



# **Perform switchover, healing, and switchback**

## ONTAP MetroCluster

NetApp  
February 13, 2026

# Table of Contents

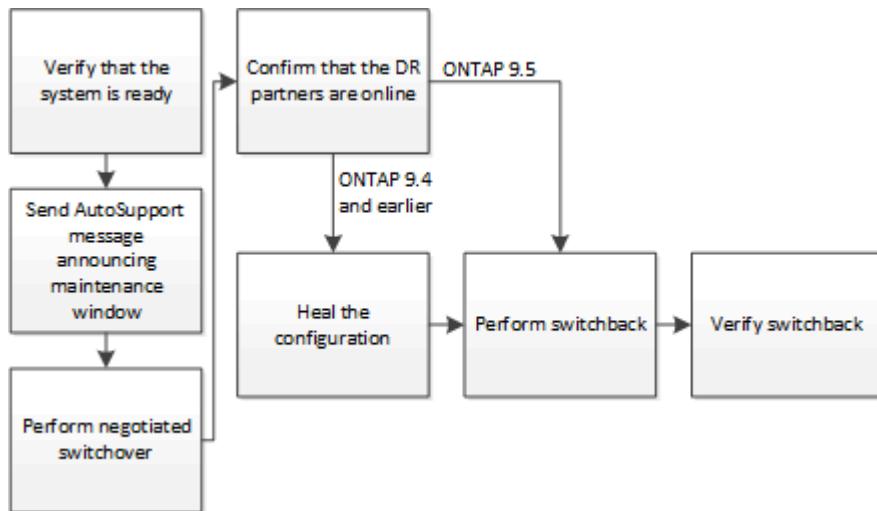
Perform switchover, healing, and switchback . . . . .	1
Perform switchover for tests or maintenance . . . . .	1
Performing switchover for tests or maintenance . . . . .	1
Limitations when the MetroCluster configuration is in switchover . . . . .	1
Verifying that your system is ready for a switchover . . . . .	2
Sending a custom AutoSupport message prior to negotiated switchover . . . . .	3
Performing a negotiated switchover . . . . .	3
Verify that the SVMs are running and the aggregates are online . . . . .	5
Heal the configuration . . . . .	6
Performing a switchback . . . . .	10
Verifying a successful switchback . . . . .	13
Commands for switchover, healing, and switchback . . . . .	14
Use System Manager to perform switchover and switchback (MetroCluster IP configurations only) . . . . .	15
Overview of switchover and switchback . . . . .	15
Use System Manager in ONTAP 9.6 or 9.7 for switchover and switchback . . . . .	15
Use System Manager in ONTAP 9.8 or later for switchover and switchback . . . . .	16
Monitoring the MetroCluster configuration . . . . .	17
Checking the MetroCluster configuration . . . . .	17
Commands for checking and monitoring the MetroCluster configuration . . . . .	20
Using the MetroCluster Tiebreaker or ONTAP Mediator to monitor the configuration . . . . .	21
How the NetApp MetroCluster Tiebreaker software detects failures . . . . .	21
Monitoring and protecting the file system consistency using NVFAIL . . . . .	23
How NVFAIL impacts access to NFS volumes or LUNs . . . . .	23
Commands for monitoring data loss events . . . . .	24
Accessing volumes in NVFAIL state after a switchover . . . . .	25
Recovering LUNs in NVFAIL states after switchover . . . . .	25
Where to find additional information . . . . .	26
MetroCluster and miscellaneous information . . . . .	26

# Perform switchover, healing, and switchback

## Perform switchover for tests or maintenance

### 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.



Beginning with ONTAP 9.6, switchover and switchback operations can be performed on MetroCluster IP configurations with ONTAP System Manager.

### Limitations when the MetroCluster configuration is in switchover

When the system is in switchover, certain operations should not be performed. Learn about restricted operations when the system is in switchover.

#### Restricted operations in switchover

The following operations are not supported when the system is in switchover:

- Creating or deleting aggregates and volumes
- Creating or deleting SVMs
- Creating or deleting LIFs
- Adding or removing disks (only if you are replacing them as part of a recovery procedure)
- Performing configuration changes to SnapMirror SVM DR
- Modifying existing broadcast domains or creating new broadcast domains
- Modifying network subnets

## Hardware replacement in switchover

Use the following procedures to replace controller hardware when the system is in switchover:

- If you need to replace a controller of the same type, at the site not in switchover, follow the procedure to [Recover from a multi-controller or storage failure](#).
  - If you need to replace the controller modules and chassis while the nodes are switched over at the surviving site, shut down both controllers and then perform the procedure to [Recover from a multi-controller or storage failure](#).
- If you need to replace a controller with a different type of controller, follow the procedure for your configuration in [Choose a controller upgrade procedure](#).
  - If your system is in switchover because of a controller failure or if you experience a controller failure while in switchover, you must first replace the controller hardware, perform a switchback, and then perform a controller upgrade:
    1. To replace the controller hardware and perform the switchback, follow [Recover from a multi-controller or storage failure](#).
    2. After you have replaced the hardware, perform a controller upgrade using the procedures in [Choose a controller upgrade procedure](#).

## 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. Issue these commands from the site that will remain up and operational:

1. Set the privilege level to advanced: `set -privilege advanced`
2. From the site that will remain up and operational, 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.



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.
```



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.

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.

- 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.

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:



Before performing maintenance tasks, you must remove monitoring if the MetroCluster configuration is monitored with the Tiebreaker or Mediator utility. [Remove ONTAP Mediator or Tiebreaker monitoring before performing maintenance tasks](#)

- 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.
    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. Issue these commands from the site that will remain up and operational.
    2. From the site that will remain up and operational, implement the switchover: `metrocluster switchover`

The operation can take several minutes to complete.

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.

## Verify that the SVMs are running and the aggregates are 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.

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.

### Steps

1. 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

.

.

.

mccl-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

```

2. 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

```

## Heal the configuration

### Heal the configuration in a MetroCluster FC configuration

## Healing the configuration in a MetroCluster FC configuration

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

- 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).

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

### 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.

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.

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-veto 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 -aggregateaggregate-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.



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

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).

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-vetoed 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-vetoed 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.

- 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.

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 a simulated switchback to verify the system is ready: `metrocluster switchback -simulate`
5. Check the configuration:

```
metrocluster check run
```

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

```

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

```

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

Component          Result
-----
nodes              ok
lifs               ok
config-replication ok
aggregates         ok
clusters           ok
connections        ok
6 entries were displayed.
```

6. Perform the switchback by running the metrocluster switchback command from any node in the surviving cluster: `metrocluster switchback`
7. 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.

8. 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.

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
```

2. Verify that all sync-destination SVMs on the surviving cluster are dormant (showing an operational state of "stopped"):

```
vserver show -subtype sync-destination
```

```
node_B_1::> vserver show -subtype sync-destination
                           Admin      Operational  Root
Vserver      Type      Subtype      State      State      Volume
Aggregate
-----
...
cluster_A-vs1a-mc data sync-destination
                           running     stopped    vs1a_vol    aggr_b2
```

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

3. Verify the sync-source SVMs on the disaster cluster are up and running:

```
vserver show -subtype sync-source
```

node_A_1::> vserver show -subtype sync-source					
Vserver	Type	Subtype	Admin State	Operational State	Root Volume
Aggregate					
-----	-----	-----	-----	-----	-----
...					
vs1a	data	sync-source	running	running	vs1a_vol aggr_b2

4. 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 <code>metrocluster operation show</code> command.

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.

## 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.	<code>metrocluster switchover -simulate</code> at the advanced privilege level
Verify that switchback can be performed without errors or vetoes.	<code>metrocluster switchback -simulate</code> at the advanced privilege level
Switch over to the partner nodes (negotiated switchover).	<code>metrocluster switchover</code>
Switch over to the partner nodes (forced switchover).	<code>metrocluster switchover -forced-on-disaster true</code>

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

## Use System Manager to perform switchover and switchback (MetroCluster IP configurations only)

You can switch over control from one MetroCluster IP site to the other to perform maintenance or recover from an issue.



Switchover and switchback procedures are supported only for MetroCluster IP configurations.

### Overview of switchover and switchback

A switchover can occur in two instances:

- **A planned switchover**

This switchover is initiated by a system administrator using System Manager. The planned switchover allows a system administrator of a local cluster to switch control so that the data services of the remote cluster are handled by the local cluster. Then, a system administrator at the remote cluster location can perform maintenance on the remote cluster.

- **An unplanned switchover**

In some cases, when a MetroCluster cluster goes down or the connections between the clusters are down, ONTAP automatically initiates a switchover so that the cluster that is still running handles the data handling responsibilities of the cluster that is down.

At other times, when ONTAP cannot determine the status of one of the clusters, the system administrator of the site that is working initiates the switchover to take control of the data handling responsibilities of the other site.

For any type of switchover procedure, the data servicing capability is returned to the cluster by using a *switchback* process.

The switchover and switchback process you follow depends on your ONTAP version:

- [Use System Manager in ONTAP 9.6 or 9.7 for switchover and switchback](#)
- [Use System Manager in ONTAP 9.8 or later for switchover and switchback](#)

### Use System Manager in ONTAP 9.6 or 9.7 for switchover and switchback

#### Steps

1. Log in to System Manager in ONTAP 9.6 or 9.7.
2. Click **(Return to classic version)**.

### 3. Click **Configuration > MetroCluster**.

System Manager verifies whether a negotiated switchover is possible.

### 4. Perform one of the following substeps when the validation process has completed:

- If validation fails, but Site B is up, then an error has occurred. For example, there might be a problem with a subsystem, or NVRAM mirroring might not be synchronized.

i. Fix the issue that is causing the error, click **Close**, and then start again at Step 2.

ii. Halt the Site B nodes, click **Close**, and then perform the steps in [Performing an unplanned switchover](#).

- If validation fails, and Site B is down, then there is most likely a connection problem. Verify that Site B is down, then perform the steps in [Performing an unplanned switchover](#).

### 5. Click **Switchover from Site B to Site A** to initiate the switchover process.

### 6. Click **Switch to the new experience**.

## Use System Manager in ONTAP 9.8 or later for switchover and switchback

### Perform a planned switchover (ONTAP 9.8 or later)

#### Steps

- Log in to System Manager in ONTAP 9.8 or later.
- Select **Dashboard**. In the **MetroCluster** section, the two clusters are shown with a connection.
- In the local cluster (shown on the left), click , and select **Switchover remote data services to the local site**.

After the switchover request is validated, control is transferred from the remote site to the local site. The local site performs data service requests for both clusters.

The remote cluster reboots, but the storage components are not active, and the cluster does not service data requests. It is now available for planned maintenance.



The remote cluster should not be used for data servicing until you perform a switchback.

### Perform an unplanned switchover (ONTAP 9.8 or later)

An unplanned switchover might be initiated automatically by ONTAP. If ONTAP cannot determine if a switchback is needed, the system administrator of the MetroCluster site that is still running initiates the switchover with the following steps:

#### Steps

- Log in to System Manager in ONTAP 9.8 or later.
- Select **Dashboard**.

In the **MetroCluster** section, the connection between the two clusters is shown with an "X" on it. This means that a connection cannot be detected and that either the connections or the cluster is down.

- In the local cluster (shown on the left), click , and select **Switchover remote data services to the local site**.

If the switchover fails with an error, click on the "View details" link in the error message and confirm the unplanned switchover.

After the switchover request is validated, control is transferred from the remote site to the local site. The local site performs data service requests for both clusters.

The cluster must be repaired before it is brought online again.



After the remote cluster is brought online, it should not be used for data servicing until you perform a switchback.

## Perform a switchback (ONTAP 9.8 or later)

### Before you begin

If the remote cluster was down due to planned maintenance or due to a disaster, it should now be up and running and waiting for the switchback.

### Steps

1. On the local cluster, log in to System Manager in ONTAP 9.8 or later.
2. Select **Dashboard**.

In the **MetroCluster** section, the two clusters are shown.

3. In the local cluster (shown on the left), click and select **Take back control**.

The data is *healed* first, to verify the data is synchronized and mirrored between both clusters.

4. When the data healing is complete, click and select **Initiate switchback**.

When the switchback is complete, both clusters are active and servicing data requests. Additionally, the data is being mirrored and synchronized between the clusters.

## 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.  <b>Note:</b> 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.	<code>metrocluster interconnect mirror show</code>

## Commands for monitoring MetroCluster SVMs

If you want to...	Use this command...
View all SVMs in both sites in the MetroCluster configuration.	<code>metrocluster vserver show</code>

## Using the MetroCluster Tiebreaker or ONTAP Mediator to monitor the configuration

See [Differences between ONTAP Mediator and MetroCluster Tiebreaker](#) to understand the differences between these two methods of monitoring your MetroCluster configuration and initiating an automatic switchover.

Use these links to install and configure Tiebreaker or Mediator:

- [Install and configure the MetroCluster Tiebreaker software](#)
- [Prepare to install ONTAP Mediator](#)

## How the NetApp MetroCluster Tiebreaker software detects failures

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 IP 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.



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

### How the ONTAP Mediator supports automatic unplanned switchover

[Learn about how the ONTAP Mediator supports automatic unplanned switchover in MetroCluster IP configurations.](#)

## 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.

<p>ONTAP detects NVRAM errors</p>	<ul style="list-style-type: none"> <li>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.</li> </ul> <p>ONTAP then sends an error message to the system console and log file.</p> <ul style="list-style-type: none"> <li>When the application restarts, files are available to CIFS clients even if you have not verified that they are valid.</li> </ul> <p>For NFS clients, files remain inaccessible until you reset the <code>in-nvfailed-state</code> option on the affected volume.</p>
<p>If one of the following parameters is used:</p> <ul style="list-style-type: none"> <li><code>dr-force-nvfail</code> volume option is set</li> <li><code>force-nvfail-all</code> switchover command option is set.</li> </ul>	<p>You can unset the <code>dr-force-nvfail</code> option after the switchover, if the administrator is not expecting to force NVFAIL processing for possible future disaster switchover operations. For NFS clients, files remain inaccessible until you reset the <code>in-nvfailed-state</code> option on the affected volume.</p> <p> Using the <code>force-nvfail-all</code> option causes the <code>dr-force-nvfail</code> option to be set on all of the DR volumes processed during the disaster switchover.</p>
<p>ONTAP detects NVRAM errors on a volume that contains LUNs</p>	<p>LUNs in that volume are brought offline. The <code>in-nvfailed-state</code> 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 backup as necessary. After all of the LUNs in the volume are recovered, the <code>in-nvfailed-state</code> option on the affected volume is cleared.</p>

## Commands for monitoring data loss events

If you enable the NVFAIL option, you receive notification when a system crash caused by NVRAM inconsistencies or a MetroCluster switchover occurs.

By default, the NVFAIL parameter is not enabled.

<p>If you want to...</p>	<p>Use this command...</p>
--------------------------	----------------------------

Create a new volume with NVFAIL enabled	volume create -nvfail on
Enable NVFAIL on an existing volume	<p>volume modify</p> <p><b>Note:</b> You set the <code>-nvfail</code> option to "on" to enable NVFAIL on the created volume.</p>
Display whether NVFAIL is currently enabled for a specified volume	<p>volume show</p> <p><b>Note:</b> You set the <code>-fields</code> parameter to "nvfail" to display the NVFAIL attribute for a specified volume.</p>

#### Related information

See the man page for each command for more information.

## Accessing volumes in NVFAIL state after a switchover

After a switchover, you must clear the NVFAIL state by resetting the `-in-nvfailed-state` parameter of the `volume modify` command to remove the restriction of clients to access data.

#### Before you begin

The database or file system must not be running or trying to access the affected volume.

#### About this task

Setting `-in-nvfailed-state` parameter requires advanced-level privilege.

#### Step

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

#### After you finish

For instructions about examining database file validity, see the documentation for your specific database software.

If your database uses LUNs, review the steps to make the LUNs accessible to the host after an NVRAM failure.

#### Related information

[Monitoring and protecting the files system consistency using NVFAIL](#)

## 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.

### MetroCluster and miscellaneous information

Information	Subject
<a href="#">MetroCluster Documentation</a>	<ul style="list-style-type: none"> <li>• All MetroCluster information</li> </ul>
<a href="#">NetApp Technical Report 4375: NetApp MetroCluster for ONTAP 9.3</a>	<ul style="list-style-type: none"> <li>• A technical overview of the MetroCluster configuration and operation.</li> <li>• Best practices for MetroCluster configuration.</li> </ul>
<a href="#">Fabric-attached MetroCluster installation and configuration</a>	<ul style="list-style-type: none"> <li>• Fabric-attached MetroCluster architecture</li> <li>• Cabling the configuration</li> <li>• Configuring the FC-to-SAS bridges</li> <li>• Configuring the FC switches</li> <li>• Configuring the MetroCluster in ONTAP</li> </ul>
<a href="#">Stretch MetroCluster installation and configuration</a>	<ul style="list-style-type: none"> <li>• Stretch MetroCluster architecture</li> <li>• Cabling the configuration</li> <li>• Configuring the FC-to-SAS bridges</li> <li>• Configuring the MetroCluster in ONTAP</li> </ul>
<a href="#">MetroCluster IP installation and configuration</a>	<ul style="list-style-type: none"> <li>• MetroCluster IP architecture</li> <li>• Cabling the configuration</li> <li>• Configuring the MetroCluster in ONTAP</li> </ul>
<a href="#">MetroCluster Tiebreaker 1.21 software installation and configuration</a>	<ul style="list-style-type: none"> <li>• Monitoring the MetroCluster configuration with the MetroCluster Tiebreaker software</li> </ul>
<a href="#">Active IQ Unified Manager documentation</a> <a href="#">NetApp Documentation: Product Guides and Resources</a>	<ul style="list-style-type: none"> <li>• Monitoring the MetroCluster configuration and performance</li> </ul>

## Copy-based transition

- Transitioning data from 7-Mode storage systems to clustered storage systems

## Copyright information

Copyright © 2026 NetApp, Inc. All Rights Reserved. Printed in the U.S. No part of this document covered by copyright may be reproduced in any form or by any means—graphic, electronic, or mechanical, including photocopying, recording, taping, or storage in an electronic retrieval system—with prior written permission of the copyright owner.

Software derived from copyrighted NetApp material is subject to the following license and disclaimer:

THIS SOFTWARE IS PROVIDED BY NETAPP “AS IS” AND WITHOUT ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE, WHICH ARE HEREBY DISCLAIMED. IN NO EVENT SHALL NETAPP BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

NetApp reserves the right to change any products described herein at any time, and without notice. NetApp assumes no responsibility or liability arising from the use of products described herein, except as expressly agreed to in writing by NetApp. The use or purchase of this product does not convey a license under any patent rights, trademark rights, or any other intellectual property rights of NetApp.

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

**LIMITED RIGHTS LEGEND:** Use, duplication, or disclosure by the government is subject to restrictions as set forth in subparagraph (b)(3) of the Rights in Technical Data -Noncommercial Items at DFARS 252.227-7013 (FEB 2014) and FAR 52.227-19 (DEC 2007).

Data contained herein pertains to a commercial product and/or commercial service (as defined in FAR 2.101) and is proprietary to NetApp, Inc. All NetApp technical data and computer software provided under this Agreement is commercial in nature and developed solely at private expense. The U.S. Government has a non-exclusive, non-transferrable, nonsublicensable, worldwide, limited irrevocable license to use the Data only in connection with and in support of the U.S. Government contract under which the Data was delivered. Except as provided herein, the Data may not be used, disclosed, reproduced, modified, performed, or displayed without the prior written approval of NetApp, Inc. United States Government license rights for the Department of Defense are limited to those rights identified in DFARS clause 252.227-7015(b) (FEB 2014).

## Trademark information

NETAPP, the NETAPP logo, and the marks listed at <http://www.netapp.com/TM> are trademarks of NetApp, Inc. Other company and product names may be trademarks of their respective owners.