

# **Testing the MetroCluster configuration**

**ONTAP MetroCluster** 

Thom Illingworth, Ivana Devine July 29, 2021

This PDF was generated from https://docs.netapp.com/us-en/ontap-metrocluster/install-fc/task\_test\_the\_mcc\_configuration.html on September 24, 2021. Always check docs.netapp.com for the latest

# **Table of Contents**

Testing the MetroCluster configuration	. 1
Verifying negotiated switchover	. 1
Verifying healing and manual switchback	. 2
Loss of a single FC-to-SAS bridge	. 5
Verifying operation after power line disruption	. 7
Verifying operation after a switch fabric failure	. 8
Verifying operation after loss of a single storage shelf.	10

# **Testing the MetroCluster configuration**

You can test failure scenarios to confirm the correct operation of the MetroCluster configuration.

## Verifying negotiated switchover

You can test the negotiated (planned) switchover operation to confirm uninterrupted data availability.

#### About this task

This test validates that data availability is not affected (except for Microsoft Server Message Block (SMB) and Solaris Fibre Channel protocols) by switching the cluster over to the second data center.

This test should take about 30 minutes.

This procedure has the following expected results:

• The metrocluster switchover command will present a warning prompt.

If you respond yes to the prompt, the site the command is issued from will switch over the partner site.

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. Confirm that all nodes are in the configured state and normal mode:

```
metrocluster node show
```

2. Begin the switchover operation:

```
metrocluster switchover
```

```
cluster_A::> metrocluster switchover
Warning: negotiated switchover is about to start. It will stop all the
data Vservers on cluster "cluster_B" and
automatically re-start them on cluster "cluster_A". It will finally
gracefully shutdown cluster "cluster_B".
```

3. Confirm that the local cluster is in the configured state and switchover mode:

metrocluster node show

4. Confirm that the switchover operation was successful:

metrocluster operation show

```
cluster_A::> metrocluster operation show

cluster_A::> metrocluster operation show
   Operation: switchover
        State: successful
Start Time: 2/6/2016 13:28:50
   End Time: 2/6/2016 13:29:41
        Errors: -
```

5. Use the vserver show and network interface show commands to verify that DR SVMs and LIFs have come online.

## Verifying healing and manual switchback

You can test the healing and manual switchback operations to verify that data availability is not affected (except for SMB and Solaris FC configurations) by switching back the cluster to the original data center after a negotiated switchover.

#### About this task

This test should take about 30 minutes.

The expected result of this procedure is that services should be switched back to their home nodes.

#### **Steps**

1. Verify that healing is completed:

```
metrocluster node show
```

The following example shows the successful completion of the command:

2. Verify that all aggregates are mirrored:

```
storage aggregate show
```

The following example shows that all aggregates have a RAID Status of mirrored:

cluster_A::> stora		te show	N			
cluster Aggregates Aggregate Size Status		Used%	State	#Vols	Nodes	RAID
data_cluster 4.19TB	4.13TB	2%	online	8	node_A_1	<pre>raid_dp, mirrored, normal</pre>
root_cluster 715.5GB	212.7GB	70%	online	1	node_A_1	raid4, mirrored, normal
cluster_B Switched						
Aggregate Size Status	Available	Used%	State	#Vols	Nodes	RAID
data_cluster_B						
	4.11TB	2%	online	5	node_A_1	<pre>raid_dp, mirrored, normal</pre>
root_cluster_B		-	- unknow	n	- node_A_1	-

- 3. Boot the nodes from the disaster site.
- 4. Check the status of switchback recovery:

metrocluster node show

#### Perform the switchback:

metrocluster switchback

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

Confirm the status of the nodes:

metrocluster node show

7. Confirm the status of the metrocluster operation:

```
metrocluster operation show
```

The output should show a successful state.

```
cluster_A::> metrocluster operation show
  Operation: switchback
     State: successful
Start Time: 2/6/2016 13:54:25
    End Time: 2/6/2016 13:56:15
     Errors: -
```

## Loss of a single FC-to-SAS bridge

You can test the failure of a single FC-to-SAS bridge to make sure there is no single point of failure.

### About this task

This test should take about 15 minutes.

This procedure has the following expected results:

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



Starting with ONTAP 9.8, the storage bridge command is replaced with system bridge. The following steps show the storage bridge command, but if you are running ONTAP 9.8 or later, the system bridge command is preferred.

#### Steps

- 1. Turn off the power supplies of the bridge.
- 2. Confirm that the bridge monitoring indicates an error:

storage bridge show

```
cluster_A::> storage bridge show

Is

Monitor

Bridge Symbolic Name Vendor Model Bridge WWN Monitored
Status

------

ATTO_10.65.57.145

bridge_A_1 Atto FibreBridge 6500N

200000108662d46c true
```

3. Confirm that the drives behind the bridge are available with a single path:

storage disk error show

```
cluster A::> storage disk error show
Disk
                Error Type
                             Error Text
1.0.0
                                 1.0.0 (5000cca057729118): All paths
                onedomain
to this array LUN are connected to the same fault domain. This is a
single point of failure.
1.0.1
                                  1.0.1 (5000cca057727364): All paths
to this array LUN are connected to the same fault domain. This is a
single point of failure.
                onedomain
                                  1.0.2 (5000cca05772e9d4): All paths
to this array LUN are connected to the same fault domain. This is a
single point of failure.
1.0.23
                onedomain 1.0.23 (5000cca05772e9d4): All paths
to this array LUN are connected to the same fault domain. This is a
single point of failure.
```

## Verifying operation after power line disruption

You can test the MetroCluster configuration's response to the failure of a PDU.

#### About this task

The best practice is for each power supply unit (PSU) in a component to be connected to separate power supplies. If both PSUs are connected to the same power distribution unit (PDU) and an electrical disruption occurs, the site could down or a complete shelf might become unavailable. Failure of one power line is tested to confirm that there is no cabling mismatch that could cause a service disruption.

This test should take about 15 minutes.

This test requires turning off power to all left-hand PDUs and then all right-hand PDUs on all of the racks containing the MetroCluster components.

This procedure has the following expected results:

- Errors should be generated as the PDUs are disconnected.
- · No failover or loss of service should occur.

#### **Steps**

- 1. Turn off the power of the PDUs on the left-hand side of the rack containing the MetroCluster components.
- 2. Monitor the result on the console by using the following commands:

```
system environment sensors show -state fault
storage shelf show -errors
```

```
cluster A::> system environment sensors show -state fault
Node Sensor
                       State Value/Units Crit-Low Warn-Low Warn-Hi
Crit-Hi
node A 1
       PSU1
                      fault
                           PSU OFF
       PSU1 Pwr In OK fault
                           FAULT
node A 2
                       fault
       PSU1
                           PSU OFF
       PSU1 Pwr In OK fault
                           FAULT
4 entries were displayed.
cluster A::> storage shelf show -errors
   Shelf Name: 1.1
     Shelf UID: 50:0a:09:80:03:6c:44:d5
 Serial Number: SHFHU1443000059
Error Type
              Description
                   Critical condition is detected in storage shelf
power supply unit "1". The unit might fail. Reconnect PSU1
```

- 3. Turn the power back on to the left-hand PDUs.
- 4. Make sure that ONTAP clears the error condition.
- 5. Repeat the previous steps with the right-hand PDUs.

## Verifying operation after a switch fabric failure

You can disable a switch fabric to show that data availability is not affected by the loss.

#### About this task

This test should take about 15 minutes.

The expected result of this procedure is that disabling a fabric results in all cluster interconnect and disk traffic flowing to the other fabric.

In the examples shown, switch fabric 1 is disabled. This fabric consists of two switches, one at each MetroCluster site:

• FC switch A 1 on cluster A

• FC\_switch\_B\_1 on cluster\_B

#### Steps

- 1. Disable connectivity to one of the two switch fabrics in the MetroCluster configuration:
  - a. Disable the first switch in the fabric:

switchdisable

```
FC_switch_A_1::> switchdisable
```

b. Disable the second switch in the fabric:

switchdisable

```
FC_switch_B_1::> switchdisable
```

Monitor the result on the console of the controller modules.

You can use the following commands to check the cluster nodes to make sure that all data is still being served. The command output shows missing paths to disks. This is expected.

- vserver show
- network interface show
- · aggr show
- system node runnodename-command storage show disk -p
- storage disk error show
- 3. Reenable connectivity to one of the two switch fabrics in the MetroCluster configuration:
  - a. Reenable the first switch in the fabric:

switchenable

```
FC_switch_A_1::> switchenable
```

b. Reenable the second switch in the fabric:

switchenable

```
FC_switch_B_1::> switchenable
```

4. Wait at least 10 minutes and then repeat the above steps on the other switch fabric.

## Verifying operation after loss of a single storage shelf

You can test the failure of a single storage shelf to verify that there is no single point of failure.

#### About this task

This procedure has the following expected results:

- An error message should be reported by the monitoring software.
- · No failover or loss of service should occur.
- · Mirror resynchronization starts automatically after the hardware failure is restored.

### **Steps**

1. Check the storage failover status:

storage failover show

2. Check the aggregate status:

```
storage aggregate show
```

```
cluster A::> storage aggregate show
cluster Aggregates:
Aggregate Size Available Used% State #Vols Nodes RAID
Status
_____ _____
-----
node A 1data01 mirrored
        4.15TB 3.40TB 18% online 3 node A_1
raid dp,
mirrored,
normal
node A 1root
       707.7GB 34.29GB 95% online 1 node A_1
raid dp,
mirrored,
normal
node_A_2_data01_mirrored
        4.15TB 4.12TB 1% online 2 node_A_2
raid_dp,
mirrored,
normal
node A 2 data02 unmirrored
        raid_dp,
normal
node A 2 root
       707.7GB 34.27GB 95% online 1 node_A_2
raid dp,
mirrored,
normal
```

3. Verify that all data SVMs and data volumes are online and serving data:

vserver show -type data

```
cluster A::> vserver show -type data
cluster A::> vserver show -type data
                      Admin Operational Root
State State Volum
                              State Volume
Vserver
         Type Subtype
Aggregate
SVM1 data sync-source running SVM1_root
node A 1 data01 mirrored
SVM2
        data sync-source running SVM2 root
node A 2 data01 mirrored
cluster A::> network interface show -fields is-home false
There are no entries matching your query.
cluster A::> volume show !vol0,!MDV*
Vserver Volume Aggregate State Type Size
Available Used%
----
SVM1
       SVM1 root
                 node A 1data01 mirrored
                           online RW 10GB
9.50GB
      5%
SVM1
       SVM1 data vol
                 node A 1data01 mirrored
                          online RW
                                        10GB
9.49GB
       5%
SVM2
       SVM2 root
                 node A 2 data01 mirrored
                          online RW
                                           10GB
9.49GB
       5%
SVM2
       SVM2 data vol
                 node A 2 data02 unmirrored
                           online RW
                                            1GB
972.6MB
      5%
```

4. Identify a shelf in Pool 1 for node node\_A\_2 to power off to simulate a sudden hardware failure:

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

The shelf you select must contain drives that are part of a mirrored data aggregate.

In the following example, shelf ID 31 is selected to fail.

<pre>cluster_A::&gt; storage aggregate show -r -node node_A_2 !*root Owner Node: node_A_2</pre>					
Aggregate: node_A_2_data01_mirrored (online, raid_dp, mirrored) (block checksums)					
Plex: /node_A_2_data01_mirrored/pl RAID Group /node_A_2_data01_mirro				_	
checksums)				Usable	
Physical					
Position Disk Size Status	Pool	Туре	RPM	Size	
dparity 2.30.3	0	BSAS	7200	827.7GB	
828.0GB (normal)	0		7000	007 7CD	
parity 2.30.4 828.0GB (normal)	0	BSAS	7200	827.7GB	
data 2.30.6	0	BSAS	7200	827.7GB	
828.0GB (normal)	Ç	20110	, _ 0 0	0217	
data 2.30.8	0	BSAS	7200	827.7GB	
828.0GB (normal)					
data 2.30.5	0	BSAS	7200	827.7GB	
828.0GB (normal)					
Plex: /node A 2 data01 mirrored/pl	.ex4 (online	e, norr	mal, act	ive, pool1)	
RAID Group /node_A_2_data01_mirro	red/plex4/	rg0 (no	ormal, b	lock	
checksums)					
				Usable	
Physical	_				
Position Disk	Pool	Type	RPM	Size	
Size Status					
dparity 1.31.7	1	BSAS	7200	827.7GB	
828.0GB (normal)					
parity 1.31.6	1	BSAS	7200	827.7GB	
828.0GB (normal)	4	D. G. T. C.	7000	007 75-	
data 1.31.3	1	BSAS	7200	827.7GB	
828.0GB (normal)					

```
data 1.31.4
                                          BSAS
                                                 7200 827.7GB
828.0GB (normal)
    data 1.31.5
                                        BSAS
                                                 7200 827.7GB
                                       1
828.0GB (normal)
Aggregate: node A 2 data02 unmirrored (online, raid dp) (block
checksums)
 Plex: /node A 2 data02 unmirrored/plex0 (online, normal, active,
pool0)
  RAID Group /node A 2 data02 unmirrored/plex0/rg0 (normal, block
checksums)
                                                        Usable
Physical
    Position Disk
                                     Pool Type RPM
                                                         Size
Size Status
    ______
_____
    dparity 2.30.12
                                       0
                                          BSAS
                                                 7200 827.7GB
828.0GB (normal)
    parity 2.30.22
                                       0
                                          BSAS
                                                 7200 827.7GB
828.0GB (normal)
    data
           2.30.21
                                          BSAS
                                                 7200 827.7GB
                                       0
828.0GB (normal)
    data 2.30.20
                                       0
                                          BSAS
                                                 7200 827.7GB
828.0GB (normal)
            2.30.14
    data
                                       0
                                                 7200 827.7GB
                                          BSAS
828.0GB (normal)
15 entries were displayed.
```

- 5. Physically power off the shelf that you selected.
- 6. Check the aggregate status again:

```
storage aggregate show
storage aggregate show -r -node node A 2 !*root
```

The aggregate with drives on the powered-off shelf should have a "degraded" RAID status, and drives on the affected plex should have a "failed" status, as shown in the following example:

```
raid_dp,
mirrored,
normal
node A 1root
        707.7GB 34.29GB 95% online 1 node A 1
raid dp,
mirrored,
normal
node A 2 data01 mirrored
         4.15TB 4.12TB 1% online 2 node_A_2
raid_dp,
mirror
degraded
node A 2 data02 unmirrored
         2.18TB 2.18TB 0% online 1 node_A_2
raid_dp,
normal
node A 2 root
        707.7GB 34.27GB 95% online 1 node_A_2
raid dp,
mirror
degraded
cluster A::> storage aggregate show -r -node node A 2 !*root
Owner Node: node A 2
Aggregate: node A 2 data01 mirrored (online, raid_dp, mirror degraded)
(block checksums)
 Plex: /node A 2 data01 mirrored/plex0 (online, normal, active, pool0)
  RAID Group /node A 2 data01 mirrored/plex0/rg0 (normal, block
checksums)
                                                     Usable
Physical
                              Pool Type RPM Size
   Position Disk
Size Status
    ______
-----
                                     0 BSAS 7200 827.7GB
   dparity 2.30.3
828.0GB (normal)
```

parity		0	BSAS	7200	827.7GB
828.0GB (norma		0	Daza	7000	007 700
data		0	BSAS	1200	827.7GB
828.0GB (norma		0	DCAC	7200	827.7GB
828.0GB (norma		U	BSAS	7200	827.7GB
data		0	BSAS	7200	827.7GB
828.0GB (norma		U	DOAO	7200	027.7GB
OZO.OCD (HOTHIC	44,				
Plex: /node	A 2 data01 mirrored/plex4 (c	offli	ne, faile	ed, ina	active,
pool1)			·	·	·
-	/node A 2 data01 mirrored/pi	Lex4/	rg0 (par	tial, ı	none
checksums)					
					Usable
Physical					
Position	Disk	Pool	Type	RPM	Size
Size Status					
dparity	FAILED	-	_	_	827.7GB
- (failed)					005 505
parity	FAILED	-	-	_	827.7GB
- (failed)					007 700
	FAILED	_	_	_	827.7GB
- (failed) data	FAILED	_	_	_	827.7GB
- (failed)	FAILED	_	_	_	027.7GD
	FAILED	_	_	_	827.7GB
- (failed)					027 <b>.</b> 70D
(141164)					
Aggregate: no	ode A 2 data02 unmirrored (or	nline	, raid d	o) (blo	ock
checksums)					
Plex: /node_	A_2_data02_unmirrored/plex0	(onl:	ine, norm	mal, a	ctive,
pool0)					
RAID Group	/node_A_2_data02_unmirrored,	/plex	0/rg0 (no	ormal,	block
checksums)					
					Usable
Physical					
Position	Disk	Pool	Type	RPM	Size
Size Status					
dparity		0	BSAS	7200	827.7GB
828.0GB (norma		0	Dara	7000	007 700
parity		0	BSAS	7200	827.7GB
828.0GB (norma	A _ )				

data	2.30.21	0	BSAS	7200	827.7GB
828.0GB (nor	mal)				
data	2.30.20	0	BSAS	7200	827.7GB
828.0GB (nor	mal)				
data	2.30.14	0	BSAS	7200	827.7GB

828.0GB (normal)

15 entries were displayed.

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

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

cluster A::> vserver show -type data cluster A::> vserver show -type data Admin Operational Root Vserver Type Subtype State State Volume Aggregate SVM1 data sync-source running SVM1\_root node A 1 data01 mirrored SVM2 data sync-source running SVM2 root node A 1 data01 mirrored cluster A::> network interface show -fields is-home false There are no entries matching your query. cluster\_A::> volume show !vol0,!MDV\* Vserver Volume Aggregate State Type Size Available Used% \_\_\_\_\_\_ \_\_\_\_ \_\_\_\_ SVM1 SVM1 root node A 1data01 mirrored online RW 10GB 9.50GB 5% SVM1 SVM1 data\_vol node A 1data01 mirrored online RW 10GB 9.49GB 5% SVM2 SVM2 root node A 1data01 mirrored online RW 10GB 9.49GB 5% SVM2 SVM2 data vol node A 2 data02 unmirrored online RW 1GB 972.6MB 5%

### 8. Physically power on the shelf.

Resynchronization starts automatically.

9. Verify that resynchronization has started:

```
storage aggregate show
```

The affected aggregate should have a "resyncing" RAID status, as shown in the following example:

```
cluster A::> storage aggregate show
cluster Aggregates:
Aggregate Size Available Used% State #Vols Nodes RAID
Status
__________
_____
node A 1 data01 mirrored
       4.15TB 3.40TB 18% online 3 node A_1
raid dp,
mirrored,
normal
node A 1 root
       707.7GB 34.29GB 95% online 1 node A 1
raid dp,
mirrored,
normal
node A 2 data01 mirrored
       4.15TB 4.12TB 1% online 2 node_A_2
raid dp,
resyncing
node A 2 data02 unmirrored
       raid dp,
normal
node_A_2_root
       707.7GB 34.27GB 95% online 1 node A 2
raid dp,
resyncing
```

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

```
storage aggregate show
```

The affected aggregate should have a "normal" RAID status, as shown in the following example:

```
cluster A::> storage aggregate show
cluster Aggregates:
Aggregate Size Available Used% State #Vols Nodes RAID
_____ ____
_____
node A 1data01 mirrored
     4.15TB 3.40TB 18% online 3 node_A_1
raid_dp,
mirrored,
normal
node A 1root
     707.7GB 34.29GB 95% online 1 node_A_1
raid dp,
mirrored,
normal
node A 2 data01 mirrored
      4.15TB 4.12TB 1% online 2 node_A_2
raid dp,
normal
node A 2 data02 unmirrored
       raid_dp,
normal
node A 2 root
       707.7GB 34.27GB 95% online 1 node_A_2
raid_dp,
resyncing
```

### **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 <a href="http://www.netapp.com/TM">http://www.netapp.com/TM</a> are trademarks of NetApp, Inc. Other company and product names may be trademarks of their respective owners.