge-attached or fabric-attached MetroCluster configuration, turn on all FC-to-SAS bridges at the disaster site.
6. If any disks were replaced, leave the controllers powered off or at the LOADER prompt.
7. In a fabric-attached MetroCluster configuration, enable the non-ISL ports on the FC switches.

If the switch vendor is...	Then use these steps to enable the ports...

Brocade	**a.** Persistently enable the ports connected to the FC-to-SAS bridges:
	portpersistentenable *port-number*
	In the following example, ports 6 and 7 are enabled:
	`FC_switch_A_1:admin> portpersistentenable 6`
	`FC_switch_A_1:admin> portpersistentenable 7`
	`FC_switch_A_1:admin>`
	b. Persistently enable the ports connected to the HBAs and FC-VI adapters:
	portpersistentenable *port-number*
	In the following example, ports 6 and 7 are enabled:
	`FC_switch_A_1:admin> portpersistentenable 1`
	`FC_switch_A_1:admin> portpersistentenable 2`
	`FC_switch_A_1:admin> portpersistentenable 4`
	`FC_switch_A_1:admin> portpersistentenable 5`
	`FC_switch_A_1:admin>`
	Note: For AFF A700 and FAS9000 systems, you must persistently enable all four FC-VI ports by using the `switchcfgpersistentenable` command.
	c. Repeat substeps a and b for the second FC switch at the surviving site.

If the switch vendor is...	Then use these steps to enable the ports...

Cisco	**a.** Enter configuration mode for the interface, and then enable the ports with the `no shut` command.
	In the following example, port fc1/36 is disabled:
	`FC_switch_A_1# conf t`
	`FC_switch_A_1(config)# interface fc1/36`
	`FC_switch_A_1(config)# no shut`
	`FC_switch_A_1(config-if)# end`
	`FC_switch_A_1# copy running-config startup-config`
	b. Verify that the switch port is enabled:
	`show interface brief`
	c. Repeat substeps a and b on the other ports connected to the FC-to-SAS bridges, HBAs, and FC-VI adapters.
	d. Repeat substeps a, b, and c for the second FC switch at the surviving site.

Assigning ownership for replaced drives

If you replaced drives when restoring hardware at the disaster site or you had to zero drives or remove ownership, you must assign ownership to the affected drives.

Before you begin

The disaster site must have at least as many available drives as it did prior to the disaster.

The drives shelves and drives arrangement must meet the requirements in the "Required MetroCluster IP components and naming conventions" section of the *MetroCluster IP Installation and Configuration Guide*.

MetroCluster IP installation and configuration

About this task

These steps are performed on the cluster at the disaster site.

This procedure shows the reassignment of all drives and the creation of new plexes at the disaster site. The new plexes are remote plexes of surviving site and local plexes of disaster site.

This section provides examples for two and four-node configurations. For two-node configurations, you can ignore references to the second node at each site. For eight-node configurations, you must account for the additional nodes on the second DR group. The examples make the following assumptions:

- Site A is the disaster site.
- node_A_1 has been replaced.
- node_A_2 has been replaced.
  Present only in four-node MetroCluster configurations.
- Site B is the surviving site.
- node_B_1 is healthy.
- node_B_2 is healthy.
  Present only in four-node MetroCluster configurations.

The controller modules have the following original system IDs:

Number of nodes in MetroCluster configuration	Node	Original system ID
Four	node_A_1	4068741258
	node_A_2	4068741260
	node_B_1	4068741254
	node_B_2	4068741256
Two	node_A_1	4068741258
	node_B_1	4068741254

You should keep in mind the following points when assigning the drives:

- The *old-count-of-disks* must be at least the same number of disks for each node that were present before the disaster.
  If a lower number of disks is specified or present, the healing operations might not be completed due to insufficient space.
- The new plexes to be created are remote plexes belonging to the surviving site (node_B_x pool1) and local plexes belonging to the disaster site (node_B_x pool0).
- The total number of required drives should not include the root aggr disks.
  If *n* disks are assigned to pool1 of the surviving site, then (*n*-3) disks should be assigned to the disaster site with the assumption that the root aggregate uses three disks.
- None of the disks can be assigned to a pool that is different from the one to which all other disks on the same stack are assigned.
- Disks belonging to the surviving site are assigned to pool 1 and disks belonging to the disaster site are assigned to pool 0.

Steps

1. Assign the new, unowned drives based on whether you have a four-node or two-node MetroCluster configuration:

   - For four-node MetroCluster configurations, assign the new, unowned disks to the appropriate disk pools by using the following series of commands on the replacement nodes:

     a. Systematically assign the replaced disks for each node to their respective disk pools:

     disk assign -s *sysid* -n *old-count-of-disks* -p *pool*

     From the surviving site, you issue a `disk assign` command for each node:

     ```
     cluster_B::> disk assign -s node_B_1-sysid -n old-count-of-disks -p 1 (remote
     pool of surviving site)
     cluster_B::> disk assign -s node_B_2-sysid -n old-count-of-disks -p 1 (remote
     pool of surviving site)
     cluster_B::> disk assign -s node_A_1-old-sysid -n old-count-of-disks -p 0 (local
     pool of disaster site)
     cluster_B::> disk assign -s node_A_2-old-sysid -n old-count-of-disks -p 0 (local
     pool of disaster site)
     ```

     The following example shows the commands with the system IDs:

     ```
     cluster_B::> disk assign -s 4068741254 -n 21 -p 1
     cluster_B::> disk assign -s 4068741256 -n 21 -p 1
     cluster_B::> disk assign -s 4068741258 -n 21 -p 0
     cluster_B::> disk assign -s 4068741260 -n 21 -p 0
     ```

     b. Confirm the ownership of the disks:

storage disk show -fields owner, pool

```
storage disk show -fields owner, pool
cluster_A::> storage disk show -fields owner, pool
disk     owner          pool
-------- ------------- -----
0c.00.1  node_A_1      Pool0
0c.00.2  node_A_1      Pool0
.
.
.
0c.00.8  node_A_1      Pool1
0c.00.9  node_A_1      Pool1
.
.
.
0c.00.15 node_A_2      Pool0
0c.00.16 node_A_2      Pool0
.
.
.
0c.00.22 node_A_2      Pool1
0c.00.23 node_A_2      Pool1
.
.
.
```

- For two-node MetroCluster configurations, assign the new, unowned disks to the appropriate disk pools by using the following series of commands on the replacement node:

  a. Display the local shelf IDs:

  run local storage show shelf

  b. Assign the replaced disks for the healthy node to pool 1:

  run local disk assign -shelf *shelf-id* -n *old-count-of-disks* -p 1 -s *node_B_1-sysid* -f

  c. Assign the replaced disks for the replacement node to pool 0:

  run local disk assign -shelf *shelf-id* -n *old-count-of-disks* -p 0 -s *node_A_1-sysid* -f

2. On the surviving site, turn on automatic disk assignment again:

   storage disk option modify -autoassign on *

```
cluster_B::> storage disk option modify -autoassign on *
2 entries were modified.
```

3. On the surviving site, confirm that automatic disk assignment is on:

   storage disk option show

```
 cluster_B::> storage disk option show
 Node     BKg. FW. Upd.  Auto Copy   Auto Assign  Auto Assign Policy
--------  -------------  -----------  -----------  ------------------
node_B_1       on            on          on             default
node_B_2       on            on          on             default
2 entries were displayed.

 cluster_B::>
```

Related concepts

How MetroCluster configurations use SyncMirror to provide data redundancy on page 8

Mirrored aggregates using SyncMirror functionality provide data redundancy and contain the volumes owned by the source and destination storage virtual machine (SVM). Data is replicated into disk pools on the partner cluster. Unmirrored aggregates are also supported.

Related information

Disk and aggregate management

Performing aggregate healing and restoring mirrors (MetroCluster FC configurations)

After replacing hardware and assigning disks, you can perform the MetroCluster healing operations. You must then confirm that aggregates are mirrored and, if necessary, restart mirroring.

Steps

1. Perform the two phases of healing (aggregate healing and root healing) on the disaster site:

```
cluster_B::> metrocluster heal -phase aggregates

cluster_B::> metrocluster heal -phase root aggregates
```

2. Monitor the healing and verify that the aggregates are in either the `resyncing` or `mirrored` state:

`storage aggregate show -node local`

If the aggregate shows this state...	Then...
`resyncing`	No action is required. Let the aggregate complete resyncing.
`mirror degraded`	Proceed to step *3*.
`mirrored, normal`	No action is required.
`unknown, offline`	The root aggregate shows this state if all the disks on the disaster sites were replaced.

```
cluster_B::> storage aggregate show -node local

 Aggregate     Size Available Used% State   #Vols  Nodes      RAID Status
 --------- -------- --------- ----- ------- ------ ---------- ------------
 node_B_1_aggr1
            227.1GB   11.00GB   95% online       1 node_B_1   raid_dp,
                                                              resyncing
 NodeA_1_aggr2
            430.3GB   28.02GB   93% online       2 node_B_1   raid_dp,
                                                              mirror
                                                              degraded

 node_B_1_aggr3
            812.8GB   85.37GB   89% online       5 node_B_1   raid_dp,
                                                              mirrored,
                                                              normal
 3 entries were displayed.

cluster_B::>
```

In the following examples, the three aggregates are each in a different state:

Node	State
node_B_1_aggr1	`resyncing`
node_B_1_aggr2	`mirror degraded`
node_B_1_aggr3	`mirrored, normal`

3. If one or more plexes remain offline, additional steps are required to rebuild the mirror.

   In the preceding table, the mirror for node_B_1_aggr2 must be rebuilt.

a. View details of the aggregate to identify any failed plexes:

storage aggregate show -r -aggregate node_B_1_aggr2

In the following example, plex /node_B_1_aggr2/plex0 is in a failed state:

```
cluster_B::> storage aggregate show -r -aggregate node_B_1_aggr2

 Owner Node: node_B_1
  Aggregate: node_B_1_aggr2 (online, raid_dp, mirror degraded) (block checksums)
   Plex: /node_B_1_aggr2/plex0 (offline, failed, inactive, pool0)
    RAID Group /node_B_1_aggr2/plex0/rg0 (partial)
                                                         Usable Physical
      Position Disk                     Pool Type     RPM     Size     Size Status
      -------- ------------------------ ---- ----- ------ -------- -------- ----------

   Plex: /node_B_1_aggr2/plex1 (online, normal, active, pool1)
    RAID Group /node_B_1_aggr2/plex1/rg0 (normal, block checksums)
                                                         Usable Physical
      Position Disk                     Pool Type     RPM     Size     Size Status
      -------- ------------------------ ---- ----- ------ -------- -------- ----------
      dparity  1.44.8                    1   SAS    15000  265.6GB  273.5GB (normal)
      parity   1.41.11                   1   SAS    15000  265.6GB  273.5GB (normal)
      data     1.42.8                    1   SAS    15000  265.6GB  273.5GB (normal)
      data     1.43.11                   1   SAS    15000  265.6GB  273.5GB (normal)
      data     1.44.9                    1   SAS    15000  265.6GB  273.5GB (normal)
      data     1.43.18                   1   SAS    15000  265.6GB  273.5GB (normal)
 6 entries were displayed.

 cluster_B::>
```

b. Delete the failed plex:

storage aggregate plex delete -aggregate *aggregate-name* -plex *plex*

c. Reestablish the mirror:

storage aggregate mirror -aggregate *aggregate-name*

d. Monitor the resynchronization and mirroring status of the plex until all mirrors are reestablished and all aggregates show `mirrored, normal` status:

storage aggregate show

Reassigning disk ownership for root aggregates to replacement controller modules (MetroCluster FC configurations)

If one or both of the controller modules or NVRAM cards were replaced at the disaster site, the system ID has changed and you must reassign disks belonging to the root aggregates to the replacement controller modules.

About this task

Because the nodes are in switchover mode and healing has been done, only the disks containing the root aggregates of pool1 of the disaster site will be reassigned in this section. They are the only disks still owned by the old system ID at this point.

This section provides examples for two and four-node configurations. For two-node configurations, you can ignore references to the second node at each site. For eight-node configurations, you must account for the additional nodes on the second DR group. The examples make the following assumptions:

- Site A is the disaster site.
- node_A_1 has been replaced.
- node_A_2 has been replaced.
  Present only in four-node MetroCluster configurations.
- Site B is the surviving site.
- node_B_1 is healthy.
- node_B_2 is healthy.
  Present only in four-node MetroCluster configurations.

The old and new system IDs were identified in *Acquiring the new System ID*.

The examples in this procedure use controllers with the following system IDs:

Number of nodes	Node	Original system ID	New system ID
Four	node_A_1	4068741258	1574774970
	node_A_2	4068741260	1574774991
	node_B_1	4068741254	unchanged
	node_B_2	4068741256	unchanged
Two	node_A_1	4068741258	1574774970
	node_B_1	4068741254	unchanged

Steps

1. With the replacement node in Maintenance mode, reassign the root aggregate disks:

   disk reassign -s *old-system-ID* -d *new-system-ID*

   ```
   *> disk reassign -s 4068741258 -d 1574774970
   ```

2. View the disks to confirm the ownership change of the pool1 root aggr disks of the disaster site to the replacement node:

   disk show

   The output might show more or fewer disks, depending on how many disks are in the root aggregate and whether any of these disks failed and were replaced. If the disks were replaced, then Pool0 disks will not appear in the output.

   The pool1 root aggregate disks of the disaster site should now be assigned to the replacement node.

   ```
   *> disk show
   Local System ID: 1574774970

     DISK             OWNER             POOL   SERIAL NUMBER         HOME                DR HOME
   ------------       -------------     -----  -------------         -------------       -------------
   sw_A_1:6.126L19 node_A_1(1574774970) Pool0  serial-number  node_A_1(1574774970)
   sw_A_1:6.126L3  node_A_1(1574774970) Pool0  serial-number  node_A_1(1574774970)
   sw_A_1:6.126L7  node_A_1(1574774970) Pool0  serial-number  node_A_1(1574774970)
   sw_B_1:6.126L8  node_A_1(1574774970) Pool1  serial-number  node_A_1(1574774970)
   sw_B_1:6.126L24 node_A_1(1574774970) Pool1  serial-number  node_A_1(1574774970)
   sw_B_1:6.126L2  node_A_1(1574774970) Pool1  serial-number  node_A_1(1574774970)

   *> aggr status
           Aggr State           Status
    node_A_1_root online         raid_dp, aggr
                                 mirror degraded
                                 64-bit
   *>
   ```

3. View the aggregate status:

   aggr status

   The output might show more or fewer disks, depending on how many disks are in the root aggregate and whether any of these disks failed and were replaced. If disks were replaced, then Pool0 disks will not appear in output.

   ```
   *> aggr status
            Aggr State           Status
     node_A_1_root online         raid_dp, aggr
                                  mirror degraded
                                  64-bit
   *>
   ```

4. Delete the contents of the mailbox disks:

```
mailbox destroy local
```

5. If the aggregate is not online, bring it online:

   ```
   aggr online aggr_name
   ```

6. Halt the node to display the LOADER prompt:

   ```
   halt
   ```

Booting the new controller modules (MetroCluster FC configurations)

After aggregate healing has been completed for both the data and root aggregates, you must boot the node or nodes at the disaster site.

About this task

This task begins with the nodes showing the LOADER prompt.

Steps

1. Display the boot menu:

   ```
   boot_ontap menu
   ```

2. From the boot menu, select option 6, **Update flash from backup config**.
3. Respond y to the following prompt:

   ```
   This will replace all flash-based configuration with the last backup to disks. Are you sure you want to continue?: y
   ```

   The system will boot twice, the second time to load the new configuration.

   Note: If you did not clear the NVRAM contents of a used replacement controller, then you might see a panic with the following message:

   ```
   PANIC: NVRAM contents are invalid...
   ```

   If this occurs, repeat step 2 to boot the system to the ONTAP prompt. You will then need to perform a root recovery. Contact technical support for assistance.

4. Mirror the root aggregate on plex 0:
   a. Assign three pool0 disks to the new controller module.
   b. Mirror the root aggregate pool1 plex:

      ```
      aggr mirror root-aggr-name
      ```

   c. Assign unowned disks to pool0 on the local node
5. Refresh the MetroCluster configuration:
   a. Enter advanced privilege mode:

      ```
      set -privilege advanced
      ```

   b. Refresh the configuration:

      ```
      metrocluster configure -refresh true
      ```

   c. Return to admin privilege mode:

      ```
      set -privilege admin
      ```

6. If you have a four-node configuration, repeat the previous steps on the other node at the disaster site.

After you finish

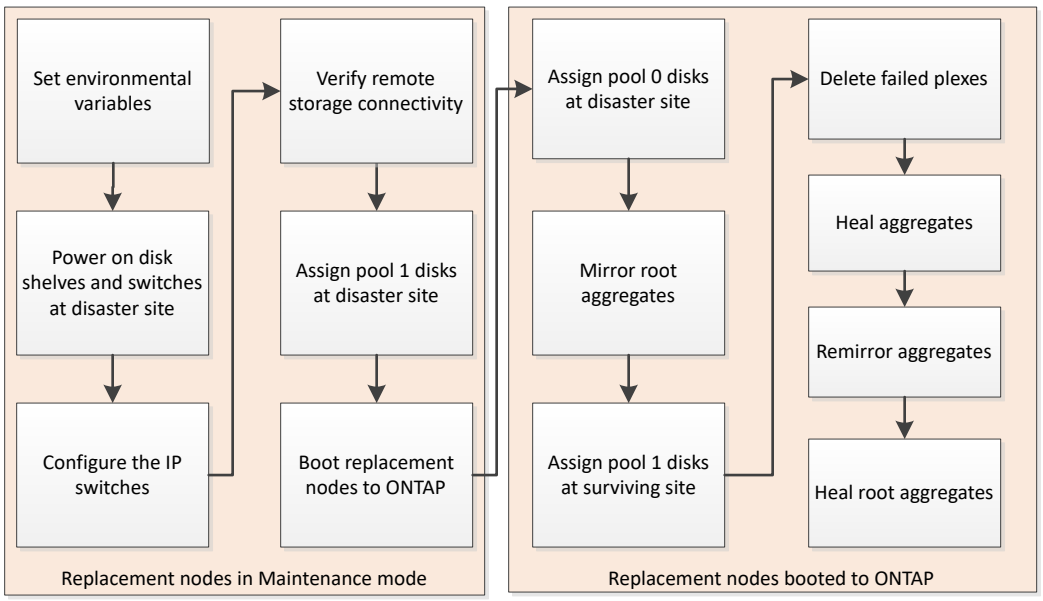
Proceed to verify the licenses on the replaced nodes.

Verifying licenses on the replaced nodes on page 128

Preparing for switchback in a MetroCluster IP configuration

You must perform certain tasks in order to prepare the MetroCluster IP configuration for the switchback operation.

About this task



Steps

1. *Setting required environmental variables in MetroCluster IP configurations* on page 107
2. *Powering on the equipment at the disaster site (MetroCluster IP configurations)* on page 111
3. *Configuring the IP switches (MetroCluster IP configurations)* on page 111
4. *Verify storage connectivity to the remote site (MetroCluster IP configurations)* on page 113
5. *Reassigning disk ownership for pool 1 disks on the disaster site (MetroCluster IP configurations)* on page 113
6. *Booting to ONTAP on replacement controller modules in MetroCluster IP configurations* on page 115
7. *Restoring connectivity from the surviving nodes to the disaster site (MetroCluster IP configurations)* on page 117
8. *Verifying automatic assignment or manually assigning pool 0 drives* on page 118
9. *Assigning pool 1 drives on the surviving site (MetroCluster IP configurations)* on page 119
10. *Deleting failed plexes owned by the surviving site (MetroCluster IP configurations)* on page 119
11. *Performing aggregate healing and restoring mirrors (MetroCluster IP configurations)* on page 122

Setting required environmental variables in MetroCluster IP configurations

In MetroCluster IP configurations, you must retrieve the IP address of the MetroCluster interfaces on the Ethernet ports, and then use them to configure the interfaces on the replacement controller modules.

About this task

This task is required only in MetroCluster IP configurations.

Commands in this task are performed from the cluster prompt of the surviving site and from the LOADER prompt of the nodes at the disaster site.

The nodes in these examples have the following IP addresses for their MetroCluster IP connections:

Note: These examples are for an AFF A700 or FAS9000 system. The interfaces vary by platform model.

Node	Port	IP address
node_A_1	e5a	172.17.26.10
	e5b	172.17.27.10
node_A_2	e5a	172.17.26.11
	e5b	172.17.27.11
node_B_1	e5a	172.17.26.13
	e5b	172.17.27.13
node_B_2	e5a	172.17.26.12
	e5b	172.17.27.12

The following table summarizes the relationships between the nodes and each node's MetroCluster IP addresses.

Node	HA partner	DR partner	DR auxiliary partner
node_A_1 • e5a: 172.17.26.10 • e5b: 172.17.27.10	node_A_2 • e5a: 172.17.26.11 • e5b: 172.17.27.11	node_B_1 • e5a: 172.17.26.13 • e5b: 172.17.27.13	node_B_2 • e5a: 172.17.26.12 • e5b: 172.17.27.12
node_A_2 • e5a: 172.17.26.11 • e5b: 172.17.27.11	node_A_1 • e5a: 172.17.26.10 • e5b: 172.17.27.10	node_B_2 • e5a: 172.17.26.12 • e5b: 172.17.27.12	node_B_1 • e5a: 172.17.26.13 • e5b: 172.17.27.13
node_B_1 • e5a: 172.17.26.13 • e5b: 172.17.27.13	node_B_2 • e5a: 172.17.26.12 • e5b: 172.17.27.12	node_A_1 • e5a: 172.17.26.10 • e5b: 172.17.27.10	node_A_2 • e5a: 172.17.26.11 • e5b: 172.17.27.11
node_B_2 • e5a: 172.17.26.12 • e5b: 172.17.27.12	node_B_1 • e5a: 172.17.26.13 • e5b: 172.17.27.13	node_A_2 • e5a: 172.17.26.11 • e5b: 172.17.27.11	node_A_1 • e5a: 172.17.26.10 • e5b: 172.17.27.10

The following table lists the platform models that use VLAN IDs on the MetroCluster IP interfaces. These models may require additional steps if you are not using the default VLAN IDs.

Platform models that use VLAN IDs with the MetroCluster IP interfaces	
• AFF A220 • AFF A250 • AFF A400 • AFF A800	• FAS500f • FAS2750 • FAS8300 • FAS8700

Steps

1. From the surviving site, gather the IP addresses of the MetroCluster interfaces on the disaster site:

   `metrocluster configuration-settings connection show`

   The required addresses are the DR Partner addresses shown in the `Destination Network Address` column.

   The following output shows the IP addresses for a configuration with AFF A700 and FAS9000 systems with the MetroCluster IP interfaces on ports e5a and e5b. The interfaces vary depending on platform type.

```
cluster_B::*> metrocluster configuration-settings connection show
DR                    Source          Destination
DR                    Source          Destination
Group Cluster Node    Network Address Network Address Partner Type Config State
----- ------- ------- --------------- --------------- ------------ ------------
1     cluster_B
              node_B_1
                 Home Port: e5a
                      172.17.26.13    172.17.26.12    HA Partner   completed
                 Home Port: e5a
                      172.17.26.13    172.17.26.10    DR Partner   completed
                 Home Port: e5a
                      172.17.26.13    172.17.26.11    DR Auxiliary completed
                 Home Port: e5b
                      172.17.27.13    172.17.27.12    HA Partner   completed
                 Home Port: e5b
                      172.17.27.13    172.17.27.10    DR Partner   completed
                 Home Port: e5b
                      172.17.27.13    172.17.27.11    DR Auxiliary completed
              node_B_2
                 Home Port: e5a
                      172.17.26.12    172.17.26.13    HA Partner   completed
                 Home Port: e5a
                      172.17.26.12    172.17.26.11    DR Partner   completed
                 Home Port: e5a
                      172.17.26.12    172.17.26.10    DR Auxiliary completed
                 Home Port: e5b
                      172.17.27.12    172.17.27.13    HA Partner   completed
                 Home Port: e5b
                      172.17.27.12    172.17.27.11    DR Partner   completed
                 Home Port: e5b
                      172.17.27.12    172.17.27.10    DR Auxiliary completed
12 entries were displayed.
```

2. If the systems use VLAN IDs with the MetroCluster IP interfaces (see the list above), and if you are not using the default VLAN IDs, determine the VLAN IDs from the surviving site:

   `metrocluster configuration-settings interface show`

   The VLAN IDs are included in the Network Address column of the output.

   In this example the interfaces are e0a with the VLAN ID 120 and e0b with the VLAN ID 130:

```
Cluster-A::*> metrocluster configuration-settings interface show
DR                                                                        Config
Group Cluster Node     Network Address Netmask         Gateway         State
----- ------- ------- --------------- --------------- --------------- ---------
1
      cluster_A
              node_A_1
                  Home Port: e0a-120
                          172.17.26.10  255.255.255.0  -            completed
                  Home Port: e0b-130
                          172.17.27.10  255.255.255.0  -            completed
```

3. If the disaster site nodes use VLAN IDs (see the list above), at the LOADER prompt for each of the disaster site nodes, set the following bootargs:

   setenv bootarg.mcc.port_a_ip_config *local-IP-address/local-IP-mask*,0,*HA-partner-IP-address*,*DR-partner-IP-address*,*DR-aux-partnerIP-address*,*vlan-id*

   setenv bootarg.mcc.port_b_ip_config *local-IP-address/local-IP-mask*,0,*HA-partner-IP-address*,*DR-partner-IP-address*,*DR-aux-partnerIP-address*,*vlan-id*

   Note: If the interfaces are using the default VLANs, the *vlan-id* is not necessary.

   The following commands set the values for node_A_1 using VLAN 120 for the first network and VLAN 130 for the second network:

   ```
   setenv bootarg.mcc.port_a_ip_config
   172.17.26.10/23,0,172.17.26.11,172.17.26.13,172.17.26.12,120
   setenv bootarg.mcc.port_b_ip_config
   172.17.27.10/23,0,172.17.27.11,172.17.27.13,172.17.27.12,130
   ```

   The following commands set the values for node_A_2 using VLAN 120 for the first network and VLAN 130 for the second network:

   ```
   setenv bootarg.mcc.port_a_ip_config
   172.17.26.11/23,0,172.17.26.10,172.17.26.12,172.17.26.13,120
   setenv bootarg.mcc.port_b_ip_config
   172.17.27.11/23,0,172.17.27.10,172.17.27.12,172.17.27.13,130
   ```

   The following example shows the commands for node_A_1 when the default VLAN is used:

   ```
   setenv bootarg.mcc.port_a_ip_config
   172.17.26.10/23,0,172.17.26.11,172.17.26.13,172.17.26.12
   setenv bootarg.mcc.port_b_ip_config
   172.17.27.10/23,0,172.17.27.11,172.17.27.13,172.17.27.12
   ```

   The following example shows the commands for node_A_2 when the default VLAN is used:

   ```
   setenv bootarg.mcc.port_a_ip_config
   172.17.26.11/23,0,172.17.26.10,172.17.26.12,172.17.26.13
   setenv bootarg.mcc.port_b_ip_config
   172.17.27.11/23,0,172.17.27.10,172.17.27.12,172.17.27.13
   ```

4. If the disaster site nodes are not systems listed in the previous step, at the LOADER prompt for each of the disaster nodes, set the following bootargs with local_IP/mask,gateway:

   setenv bootarg.mcc.port_a_ip_config *local-IP-address/local-IP-mask*,0,*HA-partner-IP-address*,*DR-partner-IP-address*,*DR-aux-partnerIP-address*

   setenv bootarg.mcc.port_b_ip_config *local-IP-address/local-IP-mask*,0,*HA-partner-IP-address*,*DR-partner-IP-address*,*DR-aux-partnerIP-address*

   The following commands set the values for node_A_1:

   ```
   setenv bootarg.mcc.port_a_ip_config 172.17.26.10/23,0,172.17.26.11,172.17.26.13,172.17.26.12
   setenv bootarg.mcc.port_b_ip_config 172.17.27.10/23,0,172.17.27.11,172.17.27.13,172.17.27.12
   ```

   The following commands set the values for node_A_2:

   ```
   setenv bootarg.mcc.port_a_ip_config 172.17.26.11/23,0,172.17.26.10,172.17.26.12,172.17.26.13
   setenv bootarg.mcc.port_b_ip_config 172.17.27.11/23,0,172.17.27.10,172.17.27.12,172.17.27.13
   ```

5. From the surviving site, gather the UUIDs for the disaster site:

   metrocluster node show -fields node-cluster-uuid, node-uuid

   ```
   cluster_B::> metrocluster node show -fields node-cluster-uuid, node-uuid
     (metrocluster node show)
   dr-group-id cluster     node     node-uuid                            node-cluster-uuid
   ----------- ----------- -------- ------------------------------------ ------------------------------
   1           cluster_A   node_A_1 f03cb63c-9a7e-11e7-b68b-00a098908039 ee7db9d5-9a82-11e7-b68b-00a098
                                                                          908039
   1           cluster_A   node_A_2 aa9a7a7a-9a81-11e7-a4e9-00a098908c35 ee7db9d5-9a82-11e7-b68b-00a098
                                                                          908039
   1           cluster_B   node_B_1 f37b240b-9ac1-11e7-9b42-00a098c9e55d 07958819-9ac6-11e7-9b42-00a098
   ```

```
                                                                    c9e55d
1           cluster_B   node_B_2 bf8e3f8f-9ac4-11e7-bd4e-00a098ca379f 07958819-9ac6-11e7-9b42-00a098
                                                                    c9e55d
4 entries were displayed.
cluster_A::*>
```

Node	UUID
cluster_B	07958819-9ac6-11e7-9b42-00a098c9e55d
node_B_1	f37b240b-9ac1-11e7-9b42-00a098c9e55d
node_B_2	bf8e3f8f-9ac4-11e7-bd4e-00a098ca379f
cluster_A	ee7db9d5-9a82-11e7-b68b-00a098908039
node_A_1	f03cb63c-9a7e-11e7-b68b-00a098908039
node_A_2	aa9a7a7a-9a81-11e7-a4e9-00a098908c35

6. At the replacement nodes' LOADER prompt, set the UUIDs:

   setenv bootarg.mgwd.partner_cluster_uuid *partner-cluster-UUID*

   setenv bootarg.mgwd.cluster_uuid *local-cluster-UUID*

   setenv bootarg.mcc.pri_partner_uuid *DR-partner-node-UUID*

   setenv bootarg.mcc.aux_partner_uuid *DR-aux-partner-node-UUID*

   setenv bootarg.mcc_iscsi.node_uuid *local-node-UUID*

   a. Set the UUIDs on node_A_1.
      The following example shows the commands for setting the UUIDs on node_A_1:

      ```
      setenv bootarg.mgwd.cluster_uuid ee7db9d5-9a82-11e7-b68b-00a098908039
      setenv bootarg.mgwd.partner_cluster_uuid 07958819-9ac6-11e7-9b42-00a098c9e55d
      setenv bootarg.mcc.pri_partner_uuid f37b240b-9ac1-11e7-9b42-00a098c9e55d
      setenv bootarg.mcc.aux_partner_uuid bf8e3f8f-9ac4-11e7-bd4e-00a098ca379f
      setenv bootarg.mcc_iscsi.node_uuid f03cb63c-9a7e-11e7-b68b-00a098908039
      ```

   b. Set the UUIDs on node_A_2:
      The following example shows the commands for setting the UUIDs on node_A_2:

      ```
      setenv bootarg.mgwd.cluster_uuid ee7db9d5-9a82-11e7-b68b-00a098908039
      setenv bootarg.mgwd.partner_cluster_uuid 07958819-9ac6-11e7-9b42-00a098c9e55d
      setenv bootarg.mcc.pri_partner_uuid bf8e3f8f-9ac4-11e7-bd4e-00a098ca379f
      setenv bootarg.mcc.aux_partner_uuid f37b240b-9ac1-11e7-9b42-00a098c9e55d
      setenv bootarg.mcc_iscsi.node_uuid aa9a7a7a-9a81-11e7-a4e9-00a098908c35
      ```

7. If the original systems were configured for ADP, at each of the replacement nodes' LOADER prompt, enable ADP:

   setenv bootarg.mcc.adp_enabled true

8. If running ONTAP 9.5, 9.6 or 9.7, at each of the replacement nodes' LOADER prompt, enable the following variable:

   setenv bootarg.mcc.lun_part true

   a. Set the variables on node_A_1.
      The following example shows the commands for setting the values on node_A_1 when running ONTAP 9.6:

      ```
      setenv bootarg.mcc.lun_part true
      ```

   b. Set the variables on node_A_2.
      The following example shows the commands for setting the values on node_A_2 when running ONTAP 9.6:

      ```
      setenv bootarg.mcc.lun_part true
      ```

9. If the original systems were configured for ADP, at each of the replacement nodes' LOADER prompt, set the original system ID (**not** the system ID of the replacement controller module) and the system ID of the DR partner of the node:

setenv bootarg.mcc.local_config_id *original-sysID*

setenv bootarg.mcc.dr_partner *dr_partner-sysID*

Determining the system IDs and VLAN IDs of the old controller modules on page 50

a. Set the variables on node_A_1.
The following example shows the commands for setting the system IDs on node_A_1:

- The old system ID of node_A_1 is 4068741258.
- The system ID of node_B_1 is 4068741254.

```
setenv bootarg.mcc.local_config_id 4068741258
setenv bootarg.mcc.dr_partner 4068741254
```

b. Set the variables on node_A_2.
The following example shows the commands for setting the system IDs on node_A_2:

- The old system ID of node_A_1 is 4068741260.
- The system ID of node_B_1 is 4068741256.

```
setenv bootarg.mcc.local_config_id 4068741260
setenv bootarg.mcc.dr_partner 4068741256
```

Powering on the equipment at the disaster site (MetroCluster IP configurations)

You must power on the disk shelves and MetroCluster IP switches components at the disaster site. The controller modules at the disaster site remain at the LOADER prompt.

About this task

The examples in this procedure assume the following:

- Site A is the disaster site.
- Site B is the surviving site.

Steps

1. Turn on the disk shelves at the disaster site and make sure that all disks are running.
2. Turn on the MetroCluster IP switches if they are not already on.

Configuring the IP switches (MetroCluster IP configurations)

You must configure any IP switches that were replaced.

About this task

This task applies to MetroCluster IP configurations only.

This must be done on both switches. Verify after configuring the first switch that storage access on the surviving site is not impacted.

Note: You must not proceed with the second switch if storage access on the surviving site is impacted.

Steps

1. Refer to the *MetroCluster IP Installation and Configuration Guide* for procedures for cabling and configuring a replacement switch.

   MetroCluster IP installation and configuration

   You can use the procedures in the following sections:

   - Cabling the IP switches
   - Configuring the IP switches
2. If the ISLs were disabled at the surviving site, enable the ISLs and verify that the ISLs are online.

a. Enable the of the ISL interfaces on the first switch:

`no shutdown`

The following examples show the commands for a Broadcom IP switch or a Cisco IP switch.

Switch vendor	Commands

Broadcom	`(IP_Switch_A_1)> enable`
	`(IP_switch_A_1)# configure`
	`(IP_switch_A_1)(Config)# interface 0/13-0/16`
	`(IP_switch_A_1)(Interface 0/13-0/16 )# no shutdown`
	`(IP_switch_A_1)(Interface 0/13-0/16 )# exit`
	`(IP_switch_A_1)(Config)# exit`

Cisco	`IP_switch_A_1# conf t`
	`IP_switch_A_1(config)# int eth1/15-eth1/20`
	`IP_switch_A_1(config)# no shutdown`
	`IP_switch_A_1(config)# copy running startup`
	`IP_switch_A_1(config)# show interface brief`

b. Enable the of the ISL interfaces on the partner switch:

`no shutdown`

The following examples show the commands for a Broadcom IP switch or a Cisco IP switch.

Switch vendor	Commands

Broadcom	`(IP_Switch_A_2)> enable`
	`(IP_switch_A_2)# configure`
	`(IP_switch_A_2)(Config)# interface 0/13-0/16`
	`(IP_switch_A_2)(Interface 0/13-0/16 )# no shutdown`
	`(IP_switch_A_2)(Interface 0/13-0/16 )# exit`
	`(IP_switch_A_2)(Config)# exit`

Cisco	`IP_switch_A_2# conf t`
	`IP_switch_A_2(config)# int eth1/15-eth1/20`
	`IP_switch_A_2(config)# no shutdown`
	`IP_switch_A_2(config)# copy running startup`
	`IP_switch_A_2(config)# show interface brief`

c. Verify that the interfaces are enabled:

`show interface brief`

The following example shows the output for a Cisco switch.

```
IP_switch_A_2(config)# show interface brief

--------------------------------------------------------
Port VRF Status IP Address Speed MTU
--------------------------------------------------------
mt0 -- up 10.10.99.10 100 1500
--------------------------------------------------------
Ethernet    VLAN Type Mode    Status Reason Speed   Port
Interface                                         Ch
#
--------------------------------------------------------
.
.
.
Eth1/15    10   eth   access  up     none   40G(D)  --
Eth1/16    10   eth   access  up     none   40G(D)  --
Eth1/17    10   eth   access  down   none   auto(D) --
```

```
Eth1/18    10   eth   access  down   none   auto(D) --
Eth1/19    10   eth   access  down   none   auto(D) --
Eth1/20    10   eth   access  down   none   auto(D) --
.
.
.
IP_switch_A_2#
```

After you finish

Proceed to *Preparing for switchback in a MetroCluster IP configuration* on page 106.

Verify storage connectivity to the remote site (MetroCluster IP configurations)

You must confirm that the replaced nodes have connectivity to the disk shelves at the surviving site.

About this task

This task is performed on the replacement nodes at the disaster site.

This task is performed in Maintenance mode.

Steps

1. Display the disks that are owned by the original system ID.

   disk show -s *old-system-ID*

   The remote disks can be recognized by the 0m device. 0m indicates that the disk is connected via the MetroCluster iSCSI connection. These disks must be reassigned later in the recovery procedure.

```
*> disk show -s 4068741256
Local System ID: 1574774970

  DISK     OWNER                 POOL  SERIAL NUMBER   HOME                  DR HOME
---------- --------------------- ----- -------------   --------------------- ----------------------
0m.i0.0L11 node_A_2 (4068741256) Pool1 S396NA0HA02128  node_A_2 (4068741256) node_A_2  (4068741256)
0m.i0.1L38 node_A_2 (4068741256) Pool1 S396NA0J148778  node_A_2 (4068741256) node_A_2  (4068741256)
0m.i0.0L52 node_A_2 (4068741256) Pool1 S396NA0J148777  node_A_2 (4068741256) node_A_2  (4068741256)
...
...
NOTE: Currently 49 disks are unowned. Use 'disk show -n' for additional information.
*>
```

2. Repeat this step on the other replacement nodes

Reassigning disk ownership for pool 1 disks on the disaster site (MetroCluster IP configurations)

If one or both of the controller modules or NVRAM cards were replaced at the disaster site, the system ID has changed and you must reassign disks belonging to the root aggregates to the replacement controller modules.

About this task

Because the nodes are in switchover mode, only the disks containing the root aggregates of pool1 of the disaster site will be reassigned in this task. They are the only disks still owned by the old system ID at this point.

This task is performed on the replacement nodes at the disaster site.

This task is performed in Maintenance mode.

The examples make the following assumptions:

- Site A is the disaster site.
- node_A_1 has been replaced.
- node_A_2 has been replaced.
- Site B is the surviving site.

- node_B_1 is healthy.
- node_B_2 is healthy.

The old and new system IDs were identified in *Acquiring the new System ID*.

The examples in this procedure use controllers with the following system IDs:

Node	Original system ID	New system ID
node_A_1	4068741258	1574774970
node_A_2	4068741260	1574774991
node_B_1	4068741254	unchanged
node_B_2	4068741256	unchanged

Steps

1. With the replacement node in Maintenance mode, reassign the root aggregate disks, using the correct command, depending on whether your system is configured with ADP and your ONTAP version.

   You can proceed with the reassignment when prompted.

System is using ADP	Use this command for disk reassignment:
Yes (ONTAP 9.8)	`disk reassign -s old-system-ID -d new-system-ID -r dr-partner-system-ID`
Yes (ONTAP 9.7.x and earlier)	`disk reassign -s old-system-ID -d new-system-ID -p old-partner-system-ID`
No	`disk reassign -s old-system-ID -d new-system-ID`

   The following example shows reassignment of drives on a non-ADP system:

```
*> disk reassign -s 4068741256 -d 1574774970
Partner node must not be in Takeover mode during disk reassignment from maintenance mode.
Serious problems could result!!
Do not proceed with reassignment if the partner is in takeover mode. Abort reassignment (y/n)? n

After the node becomes operational, you must perform a takeover and giveback of the HA partner node to ensure
disk reassignment is successful.
Do you want to continue (y/n)? y
Disk ownership will be updated on all disks previously belonging to Filer with sysid 537037643.
Do you want to continue (y/n)? y
disk reassign parameters: new_home_owner_id 537070473 , new_home_owner_name
Disk 0m.i0.3L14 will be reassigned.
Disk 0m.i0.1L6 will be reassigned.
Disk 0m.i0.1L8 will be reassigned.
Number of disks to be reassigned: 3
```

2. Destroy the contents of the mailbox disks:

   `mailbox destroy local`

   You can proceed with the destroy operation when prompted.

   The following example shows the output for the `mailbox destroy local` command:

```
*> mailbox destroy local
Destroying mailboxes forces a node to create new empty mailboxes,
which clears any takeover state, removes all knowledge
of out-of-date plexes of mirrored volumes, and will prevent
management services from going online in 2-node cluster
HA configurations.
Are you sure you want to destroy the local mailboxes? y
..............Mailboxes destroyed.
*>
```

3. If disks have been replaced, there will be failed local plexes that must be deleted.
   a. Display the aggregate status:

      aggr status
      In the following example, plex node_A_1_aggr0/plex0 has failed.

      ```
      *> aggr status
      Aug 18 15:00:07 [node_B_1:raid.vol.mirror.degraded:ALERT]: Aggregate node_A_1_aggr0 is
         mirrored and one plex has failed. It is no longer protected by mirroring.
      Aug 18 15:00:07 [node_B_1:raid.debug:info]: Mirrored aggregate node_A_1_aggr0 has plex0
         clean(-1), online(0)
      Aug 18 15:00:07 [node_B_1:raid.debug:info]: Mirrored aggregate node_A_1_aggr0 has plex2
         clean(0), online(1)
      Aug 18 15:00:07 [node_B_1:raid.mirror.vote.noRecord1Plex:error]: WARNING: Only one plex
         in aggregate node_A_1_aggr0 is available. Aggregate might contain stale data.
      Aug 18 15:00:07 [node_B_1:raid.debug:info]: volobj_mark_sb_recovery_aggrs: tree:
         node_A_1_aggr0 vol_state:1 mcc_dr_opstate: unknown
      Aug 18 15:00:07 [node_B_1:raid.fsm.commitStateTransit:debug]: /node_A_1_aggr0 (VOL):
         raid state change UNINITD -> NORMAL
      Aug 18 15:00:07 [node_B_1:raid.fsm.commitStateTransit:debug]: /node_A_1_aggr0 (MIRROR):
         raid state change UNINITD -> DEGRADED
      Aug 18 15:00:07 [node_B_1:raid.fsm.commitStateTransit:debug]: /node_A_1_aggr0/plex0
         (PLEX): raid state change UNINITD -> FAILED
      Aug 18 15:00:07 [node_B_1:raid.fsm.commitStateTransit:debug]: /node_A_1_aggr0/plex2
         (PLEX): raid state change UNINITD -> NORMAL
      Aug 18 15:00:07 [node_B_1:raid.fsm.commitStateTransit:debug]: /node_A_1_aggr0/plex2/rg0
         (GROUP): raid state change UNINITD -> NORMAL
      Aug 18 15:00:07 [node_B_1:raid.debug:info]: Topology updated for aggregate node_A_1_aggr0
         to plex plex2
      *>
      ```

   b. Delete the failed plex:

      aggr destroy *plex-id*

      ```
      *> aggr destroy node_A_1_aggr0/plex0
      ```

4. Halt the node to display the LOADER prompt:

   halt

5. Repeat these steps on the other node at the disaster site.

Booting to ONTAP on replacement controller modules in MetroCluster IP configurations

You must boot the replacement nodes at the disaster site to the ONTAP operating system.

About this task

This task begins with the nodes at the disaster site in Maintenance mode.

Steps

1. On one of the replacement nodes, exit to the LOADER prompt:

   halt

2. Display the boot menu:

   boot_ontap menu

3. From the boot menu, select option 6, **Update flash from backup config**.

   The system boots twice. You should respond **yes** when prompted to continue. After the second boot, you should respond **y** when prompted about the system ID mismatch.

   Note: If you did not clear the NVRAM contents of a used replacement controller module, then you might see the following panic message: `PANIC: NVRAM contents are invalid....`

   If this occurs, boot the system to the ONTAP prompt again (**boot_ontap menu**). You then need to perform a root recovery. Contact technical support for assistance.

   Confirmation to continue prompt:

```
Selection (1-9)? 6

This will replace all flash-based configuration with the last backup to
disks. Are you sure you want to continue?: yes
```

System ID mismatch prompt:

```
WARNING: System ID mismatch. This usually occurs when replacing a boot device or NVRAM
cards!
Override system ID? {y|n} y
```

4. From the surviving site, verify that the correct partner system IDs have been applied to the nodes:

   `metrocluster node show -fields node-systemid,ha-partner-systemid,dr-partner-systemid,dr-auxiliary-systemid`

   In this example, the following new system IDs should appear in the output:

   - Node_A_1: 1574774970
   - Node_A_2: 1574774991

   The `ha-partner-systemid` column should show the new system IDs.

```
metrocluster node show -fields node-systemid,ha-partner-systemid,dr-partner-systemid,dr-auxiliary-systemid

dr-group-id cluster    node      node-systemid ha-partner-systemid dr-partner-systemid dr-auxiliary-systemid
----------- ---------- --------  ------------- ------------------- ------------------- ---------------------
1           Cluster_A  Node_A_1  1574774970    1574774991          4068741254          4068741256
1           Cluster_A  Node_A_2  1574774991    1574774970          4068741256          4068741254
1           Cluster_B  Node_B_1  -             -                   -                   -
1           Cluster_B  Node_B_2  -             -                   -                   -
4 entries were displayed.
```

5. If the partner system IDs were not correctly set, you must manually set the correct value:
   a. Halt and display the LOADER prompt on the node.
   b. Verify the *`partner-sysID`* bootarg's current value:

      `printenv`

   c. Set the value to the correct partner system ID:

      `setenv partner-sysid partner-sysID`

   d. Boot the node:

      `boot_ontap`

   e. Repeat these substeps on the other node, if necessary.

6. Confirm that the replacement nodes at the disaster site are ready for switchback:

   `metrocluster node show`

   The replacement nodes should be in `waiting for switchback recovery` mode. If they are in `normal` mode instead, you can reboot the replacement nodes. After that boot, the nodes should be in `waiting for switchback recovery` mode.

   The following example shows that the replacement nodes are ready for switchback:

```
cluster_B::> metrocluster node show
DR                           Configuration  DR
Group Cluster Node           State          Mirroring Mode
----- ------- ------------------ -------------- --------- --------------------
1     cluster_B
              node_B_1           configured     enabled   switchover completed
              node_B_2           configured     enabled   switchover completed
      cluster_A
              node_A_1           configured     enabled   waiting for switchback recovery
              node_A_2           configured     enabled   waiting for switchback recovery
4 entries were displayed.

cluster_B::>
```

7. Verify the MetroCluster connection configuration settings:

   `metrocluster configuration-settings connection show`

   The configuration state should indicate `completed`.

```
cluster_B::*> metrocluster configuration-settings connection show
DR                    Source          Destination
Group Cluster Node    Network Address Network Address Partner Type Config State
----- ------- ------- --------------- --------------- ------------ ------------
1     cluster_B
              node_B_2
                 Home Port: e5a
                      172.17.26.13    172.17.26.12    HA Partner   completed
                 Home Port: e5a
                      172.17.26.13    172.17.26.10    DR Partner   completed
                 Home Port: e5a
                      172.17.26.13    172.17.26.11    DR Auxiliary completed
                 Home Port: e5b
                      172.17.27.13    172.17.27.12    HA Partner   completed
                 Home Port: e5b
                      172.17.27.13    172.17.27.10    DR Partner   completed
                 Home Port: e5b
                      172.17.27.13    172.17.27.11    DR Auxiliary completed
              node_B_1
                 Home Port: e5a
                      172.17.26.12    172.17.26.13    HA Partner   completed
                 Home Port: e5a
                      172.17.26.12    172.17.26.11    DR Partner   completed
                 Home Port: e5a
                      172.17.26.12    172.17.26.10    DR Auxiliary completed
                 Home Port: e5b
                      172.17.27.12    172.17.27.13    HA Partner   completed
                 Home Port: e5b
                      172.17.27.12    172.17.27.11    DR Partner   completed
                 Home Port: e5b
                      172.17.27.12    172.17.27.10    DR Auxiliary completed
      cluster_A
              node_A_2
                 Home Port: e5a
                      172.17.26.11    172.17.26.10    HA Partner   completed
                 Home Port: e5a
                      172.17.26.11    172.17.26.12    DR Partner   completed
                 Home Port: e5a
                      172.17.26.11    172.17.26.13    DR Auxiliary completed
                 Home Port: e5b
                      172.17.27.11    172.17.27.10    HA Partner   completed
                 Home Port: e5b
                      172.17.27.11    172.17.27.12    DR Partner   completed
                 Home Port: e5b
                      172.17.27.11    172.17.27.13    DR Auxiliary completed
              node_A_1
                 Home Port: e5a
                      172.17.26.10    172.17.26.11    HA Partner   completed
                 Home Port: e5a
                      172.17.26.10    172.17.26.13    DR Partner   completed
                 Home Port: e5a
                      172.17.26.10    172.17.26.12    DR Auxiliary completed
                 Home Port: e5b
                      172.17.27.10    172.17.27.11    HA Partner   completed
                 Home Port: e5b
                      172.17.27.10    172.17.27.13    DR Partner   completed
                 Home Port: e5b
                      172.17.27.10    172.17.27.12    DR Auxiliary completed
24 entries were displayed.

cluster_B::*>
```

8. Repeat the previous steps on the other node at the disaster site.

Restoring connectivity from the surviving nodes to the disaster site (MetroCluster IP configurations)

You must restore the MetroCluster iSCSI initiator connections from the surviving nodes.

About this task

This procedure is only required on MetroCluster IP configurations.

Steps

1. From either surviving node's prompt, change to the advanced privilege level:

   `set -privilege advanced`

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

2. Connect the iSCSI initiators on both surviving nodes in the DR group:

   `storage iscsi-initiator connect -node *surviving-node* -label *`

The following example shows the commands for connecting the initiators on site B:

```
site_B::*> storage iscsi-initiator connect -node node_B_1 -label *
site_B::*> storage iscsi-initiator connect -node node_B_2 -label *
```

3. Return to the admin privilege level:

   `set -privilege admin`

Verifying automatic assignment or manually assigning pool 0 drives

On systems configured for ADP, you must verify that pool 0 drives have been automatically assigned. On systems configured that are not configured for ADP, you must manually assign the pool 0 drives.

Choices

- *Verifying drive assignment of pool 0 drives on ADP systems at the disaster site (MetroCluster IP systems)* on page 118
- *Assigning pool 0 drives on non-ADP systems at the disaster site (MetroCluster IP configurations)* on page 119

Verifying drive assignment of pool 0 drives on ADP systems at the disaster site (MetroCluster IP systems)

If drives have been replaced at the disaster site and the system is configured for ADP, you must verify that the remote drives are visible to the nodes and have been assigned correctly.

Step

Verify that pool 0 drives are assigned automatically:

`disk show`

In the following example 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 (pool1) drives for node_B_1 and node_B_2.

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

Assigning pool 0 drives on non-ADP systems at the disaster site (MetroCluster IP configurations)

If drives have been replaced at the disaster site and the system is not configured for ADP, you need to manually assign new drives to pool 0.

About this task

For ADP systems, the drives are assigned automatically.

Steps

1. On one of the replacement nodes at the disaster site, reassign the node's pool 0 drives:

   storage disk assign -n *number-of-replacement disks* -p 0

   This command assigns the newly added (and unowned) drives on the disaster site. You should assign the same number and size (or larger) of drives that the node had prior to the disaster. The `storage disk assign` man page contains about performing more granular drive assignment.
2. Repeat the step on the other replacement node at the disaster site.

Assigning pool 1 drives on the surviving site (MetroCluster IP configurations)

If drives have been replaced at the disaster site and the system is not configured for ADP, at the surviving site you need to manually assign remote drives located at the disaster site to the surviving nodes' pool 1. You must identify the number of drives to assign.

About this task

For ADP systems, the drives are assigned automatically.

Step

On the surviving site, assign the first node's pool 1 (remote) drives:

storage disk assign -n *number-of-replacement disks* -p 1 0m*

This command assigns the newly added and unowned drives on the disaster site.

The following command assigns 22 drives:

```
cluster_B::> storage disk assign -n 22 -p 1 0m*
```

Deleting failed plexes owned by the surviving site (MetroCluster IP configurations)

After replacing hardware and assigning disks, you must delete failed remote plexes that are owned by the surviving site nodes but located at the disaster site.

About this task

These steps are performed on the surviving cluster.

Steps

1. Identify the local aggregates:

   storage aggregate show -is-home true

```
cluster_B::> storage aggregate show -is-home true

cluster_B Aggregates:
Aggregate     Size Available Used% State   #Vols  Nodes            RAID Status
--------- -------- --------- ----- ------- ------ ---------------- ------------
node_B_1_aggr0 1.49TB  74.12GB 95% online       1 node_B_1         raid4,
                                                                   mirror
                                                                   degraded
node_B_2_aggr0 1.49TB  74.12GB 95% online       1 node_B_2         raid4,
                                                                   mirror
                                                                   degraded
node_B_1_aggr1 2.99TB  2.88TB   3% online      15 node_B_1         raid_dp,
                                                                   mirror
                                                                   degraded
node_B_1_aggr2 2.99TB  2.91TB   3% online      14 node_B_1         raid_tec,
                                                                   mirror
```

```
                                                                    degraded
node_B_2_aggr1 2.95TB  2.80TB   5% online      37 node_B_2         raid_dp,
                                                                    mirror
                                                                    degraded
node_B_2_aggr2 2.99TB  2.87TB   4% online      35 node_B_2         raid_tec,
                                                                    mirror
                                                                    degraded
6 entries were displayed.

cluster_B::>
```

2. Identify the failed remote plexes:

   `storage aggregate plex show`

   The following example calls out the plexes that are remote (not plex0) and have a status of `failed`:

```
cluster_B::> storage aggregate plex show -fields aggregate,status,is-online,Plex,pool
aggregate    plex  status        is-online pool
------------ ----- ------------- --------- ----
node_B_1_aggr0 plex0 normal,active true     0
node_B_1_aggr0 plex4 failed,inactive false  - <<<<---Plex at remote site
node_B_2_aggr0 plex0 normal,active true     0
node_B_2_aggr0 plex4 failed,inactive false  - <<<<---Plex at remote site
node_B_1_aggr1 plex0 normal,active true     0
node_B_1_aggr1 plex4 failed,inactive false  - <<<<---Plex at remote site
node_B_1_aggr2 plex0 normal,active true     0
node_B_1_aggr2 plex1 failed,inactive false  - <<<<---Plex at remote site
node_B_2_aggr1 plex0 normal,active true     0
node_B_2_aggr1 plex4 failed,inactive false  - <<<<---Plex at remote site
node_B_2_aggr2 plex0 normal,active true     0
node_B_2_aggr2 plex1 failed,inactive false  - <<<<---Plex at remote site
node_A_1_aggr1 plex0 failed,inactive false  -
node_A_1_aggr1 plex4 normal,active true     1
node_A_1_aggr2 plex0 failed,inactive false  -
node_A_1_aggr2 plex1 normal,active true     1
node_A_2_aggr1 plex0 failed,inactive false  -
node_A_2_aggr1 plex4 normal,active true     1
node_A_2_aggr2 plex0 failed,inactive false  -
node_A_2_aggr2 plex1 normal,active true     1
20 entries were displayed.

cluster_B::>
```

3. Take offline each of the failed plexes, and then delete them:

   a. Take offline the failed:

      `storage aggregate plex offline -aggregate` ***`aggregate-name`*** **`-plex`** ***`plex-id`***

      The following example shows the aggregate `node_B_2_aggr1/plex1` being taken offline:

```
cluster_B::> storage aggregate plex offline -aggregate node_B_1_aggr0 -plex plex4

Plex offline successful on plex: node_B_1_aggr0/plex4
```

   b. Delete the failed plex:

      `storage aggregate plex delete -aggregate` ***`aggregate-name`*** **`-plex`** ***`plex-id`***

      You can destroy the plex when prompted.

      The following example shows the plex `node_B_2_aggr1/plex1` being deleted.

```
cluster_B::> storage aggregate plex delete -aggregate  node_B_1_aggr0 -plex plex4

Warning: Aggregate "node_B_1_aggr0" is being used for the local management root
         volume or HA partner management root volume, or has been marked as
         the aggregate to be used for the management root volume after a
         reboot operation. Deleting plex "plex4" for this aggregate could lead
         to unavailability of the root volume after a disaster recovery
         procedure. Use the "storage aggregate show -fields
         has-mroot,has-partner-mroot,root" command to view such aggregates.

Warning: Deleting plex "plex4" of mirrored aggregate "node_B_1_aggr0" on node
         "node_B_1" in a MetroCluster configuration will disable its
         synchronous disaster recovery protection. Are you sure you want to
         destroy this plex? {y|n}: y
[Job 633] Job succeeded: DONE

cluster_B::>
```

   You must repeat these steps for each of the failed plexes.

4. Confirm that the plexes have been removed:

storage aggregate plex show -fields aggregate,status,is-online,plex,pool

```
cluster_B::> storage aggregate plex show -fields aggregate,status,is-online,Plex,pool
aggregate    plex  status        is-online pool
------------ ----- ------------- --------- ----
node_B_1_aggr0 plex0 normal,active true     0
node_B_2_aggr0 plex0 normal,active true     0
node_B_1_aggr1 plex0 normal,active true     0
node_B_1_aggr2 plex0 normal,active true     0
node_B_2_aggr1 plex0 normal,active true     0
node_B_2_aggr2 plex0 normal,active true     0
node_A_1_aggr1 plex0 failed,inactive false  -
node_A_1_aggr1 plex4 normal,active true     1
node_A_1_aggr2 plex0 failed,inactive false  -
node_A_1_aggr2 plex1 normal,active true     1
node_A_2_aggr1 plex0 failed,inactive false  -
node_A_2_aggr1 plex4 normal,active true     1
node_A_2_aggr2 plex0 failed,inactive false  -
node_A_2_aggr2 plex1 normal,active true     1
14 entries were displayed.

cluster_B::>
```

5. Identify the switched-over aggregates:

   storage aggregate show -is-home false

   You can also use the `storage aggregate plex show -fields aggregate,status,is-online,plex,pool` command to identify plex 0 switched-over aggregates. They will have a status of `failed, inactive`.

   The following commands show four switched-over aggregates:

   - node_A_1_aggr1
   - node_A_1_aggr2
   - node_A_2_aggr1
   - node_A_2_aggr2

```
cluster_B::> storage aggregate show -is-home false

cluster_A Switched Over Aggregates:
Aggregate     Size Available Used% State   #Vols  Nodes            RAID Status
--------- -------- --------- ----- ------- ------ ---------------- ------------
node_A_1_aggr1 2.12TB  1.88TB   11% online      91 node_B_1        raid_dp,
                                                                   mirror
                                                                   degraded
node_A_1_aggr2 2.89TB  2.64TB    9% online      90 node_B_1        raid_tec,
                                                                   mirror
                                                                   degraded
node_A_2_aggr1 2.12TB  1.86TB   12% online      91 node_B_2        raid_dp,
                                                                   mirror
                                                                   degraded
node_A_2_aggr2 2.89TB  2.64TB    9% online      90 node_B_2        raid_tec,
                                                                   mirror
                                                                   degraded
4 entries were displayed.

cluster_B::>
```

6. Identify switched-over plexes:

   storage aggregate plex show -fields aggregate,status,is-online,Plex,pool

   You want to identify the plexes with a status of `failed, inactive`.

   The following commands show four switched-over aggregates:

```
cluster_B::> storage aggregate plex show -fields aggregate,status,is-online,Plex,pool
aggregate    plex  status        is-online pool
------------ ----- ------------- --------- ----
node_B_1_aggr0 plex0 normal,active true     0
node_B_2_aggr0 plex0 normal,active true     0
```

```
node_B_1_aggr1 plex0 normal,active true     0
node_B_1_aggr2 plex0 normal,active true     0
node_B_2_aggr1 plex0 normal,active true     0
node_B_2_aggr2 plex0 normal,active true     0
node_A_1_aggr1 plex0 failed,inactive false  -  <<<<-- Switched over aggr/Plex0
node_A_1_aggr1 plex4 normal,active true     1
node_A_1_aggr2 plex0 failed,inactive false  -  <<<<-- Switched over aggr/Plex0
node_A_1_aggr2 plex1 normal,active true     1
node_A_2_aggr1 plex0 failed,inactive false  -  <<<<-- Switched over aggr/Plex0
node_A_2_aggr1 plex4 normal,active true     1
node_A_2_aggr2 plex0 failed,inactive false  -  <<<<-- Switched over aggr/Plex0
node_A_2_aggr2 plex1 normal,active true     1
14 entries were displayed.

cluster_B::>
```

7. Delete the failed plex:

   `storage aggregate plex delete -aggregate node_A_1_aggr1 -plex plex0`

   You can destroy the plex when prompted.

   The following example shows the plex `node_A_1_aggr1/plex0` being deleted:

```
cluster_B::> storage aggregate plex delete -aggregate node_A_1_aggr1 -plex plex0

Warning: Aggregate "node_A_1_aggr1" hosts MetroCluster metadata volume
         "MDV_CRS_e8457659b8a711e78b3b00a0988fe74b_A". Deleting plex "plex0"
         for this aggregate can lead to the failure of configuration
         replication across the two DR sites. Use the "volume show -vserver
         <admin-vserver> -volume MDV_CRS*" command to verify the location of
         such volumes.

Warning: Deleting plex "plex0" of mirrored aggregate "node_A_1_aggr1" on node
         "node_A_1" in a MetroCluster configuration will disable its
         synchronous disaster recovery protection. Are you sure you want to
         destroy this plex? {y|n}: y
[Job 639] Job succeeded: DONE

cluster_B::>
```

   You must repeat these steps for each of the failed aggregates.

8. Verify that there are no failed plexes remaining on the surviving site.
   The following output shows that all plexes are normal, active, and online.

```
cluster_B::> storage aggregate plex show -fields aggregate,status,is-online,Plex,pool
aggregate    plex  status        is-online pool
------------ ----- ------------- --------- ----
node_B_1_aggr0 plex0 normal,active true     0
node_B_2_aggr0 plex0 normal,active true     0
node_B_1_aggr1 plex0 normal,active true     0
node_B_2_aggr2 plex0 normal,active true     0
node_B_1_aggr1 plex0 normal,active true     0
node_B_2_aggr2 plex0 normal,active true     0
node_A_1_aggr1 plex4 normal,active true     1
node_A_1_aggr2 plex1 normal,active true     1
node_A_2_aggr1 plex4 normal,active true     1
node_A_2_aggr2 plex1 normal,active true     1
10 entries were displayed.

cluster_B::>
```

Performing aggregate healing and restoring mirrors (MetroCluster IP configurations)

After replacing hardware and assigning disks, in systems running ONTAP 9.5 or earlier you can perform the MetroCluster healing operations. In all versions of ONTAP, you must then confirm that aggregates are mirrored and, if necessary, restart mirroring.

About this task

Starting with ONTAP 9.6, the healing operations are performed automatically when the disaster site nodes boot up. The healing commands are not required.

These steps are performed on the surviving cluster.

Steps

1. If you are using ONTAP 9.6 or later, you must verify that automatic healing completed successfully:
   a. Confirm that the heal-aggr-auto and heal-root-aggr-auto operations completed:

      `metrocluster operation history show`
      The following output shows that the operations have completed successfully on cluster_A.

      ```
      cluster_B::*> metrocluster operation history show
      Operation                    State          Start Time       End Time
      ---------------------------- -------------- ---------------- ----------------
      heal-root-aggr-auto          successful      2/25/2019 06:45:58
                                                                    2/25/2019 06:46:02
      heal-aggr-auto               successful     2/25/2019 06:45:48
                                                                    2/25/2019 06:45:52
      .
      .
      .
      ```

   b. Confirm that the disaster site is ready for switchback:

      `metrocluster node show`
      The following output shows that the operations have completed successfully on cluster_A.

      ```
      cluster_B::*> metrocluster node show
      DR                           Configuration  DR
      Group Cluster Node           State          Mirroring Mode
      ----- ------- -------------- -------------- --------- --------------------
      1     cluster_A
                    node_A_1       configured     enabled   heal roots completed
                    node_A_2       configured     enabled   heal roots completed
            cluster_B
                    node_B_1       configured     enabled   waiting for switchback recovery
                    node_B_2       configured     enabled   waiting for switchback recovery
      4 entries were displayed.
      ```

2. If you are using ONTAP 9.5 or earlier, you must perform aggregate healing:
   a. Verify the state of the nodes:

      `metrocluster node show`
      The following output shows that switchover has completed, so healing can be performed.

      ```
      cluster_B::> metrocluster node show
      DR                           Configuration  DR
      Group Cluster Node           State          Mirroring Mode
      ----- ------- -------------- -------------- --------- --------------------
      1     cluster_B
                    node_B_1       configured     enabled   switchover completed
                    node_B_2       configured     enabled   switchover completed
            cluster_A
                    node_A_1       configured     enabled   waiting for switchback recovery
                    node_A_2       configured     enabled   waiting for switchback recovery
      4 entries were displayed.

      cluster_B::>
      ```

   b. Perform the aggregates healing phase:

      `metrocluster heal -phase aggregates`
      The following output shows a typical aggregates healing operation.

      ```
      cluster_B::*> metrocluster heal -phase aggregates
      [Job 647] Job succeeded: Heal Aggregates is successful.

      cluster_B::*> metrocluster operation show
        Operation: heal-aggregates
            State: successful
       Start Time: 10/26/2017 12:01:15
         End Time: 10/26/2017 12:01:17
           Errors: -

      cluster_B::*>
      ```

   c. Verify that heal aggregates has completed and the disaster site is ready for switchback:

      `metrocluster node show`

The following output shows that the heal aggregates phase has completed on cluster_A.

```
cluster_B::> metrocluster node show
DR                               Configuration  DR
Group Cluster Node               State          Mirroring Mode
----- ------- ------------------ -------------- --------- --------------------
1     cluster_A
              node_A_1           configured     enabled   heal aggregates completed
              node_A_2           configured     enabled   heal aggregates completed
      cluster_B
              node_B_1           configured     enabled   waiting for switchback recovery
              node_B_2           configured     enabled   waiting for switchback recovery
4 entries were displayed.

cluster_B::>
```

3. If disks have been replaced, you must mirror the local and switched over aggregates:
   a. Display the aggregates:

      `storage aggregate show`

      ```
      cluster_B::> storage aggregate show
      cluster_B Aggregates:
      Aggregate     Size Available Used% State   #Vols  Nodes            RAID Status
      --------- -------- --------- ----- ------- ------ ---------------- ------------
      node_B_1_aggr0 1.49TB  74.12GB   95% online     1 node_B_1         raid4,
                                                                         normal
      node_B_2_aggr0 1.49TB  74.12GB   95% online     1 node_B_2         raid4,
                                                                         normal
      node_B_1_aggr1 3.14TB  3.04TB    3% online     15 node_B_1         raid_dp,
                                                                         normal
      node_B_1_aggr2 3.14TB  3.06TB    3% online     14 node_B_1         raid_tec,
                                                                         normal
      node_B_1_aggr1 3.14TB  2.99TB    5% online     37 node_B_2         raid_dp,
                                                                         normal
      node_B_1_aggr2 3.14TB  3.02TB    4% online     35 node_B_2         raid_tec,
                                                                         normal

      cluster_A Switched Over Aggregates:
      Aggregate     Size Available Used% State   #Vols  Nodes            RAID Status
      --------- -------- --------- ----- ------- ------ ---------------- ------------
      node_A_1_aggr1 2.36TB  2.12TB   10% online     91 node_B_1         raid_dp,
                                                                         normal
      node_A_1_aggr2 3.14TB  2.90TB    8% online     90 node_B_1         raid_tec,
                                                                         normal
      node_A_2_aggr1 2.36TB  2.10TB   11% online     91 node_B_2         raid_dp,
                                                                         normal
      node_A_2_aggr2 3.14TB  2.89TB    8% online     90 node_B_2         raid_tec,
                                                                         normal
      12 entries were displayed.

      cluster_B::>
      ```

   b. Mirror the aggregate:

      `storage aggregate mirror -aggregate` ***`aggregate-name`***

      The following output shows a typical mirroring operation.

      ```
      cluster_B::> storage aggregate mirror -aggregate node_B_1_aggr1

      Info: Disks would be added to aggregate "node_B_1_aggr1" on node "node_B_1" in
            the following manner:

            Second Plex

              RAID Group rg0, 6 disks (block checksum, raid_dp)
                Position   Disk                      Type                  Size
                ---------- ------------------------- ---------- ---------------
                dparity    5.20.6                    SSD                      -
                parity     5.20.14                   SSD                      -
                data       5.21.1                    SSD                894.0GB
                data       5.21.3                    SSD                894.0GB
                data       5.22.3                    SSD                894.0GB
                data       5.21.13                   SSD                894.0GB

            Aggregate capacity available for volume use would be 2.99TB.

      Do you want to continue? {y|n}: y
      ```

   c. Repeat the previous step for each of the aggregates from the surviving site.
   d. Wait for the aggregates to resynchronize; you can check the status with the `storage aggregate show` command.

      The following output shows that a number of aggregates are resynchronizing.

```
cluster_B::> storage aggregate show

cluster_B Aggregates:
Aggregate     Size Available Used% State   #Vols  Nodes            RAID Status
--------- -------- --------- ----- ------- ------ ---------------- ------------
node_B_1_aggr0 1.49TB  74.12GB   95% online     1 node_B_1         raid4,
                                                                   mirrored,
                                                                   normal
node_B_2_aggr0 1.49TB  74.12GB   95% online     1 node_B_2         raid4,
                                                                   mirrored,
                                                                   normal
node_B_1_aggr1 2.86TB  2.76TB    4% online     15 node_B_1         raid_dp,
                                                                   resyncing
node_B_1_aggr2 2.89TB  2.81TB    3% online     14 node_B_1         raid_tec,
                                                                   resyncing
node_B_2_aggr1 2.73TB  2.58TB    6% online     37 node_B_2         raid_dp,
                                                                   resyncing
node_B-2_aggr2 2.83TB  2.71TB    4% online     35 node_B_2         raid_tec,
                                                                   resyncing

cluster_A Switched Over Aggregates:
Aggregate     Size Available Used% State   #Vols  Nodes            RAID Status
--------- -------- --------- ----- ------- ------ ---------------- ------------
node_A_1_aggr1 1.86TB  1.62TB   13% online     91 node_B_1         raid_dp,
                                                                   resyncing
node_A_1_aggr2 2.58TB  2.33TB   10% online     90 node_B_1         raid_tec,
                                                                   resyncing
node_A_2_aggr1 1.79TB  1.53TB   14% online     91 node_B_2         raid_dp,
                                                                   resyncing
node_A_2_aggr2 2.64TB  2.39TB    9% online     90 node_B_2         raid_tec,
                                                                   resyncing
12 entries were displayed.
```

e. Confirm that all aggregates are online and have resynchronized:

storage aggregate plex show

The following output shows that all aggregates have resynchronized.

```
cluster_A::> storage aggregate plex show
  ()
                      Is      Is         Resyncing
Aggregate Plex        Online  Resyncing    Percent Status
--------- --------- ------- ---------- --------- ---------------
node_B_1_aggr0 plex0 true    false              - normal,active
node_B_1_aggr0 plex8 true    false              - normal,active
node_B_2_aggr0 plex0 true    false              - normal,active
node_B_2_aggr0 plex8 true    false              - normal,active
node_B_1_aggr1 plex0 true    false              - normal,active
node_B_1_aggr1 plex9 true    false              - normal,active
node_B_1_aggr2 plex0 true    false              - normal,active
node_B_1_aggr2 plex5 true    false              - normal,active
node_B_2_aggr1 plex0 true    false              - normal,active
node_B_2_aggr1 plex9 true    false              - normal,active
node_B_2_aggr2 plex0 true    false              - normal,active
node_B_2_aggr2 plex5 true    false              - normal,active
node_A_1_aggr1 plex4 true    false              - normal,active
node_A_1_aggr1 plex8 true    false              - normal,active
node_A_1_aggr2 plex1 true    false              - normal,active
node_A_1_aggr2 plex5 true    false              - normal,active
node_A_2_aggr1 plex4 true    false              - normal,active
node_A_2_aggr1 plex8 true    false              - normal,active
node_A_2_aggr2 plex1 true    false              - normal,active
node_A_2_aggr2 plex5 true    false              - normal,active
20 entries were displayed.
```

4. On systems running ONTAP 9.5 and earlier, perform the root-aggregates healing phase:

metrocluster heal -phase root-aggregates

```
cluster_B::> metrocluster heal -phase root-aggregates
[Job 651] Job is queued: MetroCluster Heal Root Aggregates Job.Oct 26 13:05:00
[Job 651] Job succeeded: Heal Root Aggregates is successful.
```

5. Verify that heal root-aggregates has completed and the disaster site is ready for switchback:
The following output shows that the heal roots phase has completed on cluster_A.

```
cluster_B::> metrocluster node show
DR                               Configuration  DR
Group Cluster Node               State          Mirroring Mode
----- ------- ------------------ -------------- --------- --------------------
1     cluster_A
              node_A_1           configured     enabled   heal roots completed
              node_A_2           configured     enabled   heal roots completed
      cluster_B
              node_B_1           configured     enabled   waiting for switchback recovery
              node_B_2           configured     enabled   waiting for switchback recovery
4 entries were displayed.

cluster_B::>
```

After you finish

Proceed to verify the licenses on the replaced nodes.

Verifying licenses on the replaced nodes on page 128

Preparing the nodes for switchback in a mixed configuration (recovery during transition)

You must perform certain tasks in order to prepare the mixed MetroCluster IP and FC configuration for the switchback operation. This procedure only applies to configurations that encountered a failure during the MetroCluster FC to IP transition process.

About this task

This procedure should only be used when performing recovery on a system that was in mid-transition when the failure occurred.

In this scenario, the MetroCluster is a mixed configuration::

- One DR group consists of fabric-attached MetroCluster FC nodes.
  You must perform the MetroCluster FC recovery steps on these nodes.
- One DR group consists of MetroCluster IP nodes.
  You must perform the MetroCluster IP recovery steps on these nodes.

Perform the steps in the following order.

Steps

1. Prepare the FC nodes for switchback by performing the following tasks in order:
   a. *Verifying port configuration (MetroCluster FC configurations only)* on page 57
   b. *Configuring the FC-to-SAS bridges (MetroCluster FC configurations only)* on page 58
   c. *Configuring the FC switches (MetroCluster FC configurations only)* on page 60
   d. *Verifying the storage configuration* on page 97 (only perform these steps on replaced drives on the MetroCluster FC nodes)
   e. *Powering on the equipment at the disaster site* on page 97 (only perform these steps on replaced drives on the MetroCluster FC nodes)
   f. *Assigning ownership for replaced drives* on page 99 (only perform these steps on replaced drives on the MetroCluster FC nodes)
   g. Perform the steps in *Reassigning disk ownership for root aggregates to replacement controller modules (MetroCluster FC configurations)* on page 103, up to and including the step to issue the `mailbox destroy` command.
   h. Destroy the local plex (plex 0) of the root aggregate:

      `aggr destroy plex-id`
   i. If the root aggr is not online, online it.
2. Boot the MetroCluster FC nodes.

   You must perform these steps on both of the MetroCluster FC nodes.
   a. Display the boot menu:

      `boot_ontap menu`

b. From the boot menu, select option 6, **Update flash from backup config**.

c. Respond y to the following prompt:

```
This will replace all flash-based configuration with the last backup to disks. Are
you sure you want to continue?: y
```

The system will boot twice, the second time to load the new configuration.

Note: If you did not clear the NVRAM contents of a used replacement controller, then you might see a panic with the following message:

```
PANIC: NVRAM contents are invalid...
```

If this occurs, repeat these substeps to boot the system to the ONTAP prompt. You will then need to perform a root recovery. Contact technical support for assistance.

3. Mirror the root aggregate on plex 0:

You must perform these steps on both of the MetroCluster FC nodes.

a. Assign three pool0 disks to the new controller module.

b. Mirror the root aggregate pool1 plex:

`aggr mirror root-aggr-name`

c. Assign unowned disks to pool0 on the local node

4. Return to Maintenance mode.

You must perform these steps on both of the MetroCluster FC nodes.

a. Halt the node:

`halt`

b. Boot the node to Maintenance mode:

`boot_ontap maint`

5. Delete the contents of the mailbox disks:

`mailbox destroy local`

You must perform these steps on both of the MetroCluster FC nodes.

6. Halt the nodes:

`halt`

7. After the nodes boot up, verify the status of the node:

`metrocluster node show`

```
siteA::*> metrocluster node show
DR                              Configuration  DR
Group Cluster Node              State          Mirroring Mode
----- ------- ----------------- -------------- --------- --------------------
1     siteA
              wmc66-a1          configured     enabled   waiting for switchback recovery
              wmc66-a2          configured     enabled   waiting for switchback recovery
      siteB
              wmc66-b1          configured     enabled   switchover completed
              wmc66-b2          configured     enabled   switchover completed
2     siteA
              wmc55-a1          -              -         -
              wmc55-a2          unreachable    -         -
      siteB
              wmc55-b1          configured     enabled   switchover completed
              wmc55-b2          configured
```

8. Prepare the MetroCluster IP nodes for switchback by performing the tasks in *Preparing for switchback in a MetroCluster IP configuration* on page 106 up to and including *Deleting failed plexes owned by the surviving site (MetroCluster IP configurations)* on page 119.

9. On the MetroCluster FC nodes, perform the steps in *Performing aggregate healing and restoring mirrors (MetroCluster FC configurations)* on page 102.

10. On the MetroCluster IP nodes, perform the steps in *Performing aggregate healing and restoring mirrors (MetroCluster IP configurations)* on page 122.

11. Proceed through the remaining tasks of the recovery process starting with *Reestablishing object stores for FabricPool configurations* on page 128.

Reestablishing object stores for FabricPool configurations

If one of the object stores in a FabricPool mirror was co-located with the MetroCluster disaster site and was destroyed, you must reestablish the object store and the FabricPool mirror.

About this task

- If the object-stores are remote and a MetroCluster site is destroyed, you do not need to rebuild the object store, and the original object store configurations as well as cold data contents are retained.
- For more information about FabricPool configurations, see the *Disks and Aggregates Power Guide*.
  Disk and aggregate management

.

Step

Follow the procedure "Replacing a FabricPool mirror on a MetroCluster configuration" in the *Disks and Aggregates Power Guide*.

Disk and aggregate management

Verifying licenses on the replaced nodes

You must install new licenses for the replacement nodes if the impaired nodes were using ONTAP features that require a standard (node-locked) license. For features with standard licenses, each node in the cluster should have its own key for the feature.

About this task

Until you install license keys, features requiring standard licenses continue to be available to the replacement node. However, if the impaired node was the only node in the cluster with a license for the feature, no configuration changes to the feature are allowed. Also, using unlicensed features on the node might put you out of compliance with your license agreement, so you should install the replacement license key or keys on the replacement node as soon as possible.

The licenses keys must be in the 28-character format.

You have a 90-day grace period in which to install the license keys. After the grace period, all old licenses are invalidated. After a valid license key is installed, you have 24 hours to install all of the keys before the grace period ends.

Note: If all nodes at a site have been replaced (a single node in the case of a two-node MetroCluster configuration), license keys must be installed on the replacement node or nodes prior to switchback.

Steps

1. Identify the licenses on the node:

   `license show`
   The following example displays the information about licenses in the system:

```
cluster_B::>  license show
         (system license show)

Serial Number: 1-80-00050
Owner: site1-01
Package           Type       Description            Expiration
-------           -------    -------------          -----------
Base              license    Cluster Base License        -
```

```
NFS             site        NFS License                -
CIFS            site        CIFS License               -
iSCSI           site        iSCSI License              -
FCP             site        FCP License                -
FlexClone       site        FlexClone License          -

6 entries were displayed.
```

2. Verify that the licenses are good for the node after switchback:

   `metrocluster check license show`

   The following example displays the licenses that are good for the node:

```
cluster_B::> metrocluster check license show

Cluster           Check                          Result
-------           -------                        -------------
Cluster_B         negotiated-switchover-ready    not-applicable
NFS               switchback-ready               not-applicable
CIFS              job-schedules                  ok
iSCSI             licenses                       ok
FCP               periodic-check-enabled         ok
```

3. If you need new license keys, obtain replacement license keys on the NetApp Support Site in the My Support section under Software licenses.

   Note: The new license keys that you require are automatically generated and sent to the email address on file. If you fail to receive the email with the license keys within 30 days, you should contact technical support.

4. Install each license key:

   `system license add -license-code *license-key*, *license-key*...`

5. Remove the old licenses, if desired:
   a. Check for unused licenses:

      `license clean-up -unused -simulate`

   b. If the list looks correct, remove the unused licenses:

      `license clean-up -unused`

Performing a switchback

After you heal the MetroCluster configuration, you can perform the MetroCluster switchback operation. The MetroCluster switchback operation returns the configuration to its normal operating state, with the sync-source storage virtual machines (SVMs) on the disaster site active and serving data from the local disk pools.

Before you begin

- The disaster cluster must have successfully switched over to the surviving cluster.
- Healing must have been performed on the data and root aggregates.
- The surviving cluster nodes must not be in the HA failover state (all nodes must be up and running for each HA pair).
- The disaster site controller modules must be completely booted and not in the HA takeover mode.
- The root aggregate must be mirrored.
- The Inter-Switch Links (ISLs) must be online.
- Any required licenses must be installed on the system.

Steps

1. Confirm that all nodes are in the enabled state:

   `metrocluster node show`

   The following example displays the nodes that are in the enabled state:

```
cluster_B::>  metrocluster node show

DR                        Configuration  DR
Group Cluster Node        State          Mirroring Mode
----- ------- ----------- -------------- --------- --------------------
1     cluster_A
              node_A_1    configured     enabled   heal roots completed
              node_A_2    configured     enabled   heal roots completed
      cluster_B
              node_B_1    configured     enabled   waiting for switchback recovery
              node_B_2    configured     enabled   waiting for switchback recovery
4 entries were displayed.
```

2. Confirm that resynchronization is complete on all SVMs:

   `metrocluster vserver show`

3. Verify that any automatic LIF migrations being performed by the healing operations have been successfully completed: `metrocluster check lif show`
4. Perform the switchback by running the `metrocluster switchback` command from any node in the surviving cluster.
5. Check the progress of the switchback operation:

   `metrocluster show`

   The switchback operation is still in progress when the output displays `waiting-for-switchback`:

```
cluster_B::> metrocluster show
Cluster                   Entry Name          State
------------------------- ------------------- -----------
 Local: cluster_B         Configuration state configured
                          Mode                switchover
                          AUSO Failure Domain -
Remote: cluster_A         Configuration state configured
                          Mode                waiting-for-switchback
                          AUSO Failure Domain -
```

   The switchback operation is complete when the output displays `normal`:

```
cluster_B::> metrocluster show
Cluster                   Entry Name          State
------------------------- ------------------- -----------
 Local: cluster_B         Configuration state configured
                          Mode                normal
                          AUSO Failure Domain -
Remote: cluster_A         Configuration state configured
                          Mode                normal
                          AUSO Failure Domain -
```

   If a switchback takes a long time to finish, you can check on the status of in-progress baselines by using the `metrocluster config-replication resync-status show` command. This command is at the advanced privilege level.
6. Reestablish any SnapMirror or SnapVault configurations.

   In ONTAP 8.3, you need to manually reestablish a lost SnapMirror configuration after a MetroCluster switchback operation. In ONTAP 9.0 and later, the relationship is reestablished automatically.

Verifying a successful switchback

After performing the switchback, you want to confirm that all aggregates and storage virtual machines (SVMs) are switched back and online.

Steps

1. Verify that the switched-over data aggregates are switched back:

`storage aggregate show`

In the following example, aggr_b2 on node B2 has switched back:

```
node_B_1::> storage aggregate show
Aggregate     Size Available Used% State   #Vols  Nodes            RAID Status
--------- -------- --------- ----- ------- ------ ---------------- ------------
...
aggr_b2    227.1GB   227.1GB    0% online       0 node_B_2   raid_dp,
                                                               mirrored,
                                                               normal

node_A_1::> aggr show
Aggregate     Size Available Used% State   #Vols  Nodes            RAID Status
--------- -------- --------- ----- ------- ------ ---------------- ------------
...
aggr_b2          -         -     - unknown      - node_A_1
```

If the disaster site included unmirrored aggregates and the unmirrored aggregates are no longer present, the aggregate may show up with a `State` of `unknown` in the output of the `storage aggregate show` command. Contact technical support to remove the out-of-date entries for the unmirrored aggregates.

2. Verify that all sync-destination SVMs on the surviving cluster are dormant (showing an `Admin State` of `stopped`) and the sync-source SVMs on the disaster cluster are up and running:

   `vserver show -subtype sync-source`

```
node_B_1::> vserver show -subtype sync-source
                              Admin      Root                        Name    Name
Vserver     Type    Subtype    State      Volume     Aggregate       Service Mapping
----------- ------- ---------- ---------- ---------- ----------      ------- -------
...
vs1a        data    sync-source
                              running    vs1a_vol   node_B_2        file    file
                                                                    aggr_b2

node_A_1::> vserver show -subtype sync-destination
                              Admin      Root                          Name    Name
Vserver            Type    Subtype    State      Volume     Aggregate  Service Mapping
-----------        ------- ---------- ---------- ---------- ---------- ------- -------
...
cluster_A-vs1a-mc  data    sync-destination
                                    stopped    vs1a_vol   sosb_      file    file
                                                                      aggr_b2
```

   Sync-destination aggregates in the MetroCluster configuration have the suffix "-mc" automatically appended to their name to help identify them.

3. Confirm that the switchback operations succeeded by using the `metrocluster operation show` command.

If the command output shows...	Then...
That the switchback operation state is successful.	The switchback process is complete and you can proceed with operation of the system.
That the switchback operation or switchback-continuation-agent operation is partially successful.	Perform the suggested fix provided in the output of the `metrocluster operation show` command.

After you finish

You must repeat the previous sections to perform the switchback in the opposite direction. If site_A did a switchover of site_B, have site_B do a switchover of site_A.

Mirroring the root aggregates of the replacement nodes

If disks were replaced, you must mirror the root aggregates of the new nodes on the disaster site.

Steps

1. On the disaster site, identify the aggregates which are not mirrored:

`storage aggregate show`

```
cluster_A::> storage aggregate show

Aggregate     Size Available Used% State   #Vols  Nodes            RAID Status
--------- -------- --------- ----- ------- ------ ---------------- ------------
node_A_1_aggr0
            1.49TB   74.12GB   95% online       1 node_A_1         raid4,
                                                                    normal
node_A_2_aggr0
            1.49TB   74.12GB   95% online       1 node_A_2         raid4,
                                                                    normal
node_A_1_aggr1
            1.49TB   74.12GB   95% online       1 node_A_1         raid 4, normal
                                                                    mirrored
node_A_2_aggr1
            1.49TB   74.12GB   95% online       1 node_A_2         raid 4, normal
                                                                    mirrored
4 entries were displayed.

cluster_A::>
```

2. Mirror one of the root aggregates:

   `storage aggregate mirror -aggregate` *`root-aggregate`*

   The following example shows how the command selects disks and prompts for confirmation when mirroring the aggregate.

```
cluster_A::> storage aggregate mirror -aggregate node_A_2_aggr0

Info: Disks would be added to aggregate "node_A_2_aggr0" on node "node_A_2" in
      the following manner:

      Second Plex

        RAID Group rg0, 3 disks (block checksum, raid4)
          Position   Disk                      Type                  Size
          ---------- ------------------------- ---------- ---------------
          parity     2.10.0                    SSD                     -
          data       1.11.19                   SSD                894.0GB
          data       2.10.2                    SSD                894.0GB

      Aggregate capacity available for volume use would be 1.49TB.

Do you want to continue? {y|n}: y

cluster_A::>
```

3. Verify that mirroring of the root aggregate is complete:

   `storage aggregate show`

   The following example shows that the root aggregates are mirrored.

```
cluster_A::> storage aggregate show

Aggregate     Size Available Used% State   #Vols  Nodes       RAID Status
--------- -------- --------- ----- ------- ------ ----------- ------------
node_A_1_aggr0
            1.49TB   74.12GB   95% online       1 node_A_1    raid4,
                                                               mirrored,
                                                               normal
node_A_2_aggr0
            2.24TB   838.5GB   63% online       1 node_A_2    raid4,
                                                               mirrored,
                                                               normal
node_A_1_aggr1
            1.49TB   74.12GB   95% online       1 node_A_1    raid4,
                                                               mirrored,

normal
node_A_2_aggr1
            1.49TB   74.12GB   95% online       1 node_A_2    raid4
                                                               mirrored,
                                                               normal
4 entries were displayed.

cluster_A::>
```

4. Repeat these steps for the other root aggregates.

   Any root aggregate that does not have a status of `mirrored` must be mirrored.

Reconfiguring the ONTAP Mediator service (MetroCluster IP configurations)

If you have a MetroCluster IP configuration that was configured with the ONTAP Mediator service, you must remove and reconfigure the association with the mediator.

Before you begin

- You must have the IP address and username and password for the ONTAP Mediator service.
- The ONTAP Mediator service must be configured and operating on the Linux host.

Steps

1. Remove the existing ONTAP Mediator configuration: **metrocluster configuration-settings mediator remove**
2. Reconfigure the ONTAP Mediator configuration: **metrocluster configuration-settings mediator add -mediator-address *mediator-IP-address***

Verifying the health of the MetroCluster configuration

You should check the health of the MetroCluster configuration to verify proper operation.

Steps

1. Check that the MetroCluster is configured and in normal mode on each cluster:

   metrocluster show

```
cluster_A::> metrocluster show
Cluster                   Entry Name          State
------------------------- ------------------- -----------
 Local: cluster_A         Configuration state configured
                          Mode                normal
                          AUSO Failure Domain auso-on-cluster-disaster
Remote: cluster_B         Configuration state configured
                          Mode                normal
                          AUSO Failure Domain auso-on-cluster-disaster
```

2. Check that mirroring is enabled on each node:

   metrocluster node show

```
cluster_A::> metrocluster node show
DR                       Configuration  DR
Group Cluster Node          State          Mirroring Mode
----- ------- ------------- -------------- --------- --------------------
1     cluster_A
              node_A_1       configured     enabled   normal
      cluster_B
              node_B_1       configured     enabled   normal
2 entries were displayed.
```

3. Check that the MetroCluster components are healthy:

   metrocluster check run

```
cluster_A::> metrocluster check run

Last Checked On: 10/1/2014 16:03:37

Component           Result
------------------- ---------
nodes               ok
lifs                ok
config-replication  ok
aggregates          ok
4 entries were displayed.

Command completed. Use the "metrocluster check show -instance" command or sub-commands
in "metrocluster check" directory for detailed results.
```

```
To check if the nodes are ready to do a switchover or switchback operation, run
"metrocluster switchover -simulate" or "metrocluster switchback -simulate",
respectively.
```

4. Check that there are no health alerts:

   system health alert show

5. Simulate a switchover operation:

   a. From any node's prompt, change to the advanced privilege level:

      set -privilege advanced

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

   b. Perform the switchover operation with the `-simulate` parameter:

      metrocluster switchover -simulate

   c. Return to the admin privilege level:

      set -privilege admin

6. For MetroCluster IP configurations using the ONTAP Mediator service, confirm that the Mediator service is up and operating.

   a. Check that the Mediator disks are visible to the system:

      storage failover mailbox-disk show
      The following example shows that the mailbox disks have been recognized.

```
node_A_1::*> storage failover mailbox-disk show
                    Mailbox
Node                Owner     Disk    Name        Disk UUID
------------     ------   -----   -----        ----------------
sti113-vsim-ucs626g
.
.
     local     0m.i2.3L26      7BBA77C9:AD702D14:831B3E7E:0B0730EE:
00000000:00000000:00000000:00000000:00000000:00000000
     local     0m.i2.3L27      928F79AE:
631EA9F9:4DCB5DE6:3402AC48:00000000:00000000:00000000:00000000:00000000
     local     0m.i1.0L60      B7BCDB3C:
297A4459:318C2748:181565A3:00000000:00000000:00000000:00000000:00000000
.
.
.
     partner   0m.i1.0L14
EA71F260:D4DD5F22:E3422387:61D475B2:00000000:00000000:00000000:00000000:00000000:00000000
     partner   0m.i2.3L64
4460F436:AAE5AB9E:D1ED414E:ABF811F7:00000000:00000000:00000000:00000000:00000000:00000000
28 entries were displayed.
```

   b. Change to the advanced privilege level:

      set -privilege advanced

   c. Check that the mailbox LUNs are visible to the system:

      storage iscsi-initiator show
      The output will show the presence of the mailbox LUNs:

```
Node    Type       Label      Target Portal     Target Name                                Admin/Op
----    ----       --------   ---------    --------- ---------------------------------       --------
.
.
.
.node_A_1
               mailbox
                     mediator 172.16.254.1    iqn.2012-05.local:mailbox.target.db5f02d6-e3d3    up/up
.
.
.
17 entries were displayed.
```

   d. Return to the administrative privilege level:

      set -privilege admin

Recovering from a non-controller failure

After the equipment at the disaster site has undergone any required maintenance or replacement, but no controllers were replaced, you can begin the process of returning the MetroCluster configuration to a fully redundant state. This includes healing the configuration (first the data aggregates and then the root aggregates) and performing the switchback operation.

Before you begin

- All MetroCluster hardware in the disaster cluster must be functional.
- The overall MetroCluster configuration must be in switchover.
- In a fabric-attached MetroCluster configuration, the ISL must be up and operating between the MetroCluster sites.

Steps

1. *Healing the configuration in a MetroCluster FC configuration* on page 135
2. *Verifying that your system is ready for a switchback* on page 137
3. *Performing a switchback* on page 138
4. *Verifying a successful switchback* on page 139
5. *Deleting stale aggregate listings after switchback* on page 140

Healing the configuration in a MetroCluster FC configuration

Following a switchover, you must perform the healing operations in specific order to restore MetroCluster functionality.

Before you begin

- Switchover must have been performed and the surviving site must be serving data.
- Nodes on the disaster site must be halted or remain powered off.
  They must not be fully booted during the healing process.
- Storage at the disaster site must be accessible (shelves are powered up, functional, and accessible).
- In fabric-attached MetroCluster configurations, inter-switch links (ISLs) must be up and operating.
- In four-node MetroCluster configurations, nodes in the surviving site must not be in HA failover state (all nodes must be up and running for each HA pair).

About this task

The healing operation must first be performed on the data aggregates, and then on the root aggregates.

Steps

1. *Healing the data aggregates* on page 136
2. *Healing the root aggregates after a disaster* on page 136

Healing the data aggregates

You must heal the data aggregates after repairing and replacing any hardware on the disaster site. This process resynchronizes the data aggregates and prepares the (now repaired) disaster site for normal operation. You must heal the data aggregates prior to healing the root aggregates.

About this task

The following example shows a forced switchover, where you bring the switched-over aggregate online. All configuration updates in the remote cluster successfully replicate to the local cluster. You power up the storage on the disaster site as part of this procedure, but you do not and must not power up the controller modules on the disaster site.

Steps

1. Verify that switchover was completed by running the `metrocluster operation show` command.

```
controller_A_1::> metrocluster operation show
  Operation: switchover
      State: successful
 Start Time: 7/25/2014 20:01:48
   End Time: 7/25/2014 20:02:14
     Errors: -
```

2. Resynchronize the data aggregates by running the `metrocluster heal -phase aggregates` command from the surviving cluster.

```
controller_A_1::> metrocluster heal -phase aggregates
[Job 130] Job succeeded: Heal Aggregates is successful.
```

   If the healing is vetoed, you have the option of reissuing the `metrocluster heal` command with the `–override-vetoes` parameter. If you use this optional parameter, the system overrides any soft vetoes that prevent the healing operation.

3. Verify that the operation has been completed by running the `metrocluster operation show` command.

```
controller_A_1::> metrocluster operation show
    Operation: heal-aggregates
      State: successful
Start Time: 7/25/2014 18:45:55
   End Time: 7/25/2014 18:45:56
     Errors: -
```

4. Check the state of the aggregates by running the `storage aggregate show` command.

```
controller_A_1::> storage aggregate show
Aggregate Size     Available Used% State   #Vols  Nodes        RAID Status
--------- -------- --------- ----- ------- ------ ------------ ------------
...
aggr_b2   227.1GB  227.1GB   0%    online  0      mcc1-a2      raid_dp, mirrored, normal...
```

5. If storage has been replaced at the disaster site, you might need to remirror the aggregates.

Healing the root aggregates after a disaster

After the data aggregates have been healed, you must heal the root aggregates in preparation for the switchback operation.

Before you begin

The data aggregates phase of the MetroCluster healing process must have been completed successfully.

Steps

1. Switch back the mirrored aggregates by running the `metrocluster heal -phase root-aggregates` command.

```
mcc1A::> metrocluster heal -phase root-aggregates
[Job 137] Job succeeded: Heal Root Aggregates is successful
```

   If the healing is vetoed, you have the option of reissuing the `metrocluster heal` command with the `–override-vetoes` parameter. If you use this optional parameter, the system overrides any soft vetoes that prevent the healing operation.

2. Ensure that the heal operation is complete by running the `metrocluster operation show` command on the destination cluster:

```
mcc1A::> metrocluster operation show
  Operation: heal-root-aggregates
      State: successful
 Start Time: 7/29/2014 20:54:41
   End Time: 7/29/2014 20:54:42
     Errors: -
```

3. Power up each controller module on the disaster site.
4. After nodes are booted, verify that the root aggregates are mirrored.

   If both plexes are present, any resynchronization will start automatically. If one plex has failed, that plex must be destroyed and the mirror recreated using the `storage aggregate mirror -aggregate` *`aggregate-name`* command to reestablish the mirror relationship.

Verifying that your system is ready for a switchback

If your system is already in the switchover state, you can use the `-simulate` option to preview the results of a switchback operation.

Steps

1. Simulate the switchback operation:
   a. From either surviving node's prompt, change to the advanced privilege level:

      `set -privilege advanced`

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

   b. Perform the switchback operation with the `-simulate` parameter:

      `metrocluster switchback -simulate`

   c. Return to the admin privilege level:

      `set -privilege admin`

2. Review the output that is returned.

   The output shows whether the switchback operation would run into errors.

Example of verification results

The following example shows the successful verification of a switchback operation:

```
cluster4::*> metrocluster switchback -simulate
  (metrocluster switchback)
[Job 130] Setting up the nodes and cluster components for the switchback
```

```
operation...DBG:backup_api.c:327:backup_nso_sb_vetocheck : MCC Switch Back
[Job 130] Job succeeded: Switchback simulation is successful.

cluster4::*> metrocluster op show
  (metrocluster operation show)
  Operation: switchback-simulate
      State: successful
 Start Time: 5/15/2014 16:14:34
   End Time: 5/15/2014 16:15:04
     Errors: -

cluster4::*> job show -name Me*
                            Owning
Job ID Name                 Vserver    Node           State
------ -------------------- ---------- -------------- ----------
130    MetroCluster Switchback
                            cluster4
                                       cluster4-01
                                                      Success
       Description: MetroCluster Switchback Job - Simulation
```

Performing a switchback

After you heal the MetroCluster configuration, you can perform the MetroCluster switchback operation. The MetroCluster switchback operation returns the configuration to its normal operating state, with the sync-source storage virtual machines (SVMs) on the disaster site active and serving data from the local disk pools.

Before you begin

- The disaster cluster must have successfully switched over to the surviving cluster.
- Healing must have been performed on the data and root aggregates.
- The surviving cluster nodes must not be in the HA failover state (all nodes must be up and running for each HA pair).
- The disaster site controller modules must be completely booted and not in the HA takeover mode.
- The root aggregate must be mirrored.
- The Inter-Switch Links (ISLs) must be online.
- Any required licenses must be installed on the system.

Steps

1. Confirm that all nodes are in the enabled state:

   `metrocluster node show`

   The following example displays the nodes that are in the enabled state:

   ```
   cluster_B::>  metrocluster node show

   DR                        Configuration  DR
   Group Cluster Node        State          Mirroring Mode
   ----- ------- ----------- -------------- --------- --------------------
   1     cluster_A
                 node_A_1    configured     enabled   heal roots completed
                 node_A_2    configured     enabled   heal roots completed
         cluster_B
                 node_B_1    configured     enabled   waiting for switchback recovery
                 node_B_2    configured     enabled   waiting for switchback recovery
   4 entries were displayed.
   ```

2. Confirm that resynchronization is complete on all SVMs:

   `metrocluster vserver show`

3. Verify that any automatic LIF migrations being performed by the healing operations have been successfully completed: `metrocluster check lif show`

4. Perform the switchback by running the `metrocluster switchback` command from any node in the surviving cluster.

5. Check the progress of the switchback operation:

   `metrocluster show`

   The switchback operation is still in progress when the output displays `waiting-for-switchback`:

```
cluster_B::> metrocluster show
Cluster                   Entry Name          State
------------------------- ------------------- -----------
 Local: cluster_B         Configuration state configured
                          Mode                switchover
                          AUSO Failure Domain -
Remote: cluster_A         Configuration state configured
                          Mode                waiting-for-switchback
                          AUSO Failure Domain -
```

   The switchback operation is complete when the output displays `normal`:

```
cluster_B::> metrocluster show
Cluster                   Entry Name          State
------------------------- ------------------- -----------
 Local: cluster_B         Configuration state configured
                          Mode                normal
                          AUSO Failure Domain -
Remote: cluster_A         Configuration state configured
                          Mode                normal
                          AUSO Failure Domain -
```

   If a switchback takes a long time to finish, you can check on the status of in-progress baselines by using the `metrocluster config-replication resync-status show` command. This command is at the advanced privilege level.

6. Reestablish any SnapMirror or SnapVault configurations.

   In ONTAP 8.3, you need to manually reestablish a lost SnapMirror configuration after a MetroCluster switchback operation. In ONTAP 9.0 and later, the relationship is reestablished automatically.

Verifying a successful switchback

After performing the switchback, you want to confirm that all aggregates and storage virtual machines (SVMs) are switched back and online.

Steps

1. Verify that the switched-over data aggregates are switched back:

   `storage aggregate show`

   In the following example, aggr_b2 on node B2 has switched back:

```
node_B_1::> storage aggregate show
Aggregate     Size Available Used% State   #Vols  Nodes            RAID Status
--------- -------- --------- ----- ------- ------ ---------------- ------------
...
aggr_b2    227.1GB   227.1GB    0% online       0 node_B_2   raid_dp,
                                                               mirrored,
                                                               normal

node_A_1::> aggr show
Aggregate     Size Available Used% State   #Vols  Nodes            RAID Status
--------- -------- --------- ----- ------- ------ ---------------- ------------
...
aggr_b2          -         -     - unknown      - node_A_1
```

   If the disaster site included unmirrored aggregates and the unmirrored aggregates are no longer present, the aggregate may show up with a `State` of `unknown` in the output of the `storage`

`aggregate show` command. Contact technical support to remove the out-of-date entries for the unmirrored aggregates.

2. Verify that all sync-destination SVMs on the surviving cluster are dormant (showing an `Admin State` of `stopped`) and the sync-source SVMs on the disaster cluster are up and running:

   `vserver show -subtype sync-source`

```
node_B_1::> vserver show -subtype sync-source
                              Admin      Root                        Name    Name
Vserver     Type    Subtype    State      Volume     Aggregate       Service Mapping
----------- ------- ---------- ---------- ---------- ----------      ------- -------
...
vs1a        data    sync-source
                              running    vs1a_vol   node_B_2        file    file
                                                                    aggr_b2

node_A_1::> vserver show -subtype sync-destination
                              Admin      Root                          Name    Name
Vserver            Type    Subtype    State      Volume     Aggregate  Service Mapping
-----------        ------- ---------- ---------- ---------- ---------- ------- -------
...
cluster_A-vs1a-mc  data    sync-destination
                                       stopped    vs1a_vol   sosb_      file    file
                                                                        aggr_b2
```

   Sync-destination aggregates in the MetroCluster configuration have the suffix "-mc" automatically appended to their name to help identify them.

3. Confirm that the switchback operations succeeded by using the `metrocluster operation show` command.

If the command output shows...	Then...
That the switchback operation state is successful.	The switchback process is complete and you can proceed with operation of the system.
That the switchback operation or switchback-continuation-agent operation is partially successful.	Perform the suggested fix provided in the output of the `metrocluster operation show` command.

After you finish

You must repeat the previous sections to perform the switchback in the opposite direction. If site_A did a switchover of site_B, have site_B do a switchover of site_A.

Deleting stale aggregate listings after switchback

In some circumstances after switchback, you might notice the presence of *stale* aggregates. Stale aggregates are aggregates that have been removed from ONTAP, but whose information remains recorded on disk. Stale aggregates are displayed in the nodeshell `aggr status -r` command but not in the `storage aggregate show` command. You can delete these records so that they no longer appear.

About this task

Stale aggregates can occur if you relocated aggregates while the MetroCluster configuration was in switchover. For example:

1. Site A switches over to Site B.
2. You delete the mirroring for an aggregate and relocate the aggregate from node_B_1 to node_B_2 for load balancing.
3. You perform aggregate healing.

At this point a stale aggregate appears on node_B_1, even though the actual aggregate has been deleted from that node. This aggregate appears in the output from the nodeshell `aggr status -r` command. It does not appear in the output of the `storage aggregate show` command.

Steps

1. Compare the output of the output of the `storage aggregate show` command and the nodeshell `aggr status -r` command:

   `storage aggregate show`

   `run local aggr status -r`

   Stale aggregates appear in the `run local aggr status -r` output but not in the `storage aggregate show` output. For example, the following aggregate might appear in the `run local aggr status -r` output:

```
Aggregate aggr05 (failed, raid_dp, partial) (block checksums)
Plex /aggr05/plex0 (offline, failed, inactive)
  RAID group /myaggr/plex0/rg0 (partial, block checksums)

 RAID Disk Device  HA  SHELF BAY CHAN Pool Type  RPM  Used (MB/blks)  Phys (MB/blks)
 --------- ------  ------------ ---- ---- ----  ----- --------------  --------------
 dparity   FAILED          N/A                        82/ -
 parity    0b.5    0b    -   -   SA:A   0 VMDISK  N/A 82/169472      88/182040
 data      FAILED          N/A                        82/ -
 data      FAILED          N/A                        82/ -
 data      FAILED          N/A                        82/ -
 data      FAILED          N/A                        82/ -
 data      FAILED          N/A                        82/ -
 data      FAILED          N/A                        82/ -
 Raid group is missing 7 disks.
```

2. Remove the stale aggregate:

   a. From either node's prompt, change to the advanced privilege level:

      `set -privilege advanced`

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

   b. Remove the stale aggregate:

      `aggregate remove-stale-record -aggregate` ***`aggregate_name`***

   c. Return to the admin privilege level:

      `set -privilege admin`

3. Confirm that the stale aggregate record was removed:

   `run local aggr status -r`

Commands for switchover, healing, and switchback

There are specific ONTAP commands for performing the MetroCluster disaster recovery processes.

If you want to...	Use this command...
Verify that switchover can be performed without errors or vetoes.	`metrocluster switchover -simulate` at the advanced privilege level
Verify that switchback can be performed without errors or vetoes.	`metrocluster switchback -simulate` at the advanced privilege level
Switch over to the partner nodes (negotiated switchover).	`metrocluster switchover`
Switch over to the partner nodes (forced switchover).	`metrocluster switchover -forced-on-disaster true`
Perform data aggregate healing.	`metrocluster heal -phase aggregates`
Perform root aggregate healing.	`metrocluster heal -phase root-aggregates`
Switch back to the home nodes.	`metrocluster switchback`

Monitoring the MetroCluster configuration

You can use ONTAP MetroCluster commands and Active IQ Unified Manager (formerly OnCommand Unified Manager) to monitor the health of a variety of software components and the state of MetroCluster operations.

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`

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

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

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.

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

   ```
   cluster_A::> metrocluster check aggregate show

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

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

```
                                          ownership-state           ok


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

18 entries were displayed.
```

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

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

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

Commands for checking and monitoring the MetroCluster configuration

There are specific ONTAP commands for monitoring the MetroCluster configuration and checking MetroCluster operations.

Commands for checking MetroCluster operations

If you want to...	Use this command...
Perform a check of the MetroCluster operations. **Note:** This command should not be used as the only command for pre-DR operation system validation.	`metrocluster check run`
View the results of the last check on MetroCluster operations.	`metrocluster show`
View results of check on configuration replication between the sites.	`metrocluster check config-replication show` `metrocluster check config-replication show-aggregate-eligibility`
View results of check on node configuration.	`metrocluster check node show`
View results of check on aggregate configuration.	`metrocluster check aggregate show`
View the LIF placement failures in the MetroCluster configuration.	`metrocluster check lif show`

Commands for monitoring the MetroCluster interconnect

If you want to...	Use this command...
Display the HA and DR mirroring status and information for the MetroCluster nodes in the cluster.	`metrocluster interconnect mirror show`

Commands for monitoring MetroCluster SVMs

If you want to...	Use this command...
View all SVMs in both sites in the MetroCluster configuration.	`metrocluster vserver show`

Detecting failures with NetApp MetroCluster Tiebreaker software

The Tiebreaker software resides on a Linux host. You need the Tiebreaker software only if you want to monitor 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, from a site failure.

After you install the Tiebreaker software on a Linux host, you can configure the clusters in a MetroCluster configuration to monitor for disaster conditions.

How the Tiebreaker software detects intersite connectivity failures

The MetroCluster Tiebreaker software alerts you if all connectivity between the sites is lost.

Types of network paths

Depending on the configuration, there are three types of network paths between the two clusters in a MetroCluster configuration:

FC network (present in fabric-attached MetroCluster configurations)

This type of network is composed of two redundant FC switch fabrics. Each switch fabric has two FC switches, with one switch of each switch fabric co-located with a cluster. Each cluster has two FC switches, one from each switch fabric. All of the nodes have FC (NV interconnect and FCP initiator) connectivity to each of the co-located FC switches. Data is replicated from cluster to cluster over the ISL.

Intercluster peering network

This type of network is composed of a redundant IP network path between the two clusters. The cluster peering network provides the connectivity that is required to mirror the storage virtual machine (SVM) configuration. The configuration of all of the SVMs on one cluster is mirrored by the partner cluster.

IP network (present in MetroCluster IP configurations)

This type of network is composed of two redundant IP switch networks. Each network has two IP switches, with one switch of each switch fabric co-located with a cluster. Each cluster has two IP switches, one from each switch fabric. All of the nodes have connectivity to each of the co-located FC switches. Data is replicated from cluster to cluster over the ISL.

Monitoring intersite connectivity

The Tiebreaker software regularly retrieves the status of intersite connectivity from the nodes. If NV interconnect connectivity is lost and the intercluster peering does not respond to pings, then the clusters assume that the sites are isolated and the Tiebreaker software triggers an alert as "AllLinksSevered". If a cluster identifies the "AllLinksSevered" status and the other cluster is not reachable through the network, then the Tiebreaker software triggers an alert as "disaster".

How the Tiebreaker software detects site failures

The NetApp MetroCluster Tiebreaker software checks the reachability of the nodes in a MetroCluster configuration and the cluster to determine whether a site failure has occurred. The Tiebreaker software also triggers an alert under certain conditions.

Components monitored by the Tiebreaker software

The Tiebreaker software monitors each controller in the MetroCluster configuration by establishing redundant connections through multiple paths to a node management LIF and to the cluster management LIF, both hosted on the IP network.

The Tiebreaker software monitors the following components in the MetroCluster configuration:

- Nodes through local node interfaces
- Cluster through the cluster-designated interfaces
- Surviving cluster to evaluate whether it has connectivity to the disaster site (NV interconnect, storage, and intercluster peering)

When there is a loss of connection between the Tiebreaker software and all of the nodes in the cluster and to the cluster itself, the cluster will be declared as "not reachable" by the Tiebreaker software. It takes around three to five seconds to detect a connection failure. If a cluster is unreachable from the Tiebreaker software, the surviving cluster (the cluster that is still reachable) must indicate that all of the links to the partner cluster are severed before the Tiebreaker software triggers an alert.

Note: All of the links are severed if the surviving cluster can no longer communicate with the cluster at the disaster site through FC (NV interconnect and storage) and intercluster peering.

Failure scenarios during which Tiebreaker software triggers an alert

The Tiebreaker software triggers an alert when the cluster (all of the nodes) at the disaster site is down or unreachable and the cluster at the surviving site indicates the "AllLinksSevered" status.

The Tiebreaker software does not trigger an alert (or the alert is vetoed) in the following scenarios:

- In an eight-node MetroCluster configuration, if one HA pair at the disaster site is down
- In a cluster with all of the nodes at the disaster site down, one HA pair at the surviving site down, and the cluster at the surviving site indicates the "AllLinksSevered" status

  The Tiebreaker software triggers an alert, but ONTAP vetoes that alert. In this situation, a manual switchover is also vetoed
- Any scenario in which the Tiebreaker software can either reach at least one node or the cluster interface at the disaster site, or the surviving site still can reach either node at the disaster site through either FC (NV interconnect and storage) or intercluster peering

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:

If...	Then...
ONTAP detects no NVRAM errors	File service starts normally.
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.

If...	Then...
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.

If you want to...	Use this command...
Create a new volume with NVFAIL enabled	`volume create -nvfail` **`on`**
Enable NVFAIL on an existing volume	`volume modify` **Note:** You set the `-nvfail` option to **`on`** to enable NVFAIL on the created volume.
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

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.

Related concepts

Monitoring and protecting the file system consistency using NVFAIL on page 147

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.

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 configuration and operation in NetApp's extensive documentation library.

MetroCluster and miscellaneous guides

Guide	Content
ONTAP 9 Documentation Center	• All MetroCluster guides
NetApp Technical Report 4375: NetApp MetroCluster for ONTAP 9.3	• A technical overview of the MetroCluster configuration and operation. • Best practices for MetroCluster configuration.
Fabric-attached MetroCluster installation and configuration	• Fabric-attached MetroCluster architecture • Cabling the configuration • Configuring the FC-to-SAS bridges • Configuring the FC switches • Configuring the MetroCluster in ONTAP
Stretch MetroCluster installation and configuration	• Stretch MetroCluster architecture • Cabling the configuration • Configuring the FC-to-SAS bridges • Configuring the MetroCluster in ONTAP
MetroCluster IP installation and configuration	• MetroCluster IP architecture • Cabling the configuration • Configuring the MetroCluster in ONTAP
MetroCluster Tiebreaker Software Installation and Configuration Guide	• Monitoring the MetroCluster configuration with the MetroCluster Tiebreaker software
Active IQ Unified Manager documentation *NetApp Documentation: Product Guides and Resources*	• Monitoring the MetroCluster configuration and performance
Copy-based transition	• Transitioning data from 7-Mode storage systems to clustered storage systems

Copyright and trademark

Copyright

Trademark