



Performing switchover for tests or maintenance

ONTAP MetroCluster

NetApp
December 10, 2021

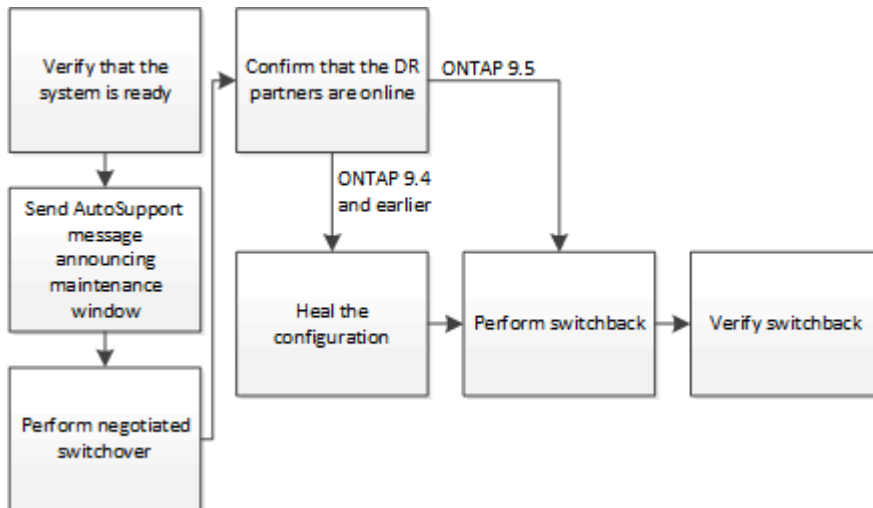
This PDF was generated from https://docs.netapp.com/us-en/ontap-metrocluster/manage/task_verify_that_your_system_is_ready_for_a_switchover.html on December 10, 2021. Always check docs.netapp.com for the latest.

Table of Contents

- Performing switchover for tests or maintenance 1
 - Verifying that your system is ready for a switchover 1
 - Sending a custom AutoSupport message prior to negotiated switchover 2
 - Performing a negotiated switchover 2
 - Output for the storage aggregate plex show command is indeterminate after a MetroCluster switchover ... 4
 - Confirming that the DR partners have come online 4
 - Healing the configuration 6
 - Performing a switchback 9
 - Verifying a successful switchback 12

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.



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

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:

- 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.

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_aggr1) 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
```

Name	Name			Admin	Operational	Root
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.

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

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 -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 -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.

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-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.

- 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 Group	Cluster	Node	Configuration State	DR 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
		node_B_2	configured	enabled	waiting for
					switchback recovery
					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

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.

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.

Copyright Information

Copyright © 2021 NetApp, Inc. All rights reserved. Printed in the U.S. No part of this document covered by copyright may be reproduced in any form or by any means-graphic, electronic, or mechanical, including photocopying, recording, taping, or storage in an electronic retrieval system- without prior written permission of the copyright owner.

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

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

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

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

RESTRICTED RIGHTS LEGEND: Use, duplication, or disclosure by the government is subject to restrictions as set forth in subparagraph (c)(1)(ii) of the Rights in Technical Data and Computer Software clause at DFARS 252.277-7103 (October 1988) and FAR 52-227-19 (June 1987).

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.