



Prepare for switchback in a MetroCluster IP configuration

ONTAP MetroCluster

NetApp
November 15, 2022

This PDF was generated from https://docs.netapp.com/us-en/ontap-metrocluster/disaster-recovery/task_prepare_for_switchback_in_a_mcc_ip_configuration_supertask.html on November 15, 2022. Always check docs.netapp.com for the latest.

Table of Contents

- Prepare for switchback in a MetroCluster IP configuration 1
 - Prepare for switchback in a MetroCluster IP configuration 1
 - Setting required environmental variables in MetroCluster IP configurations 1
 - Powering on the equipment at the disaster site (MetroCluster IP configurations) 10
 - Configuring the IP switches (MetroCluster IP configurations) 10
 - Verify storage connectivity to the remote site (MetroCluster IP configurations) 13
 - Reassigning disk ownership for pool 1 disks on the disaster site (MetroCluster IP configurations) 14
 - Booting to ONTAP on replacement controller modules in MetroCluster IP configurations 18
 - Restoring connectivity from the surviving nodes to the disaster site (MetroCluster IP configurations) 23
 - Verifying automatic assignment or manually assigning pool 0 drives 23
 - Assigning pool 1 drives on the surviving site (MetroCluster IP configurations) 25
 - Deleting failed plexes owned by the surviving site (MetroCluster IP configurations) 26
 - Performing aggregate healing and restoring mirrors (MetroCluster IP configurations) 34

Prepare for switchback in a MetroCluster IP configuration

Prepare for switchback in a MetroCluster IP configuration

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

About this task



Setting required environmental variables in MetroCluster IP configurations

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

About this task

This task is required only in MetroCluster IP configurations.

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

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



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

Node	Port	IP address
------	------	------------

node_A_1	e5a	172.17.26.10
	e5b	172.17.27.10
node_A_2	e5a	172.17.26.11
	e5b	172.17.27.11
node_B_1	e5a	172.17.26.13
	e5b	172.17.27.13
node_B_2	e5a	172.17.26.12
	e5b	172.17.27.12

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

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

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

Platform models that use VLAN IDs with the MetroCluster IP interfaces
--

<ul style="list-style-type: none"> • AFF A220 • AFF A250 • AFF A400 	<ul style="list-style-type: none"> • FAS500f • FAS2750 • FAS8300 • FAS8700
--	--

Steps

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

```
metrocluster configuration-settings connection show
```

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

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

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

```

completed          172.17.26.12      172.17.26.13      HA Partner
                    Home Port: e5a
                    172.17.26.12      172.17.26.11      DR Partner
completed
                    Home Port: e5a
                    172.17.26.12      172.17.26.10      DR Auxiliary
completed
                    Home Port: e5b
                    172.17.27.12      172.17.27.13      HA Partner
completed
                    Home Port: e5b
                    172.17.27.12      172.17.27.11      DR Partner
completed
                    Home Port: e5b
                    172.17.27.12      172.17.27.10      DR Auxiliary
completed
12 entries were displayed.

```

2. If you need to determine the VLAN ID or gateway address for the interface, determine the VLAN IDs from the surviving site:

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

- You need the VLAN ID if the platform models use VLAN IDs (see the list above), and if you are not using the default VLAN IDs.
- You need the gateway address if you are using [Layer 3 wide-area networks](#).

The VLAN IDs are included in the **Network Address** column of the output. The **Gateway** column shows the gateway IP address.

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

```

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

```

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

```

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

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

```



- If the interfaces are using the default VLANs, or the platform model does not require a VLAN (see the list above), the *vlan-id* is not necessary.
- If the configuration is not using [Layer3 wide-area networks](#), the value for *gateway-IP-address* is **0** (zero).
- If the interfaces are using the default VLANs, or the platform model does not require a VLAN (see the list above), the *vlan-id* is not necessary.
- If the configuration is not using [layer 3 backend connections](#), the value for *gateway-IP-address* is **0** (zero).

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

```
setenv bootarg.mcc.port_a_ip_config
172.17.26.10/23,0,172.17.26.11,172.17.26.13,172.17.26.12,120

setenv bootarg.mcc.port_b_ip_config
172.17.27.10/23,0,172.17.27.11,172.17.27.13,172.17.27.12,130
```

The following example shows the commands for node_A_1 without a VLAN ID:

```
setenv bootarg.mcc.port_a_ip_config
172.17.26.10/23,0,172.17.26.11,172.17.26.13,172.17.26.12

setenv bootarg.mcc.port_b_ip_config
172.17.27.10/23,0,172.17.27.11,172.17.27.13,172.17.27.12
```

4. If the disaster site nodes are not systems that use VLAN IDs, at the LOADER prompt for each of the disaster nodes, set the following bootargs with local_IP/mask,gateway:

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

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



- If the interfaces are using the default VLANs, or the platform model does not require a VLAN (see the list above), the *vlan-id* is not necessary.
- If the configuration is not using [Layer 3 wide-area networks](#), the value for *gateway-IP-address* is **0** (zero).

The following commands set the values for node_A_1. In this example, the *gateway-IP-address* and *vlan-id* values are not used.

```
setenv bootarg.mcc.port_a_ip_config
172.17.26.10/23,0,172.17.26.11,172.17.26.13,172.17.26.12

setenv bootarg.mcc.port_b_ip_config
172.17.27.10/23,0,172.17.27.11,172.17.27.13,172.17.27.12
```

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

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



```

cluster_B::> metrocluster node show -fields node-cluster-uuid, node-uuid

(metrocluster node show)
dr-group-id cluster      node      node-uuid
node-cluster-uuid
-----
1          cluster_A    node_A_1 f03cb63c-9a7e-11e7-b68b-00a098908039
ee7db9d5-9a82-11e7-b68b-00a098

908039
1          cluster_A    node_A_2 aa9a7a7a-9a81-11e7-a4e9-00a098908c35
ee7db9d5-9a82-11e7-b68b-00a098

908039
1          cluster_B    node_B_1 f37b240b-9ac1-11e7-9b42-00a098c9e55d
07958819-9ac6-11e7-9b42-00a098

c9e55d
1          cluster_B    node_B_2 bf8e3f8f-9ac4-11e7-bd4e-00a098ca379f
07958819-9ac6-11e7-9b42-00a098

c9e55d
4 entries were displayed.
cluster_A::~*>

```

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

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

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

setenv bootarg.mgwd.cluster_uuid local-cluster-UUID

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

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

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

a. Set the UUIDs on node_A_1.

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

```
setenv bootarg.mgwd.cluster_uuid ee7db9d5-9a82-11e7-b68b-00a098908039

setenv bootarg.mgwd.partner_cluster_uuid 07958819-9ac6-11e7-9b42-
00a098c9e55d

setenv bootarg.mcc.pri_partner_uuid f37b240b-9ac1-11e7-9b42-
00a098c9e55d

setenv bootarg.mcc.aux_partner_uuid bf8e3f8f-9ac4-11e7-bd4e-
00a098ca379f

setenv bootarg.mcc_iscsi.node_uuid f03cb63c-9a7e-11e7-b68b-
00a098908039
```

b. Set the UUIDs on node_A_2:

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

```
setenv bootarg.mgwd.cluster_uuid ee7db9d5-9a82-11e7-b68b-00a098908039

setenv bootarg.mgwd.partner_cluster_uuid 07958819-9ac6-11e7-9b42-
00a098c9e55d

setenv bootarg.mcc.pri_partner_uuid bf8e3f8f-9ac4-11e7-bd4e-
00a098ca379f

setenv bootarg.mcc.aux_partner_uuid f37b240b-9ac1-11e7-9b42-
00a098c9e55d

setenv bootarg.mcc.iscsi.node_uuid aa9a7a7a-9a81-11e7-a4e9-
00a098908c35
```

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

```
setenv bootarg.mcc.adp_enabled true
```

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

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

- a. Set the variables on node_A_1.

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

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

- b. Set the variables on node_A_2.

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

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

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

```
setenv bootarg.mcc.local_config_id original-sysID
```

```
setenv bootarg.mcc.dr_partner dr_partner-sysID
```

[Determining the system IDs and VLAN IDs of the old controller modules](#)

- a. Set the variables on node_A_1.

The following example shows the commands for setting the system IDs on node_A_1:

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

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

- b. Set the variables on node_A_2.

The following example shows the commands for setting the system IDs on node_A_2:

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

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

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

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

About this task

The examples in this procedure assume the following:

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

Steps

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

Configuring the IP switches (MetroCluster IP configurations)

You must configure any IP switches that were replaced.

About this task

This task applies to MetroCluster IP configurations only.

This must be done on both switches. Verify after configuring the first switch that storage access on the

surviving site is not impacted.



You must not proceed with the second switch if storage access on the surviving site is impacted.

Steps

- 1. Refer to [MetroCluster IP installation and configuration: : Differences among the ONTAP MetroCluster configurations](#) for procedures for cabling and configuring a replacement switch.

You can use the procedures in the following sections:

- Cabling the IP switches
 - Configuring the IP switches
- 2. If the ISLs were disabled at the surviving site, enable the ISLs and verify that the ISLs are online.
 - a. Enable the ISL interfaces on the first switch:

```
no shutdown
```

The following examples show the commands for a Broadcom IP switch or a Cisco IP switch.

Switch vendor	Commands
Broadcom	<pre>(IP_Switch_A_1)> enable (IP_switch_A_1)# configure (IP_switch_A_1)(Config)# interface 0/13-0/16 (IP_switch_A_1)(Interface 0/13-0/16)# no shutdown (IP_switch_A_1)(Interface 0/13-0/16)# exit (IP_switch_A_1)(Config)# exit</pre>
Cisco	<pre>IP_switch_A_1# conf t IP_switch_A_1(config)# int eth1/15-eth1/20 IP_switch_A_1(config)# no shutdown IP_switch_A_1(config)# copy running startup IP_switch_A_1(config)# show interface brief</pre>

- b. Enable the ISL interfaces on the partner switch:

```
no shutdown
```

The following examples show the commands for a Broadcom IP switch or a Cisco IP switch.

Switch vendor	Commands
Broadcom	<pre>(IP_Switch_A_2)> enable (IP_switch_A_2)# configure (IP_switch_A_2) (Config)# interface 0/13-0/16 (IP_switch_A_2) (Interface 0/13- 0/16)# no shutdown (IP_switch_A_2) (Interface 0/13- 0/16)# exit (IP_switch_A_2) (Config)# exit</pre>
Cisco	<pre>IP_switch_A_2# conf t IP_switch_A_2(config)# int eth1/15-eth1/20 IP_switch_A_2(config)# no shutdown IP_switch_A_2(config)# copy running startup IP_switch_A_2(config)# show interface brief</pre>

c. Verify that the interfaces are enabled:

```
show interface brief
```

The following example shows the output for a Cisco switch.

```

IP_switch_A_2(config)# show interface brief

-----
Port VRF Status IP Address Speed MTU
-----
mt0 -- up 10.10.99.10 100 1500
-----

Ethernet      VLAN Type Mode      Status Reason Speed   Port
Interface                                           Ch
#
-----
.
.
.
Eth1/15      10  eth  access  up       none   40G(D)  --
Eth1/16      10  eth  access  up       none   40G(D)  --
Eth1/17      10  eth  access  down     none   auto(D) --
Eth1/18      10  eth  access  down     none   auto(D) --
Eth1/19      10  eth  access  down     none   auto(D) --
Eth1/20      10  eth  access  down     none   auto(D) --
.
.
.
IP_switch_A_2#

```

Verify storage connectivity to the remote site (MetroCluster IP configurations)

You must confirm that the replaced nodes have connectivity to the disk shelves at the surviving site.

About this task

This task is performed on the replacement nodes at the disaster site.

This task is performed in Maintenance mode.

Steps

1. Display the disks that are owned by the original system ID.

```
disk show -s old-system-ID
```

The remote disks can be recognized by the 0m device. 0m indicates that the disk is connected via the MetroCluster iSCSI connection. These disks must be reassigned later in the recovery procedure.

```
*> disk show -s 4068741256
Local System ID: 1574774970
```

DISK	OWNER	POOL	SERIAL NUMBER	HOME
DR HOME				
-----	-----	-----	-----	
-----	-----	-----	-----	
0m.i0.0L11	node_A_2 (4068741256)	Pool1	S396NA0HA02128	node_A_2
(4068741256)	node_A_2 (4068741256)			
0m.i0.1L38	node_A_2 (4068741256)	Pool1	S396NA0J148778	node_A_2
(4068741256)	node_A_2 (4068741256)			
0m.i0.0L52	node_A_2 (4068741256)	Pool1	S396NA0J148777	node_A_2
(4068741256)	node_A_2 (4068741256)			
...				
...				

NOTE: Currently 49 disks are unowned. Use 'disk show -n' for additional information.

```
*>
```

2. Repeat this step on the other replacement nodes

Reassigning disk ownership for pool 1 disks on the disaster site (MetroCluster IP configurations)

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

About this task

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

This task is performed on the replacement nodes at the disaster site.

This task is performed in Maintenance mode.

The examples make the following assumptions:

- Site A is the disaster site.
- node_A_1 has been replaced.
- node_A_2 has been replaced.
- Site B is the surviving site.
- node_B_1 is healthy.
- node_B_2 is healthy.

The old and new system IDs were identified in [Determining the new System IDs of the replacement controller modules](#).

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

Node	Original system ID	New system ID
node_A_1	4068741258	1574774970
node_A_2	4068741260	1574774991
node_B_1	4068741254	unchanged
node_B_2	4068741256	unchanged

Steps

1. With the replacement node in Maintenance mode, reassign the root aggregate disks, using the correct command, depending on whether your system is configured with ADP and your ONTAP version.

You can proceed with the reassignment when prompted.

If the system is using ADP...	Use this command for disk reassignment...
Yes (ONTAP 9.8)	<code>disk reassign -s old-system-ID -d new-system-ID -r dr-partner-system-ID</code>
Yes (ONTAP 9.7.x and earlier)	<code>disk reassign -s old-system-ID -d new-system-ID -p old-partner-system-ID</code>
No	<code>disk reassign -s old-system-ID -d new-system-ID</code>

The following example shows reassignment of drives on a non-ADP system:

```
*> disk reassign -s 4068741256 -d 1574774970
Partner node must not be in Takeover mode during disk reassignment from
maintenance mode.
Serious problems could result!!
Do not proceed with reassignment if the partner is in takeover mode.
Abort reassignment (y/n)? n

After the node becomes operational, you must perform a takeover and
giveback of the HA partner node to ensure disk reassignment is
successful.
Do you want to continue (y/n)? y
Disk ownership will be updated on all disks previously belonging to
Filer with sysid 537037643.
Do you want to continue (y/n)? y
disk reassign parameters: new_home_owner_id 537070473 ,
new_home_owner_name
Disk 0m.i0.3L14 will be reassigned.
Disk 0m.i0.1L6 will be reassigned.
Disk 0m.i0.1L8 will be reassigned.
Number of disks to be reassigned: 3
```

2. Destroy the contents of the mailbox disks:

```
mailbox destroy local
```

You can proceed with the destroy operation when prompted.

The following example shows the output for the mailbox destroy local command:

```
*> mailbox destroy local
Destroying mailboxes forces a node to create new empty mailboxes,
which clears any takeover state, removes all knowledge
of out-of-date plexes of mirrored volumes, and will prevent
management services from going online in 2-node cluster
HA configurations.
Are you sure you want to destroy the local mailboxes? y
.....Mailboxes destroyed.
*>
```

3. If disks have been replaced, there will be failed local plexes that must be deleted.

a. Display the aggregate status:

```
aggr status
```

In the following example, plex node_A_1_aggr0/plex0 has failed.

```

*> aggr status
Aug 18 15:00:07 [node_B_1:raid.vol.mirror.degraded:ALERT]: Aggregate
node_A_1_aggr0 is
    mirrored and one plex has failed. It is no longer protected by
    mirroring.
Aug 18 15:00:07 [node_B_1:raid.debug:info]: Mirrored aggregate
node_A_1_aggr0 has plex0
    clean(-1), online(0)
Aug 18 15:00:07 [node_B_1:raid.debug:info]: Mirrored aggregate
node_A_1_aggr0 has plex2
    clean(0), online(1)
Aug 18 15:00:07 [node_B_1:raid.mirror.vote.noRecord1Plex:error]:
WARNING: Only one plex
    in aggregate node_A_1_aggr0 is available. Aggregate might contain
    stale data.
Aug 18 15:00:07 [node_B_1:raid.debug:info]:
volobj_mark_sb_recovery_aggrs: tree:
    node_A_1_aggr0 vol_state:1 mcc_dr_opstate: unknown
Aug 18 15:00:07 [node_B_1:raid.fsm.commitStateTransit:debug]:
/node_A_1_aggr0 (VOL):
    raid state change UNINITD -> NORMAL
Aug 18 15:00:07 [node_B_1:raid.fsm.commitStateTransit:debug]:
/node_A_1_aggr0 (MIRROR):
    raid state change UNINITD -> DEGRADED
Aug 18 15:00:07 [node_B_1:raid.fsm.commitStateTransit:debug]:
/node_A_1_aggr0/plex0
    (PLEX): raid state change UNINITD -> FAILED
Aug 18 15:00:07 [node_B_1:raid.fsm.commitStateTransit:debug]:
/node_A_1_aggr0/plex2
    (PLEX): raid state change UNINITD -> NORMAL
Aug 18 15:00:07 [node_B_1:raid.fsm.commitStateTransit:debug]:
/node_A_1_aggr0/plex2/rg0
    (GROUP): raid state change UNINITD -> NORMAL
Aug 18 15:00:07 [node_B_1:raid.debug:info]: Topology updated for
aggregate node_A_1_aggr0
    to plex plex2
*>

```

b. Delete the failed plex:

```
aggr destroy plex-id
```

```
*> aggr destroy node_A_1_aggr0/plex0
```

4. Halt the node to display the LOADER prompt:

```
halt
```

5. Repeat these steps on the other node at the disaster site.

Booting to ONTAP on replacement controller modules in MetroCluster IP configurations

You must boot the replacement nodes at the disaster site to the ONTAP operating system.

About this task

This task begins with the nodes at the disaster site in Maintenance mode.

Steps

1. On one of the replacement nodes, exit to the LOADER prompt: `halt`
2. Display the boot menu: `boot_ontap menu`
3. From the boot menu, select option 6, **Update flash from backup config**.

The system boots twice. You should respond `yes` when prompted to continue. After the second boot, you should respond `y` when prompted about the system ID mismatch.



If you did not clear the NVRAM contents of a used replacement controller module, then you might see the following panic message: `PANIC: NVRAM contents are invalid...`. If this occurs, boot the system to the ONTAP prompt again (`boot_ontap menu`). You then need to [Reset the boot_recovery and rdb_corrupt bootargs](#)

- Confirmation to continue prompt:

```
Selection (1-9)? 6
```

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

- System ID mismatch prompt:

```
WARNING: System ID mismatch. This usually occurs when replacing a  
boot device or NVRAM cards!  
Override system ID? {y|n} y
```

4. From the surviving site, verify that the correct partner system IDs have been applied to the nodes:

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

In this example, the following new system IDs should appear in the output:

- Node_A_1: 1574774970
- Node_A_2: 1574774991

The "ha-partner-systemid" column should show the new system IDs.

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

dr-group-id cluster      node      node-systemid ha-partner-systemid dr-
partner-systemid dr-auxiliary-systemid
-----
1            Cluster_A   Node_A_1   1574774970    1574774991
4068741254           4068741256
1            Cluster_A   Node_A_2   1574774991    1574774970
4068741256           4068741254
1            Cluster_B   Node_B_1   -              -
-
1            Cluster_B   Node_B_2   -              -
-
4 entries were displayed.
```

5. If the partner system IDs were not correctly set, you must manually set the correct value:

- Halt and display the LOADER prompt on the node.
- Verify the partner-sysID bootarg's current value:

```
printenv
```

- Set the value to the correct partner system ID:

```
setenv partner-sysid partner-sysID
```

- Boot the node:

```
boot_ontap
```

- Repeat these substeps on the other node, if necessary.

6. Confirm that the replacement nodes at the disaster site are ready for switchback:

```
metrocluster node show
```

The replacement nodes should be in waiting for switchback recovery mode. If they are in normal mode instead, you can reboot the replacement nodes. After that boot, the nodes should be in waiting for switchback recovery mode.

The following example shows that the replacement nodes are ready for switchback:

```

cluster_B::> metrocluster node show
DR
Group Cluster Node          Configuration State      DR
Mirroring Mode
-----
1      cluster_B
      node_B_1      configured  enabled  switchover
completed
      node_B_2      configured  enabled  switchover
completed
      cluster_A
      node_A_1      configured  enabled  waiting for
switchback recovery
      node_A_2      configured  enabled  waiting for
switchback recovery
4 entries were displayed.

cluster_B::>

```

7. Verify the MetroCluster connection configuration settings:

```
metrocluster configuration-settings connection show
```

The configuration state should indicate completed.

```

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

```

	Home Port: e5b		
completed	172.17.27.13	172.17.27.10	DR Partner
	Home Port: e5b		
completed	172.17.27.13	172.17.27.11	DR Auxiliary
	node_B_1		
	Home Port: e5a		
completed	172.17.26.12	172.17.26.13	HA Partner
	Home Port: e5a		
completed	172.17.26.12	172.17.26.11	DR Partner
	Home Port: e5a		
completed	172.17.26.12	172.17.26.10	DR Auxiliary
	Home Port: e5b		
completed	172.17.27.12	172.17.27.13	HA Partner
	Home Port: e5b		
completed	172.17.27.12	172.17.27.11	DR Partner
	Home Port: e5b		
completed	172.17.27.12	172.17.27.10	DR Auxiliary
	cluster_A		
	node_A_2		
	Home Port: e5a		
completed	172.17.26.11	172.17.26.10	HA Partner
	Home Port: e5a		
completed	172.17.26.11	172.17.26.12	DR Partner
	Home Port: e5a		
completed	172.17.26.11	172.17.26.13	DR Auxiliary
	Home Port: e5b		
completed	172.17.27.11	172.17.27.10	HA Partner
	Home Port: e5b		
completed	172.17.27.11	172.17.27.12	DR Partner
	Home Port: e5b		
completed	172.17.27.11	172.17.27.13	DR Auxiliary
	node_A_1		

```

Home Port: e5a
172.17.26.10 172.17.26.11 HA Partner
completed
Home Port: e5a
172.17.26.10 172.17.26.13 DR Partner
completed
Home Port: e5a
172.17.26.10 172.17.26.12 DR Auxiliary
completed
Home Port: e5b
172.17.27.10 172.17.27.11 HA Partner
completed
Home Port: e5b
172.17.27.10 172.17.27.13 DR Partner
completed
Home Port: e5b
172.17.27.10 172.17.27.12 DR Auxiliary
completed
24 entries were displayed.

cluster_B::*>

```

8. Repeat the previous steps on the other node at the disaster site.

Reset the boot_recovery and rdb_corrupt bootargs

If required, you can reset the boot_recovery and rdb_corrupt_bootargs

Steps

1. Halt the node back to the LOADER prompt:

```
node_A_1::*> halt -node _node-name_
```

2. Check if the following bootargs have been set:

```

LOADER> printenv bootarg.init.boot_recovery
LOADER> printenv bootarg.rdb_corrupt

```

3. If either bootarg has been set to a value, unset it and boot ONTAP:


```
LOADER> unsetenv bootarg.init.boot_recovery
LOADER> unsetenv bootarg.rdb_corrupt
LOADER> saveenv
LOADER> bye
```

Restoring connectivity from the surviving nodes to the disaster site (MetroCluster IP configurations)

You must restore the MetroCluster iSCSI initiator connections from the surviving nodes.

About this task

This procedure is only required on MetroCluster IP configurations.

Steps

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

```
set -privilege advanced
```

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

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

```
storage iscsi-initiator connect -node surviving-node -label *
```

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

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

3. Return to the admin privilege level:

```
set -privilege admin
```

Verifying automatic assignment or manually assigning pool 0 drives

On systems configured for ADP, you must verify that pool 0 drives have been automatically assigned. On systems that are not configured for ADP, you must manually assign the pool 0 drives.

Verifying drive assignment of pool 0 drives on ADP systems at the disaster site (MetroCluster IP systems)

If drives have been replaced at the disaster site and the system is configured for ADP, you must verify that the

remote drives are visible to the nodes and have been assigned correctly.

Step

1. Verify that pool 0 drives are assigned automatically:

```
disk show
```

In the following example for an AFF A800 system with no external shelves, one quarter (8 drives) were automatically assigned to node_A_1 and one quarter were automatically assigned to node_A_2. The remaining drives will be remote (pool1) drives for node_B_1 and node_B_2.

```
cluster_A::*> disk show
```

Disk Owner	Usable Size	Disk Shelf	Bay	Container Type	Type	Container Name
node_A_1:0n.12	1.75TB	0	12	SSD-NVM	shared	aggr0
node_A_1:0n.13	1.75TB	0	13	SSD-NVM	shared	aggr0
node_A_1:0n.14	1.75TB	0	14	SSD-NVM	shared	aggr0
node_A_1:0n.15	1.75TB	0	15	SSD-NVM	shared	aggr0
node_A_1:0n.16	1.75TB	0	16	SSD-NVM	shared	aggr0
node_A_1:0n.17	1.75TB	0	17	SSD-NVM	shared	aggr0
node_A_1:0n.18	1.75TB	0	18	SSD-NVM	shared	aggr0
node_A_1:0n.19	1.75TB	0	19	SSD-NVM	shared	-
node_A_2:0n.0	1.75TB	0	0	SSD-NVM	shared	
aggr0_node_A_2_0	node_A_2					
node_A_2:0n.1	1.75TB	0	1	SSD-NVM	shared	
aggr0_node_A_2_0	node_A_2					
node_A_2:0n.2	1.75TB	0	2	SSD-NVM	shared	
aggr0_node_A_2_0	node_A_2					
node_A_2:0n.3	1.75TB	0	3	SSD-NVM	shared	
aggr0_node_A_2_0	node_A_2					
node_A_2:0n.4	1.75TB	0	4	SSD-NVM	shared	
aggr0_node_A_2_0	node_A_2					
node_A_2:0n.5	1.75TB	0	5	SSD-NVM	shared	
aggr0_node_A_2_0	node_A_2					
node_A_2:0n.6	1.75TB	0	6	SSD-NVM	shared	
aggr0_node_A_2_0	node_A_2					

```
node_A_2:0n.7      1.75TB      0      7      SSD-NVM shared      -
node_A_2
node_A_2:0n.24     -            0     24     SSD-NVM unassigned  -      -
node_A_2:0n.25     -            0     25     SSD-NVM unassigned  -      -
node_A_2:0n.26     -            0     26     SSD-NVM unassigned  -      -
node_A_2:0n.27     -            0     27     SSD-NVM unassigned  -      -
node_A_2:0n.28     -            0     28     SSD-NVM unassigned  -      -
node_A_2:0n.29     -            0     29     SSD-NVM unassigned  -      -
node_A_2:0n.30     -            0     30     SSD-NVM unassigned  -      -
node_A_2:0n.31     -            0     31     SSD-NVM unassigned  -      -
node_A_2:0n.36     -            0     36     SSD-NVM unassigned  -      -
node_A_2:0n.37     -            0     37     SSD-NVM unassigned  -      -
node_A_2:0n.38     -            0     38     SSD-NVM unassigned  -      -
node_A_2:0n.39     -            0     39     SSD-NVM unassigned  -      -
node_A_2:0n.40     -            0     40     SSD-NVM unassigned  -      -
node_A_2:0n.41     -            0     41     SSD-NVM unassigned  -      -
node_A_2:0n.42     -            0     42     SSD-NVM unassigned  -      -
node_A_2:0n.43     -            0     43     SSD-NVM unassigned  -      -
32 entries were displayed.
```

Assigning pool 0 drives on non-ADP systems at the disaster site (MetroCluster IP configurations)

If drives have been replaced at the disaster site and the system is not configured for ADP, you need to manually assign new drives to pool 0.

About this task

For ADP systems, the drives are assigned automatically.

Steps

1. On one of the replacement nodes at the disaster site, reassign the node's pool 0 drives:

```
storage disk assign -n number-of-replacement disks -p 0
```

This command assigns the newly added (and unowned) drives on the disaster site. You should assign the same number and size (or larger) of drives that the node had prior to the disaster. The `storage disk assign` man page contains more information about performing more granular drive assignment.

2. Repeat the step on the other replacement node at the disaster site.

Assigning pool 1 drives on the surviving site (MetroCluster IP configurations)

If drives have been replaced at the disaster site and the system is not configured for ADP, at the surviving site you need to manually assign remote drives located at the disaster site to the surviving nodes' pool 1. You must identify the number of drives to assign.

About this task

For ADP systems, the drives are assigned automatically.

Step

1. On the surviving site, assign the first node's pool 1 (remote) drives: `storage disk assign -n number-of-replacement disks -p 1 0m*`

This command assigns the newly added and unowned drives on the disaster site.

The following command assigns 22 drives:

```
cluster_B::> storage disk assign -n 22 -p 1 0m*
```

Deleting failed plexes owned by the surviving site (MetroCluster IP configurations)

After replacing hardware and assigning disks, you must delete failed remote plexes that are owned by the surviving site nodes but located at the disaster site.

About this task

These steps are performed on the surviving cluster.

Steps

1. Identify the local aggregates: `storage aggregate show -is-home true`

```
cluster_B::> storage aggregate show -is-home true

cluster_B Aggregates:
Aggregate      Size Available Used% State   #Vols  Nodes      RAID
Status
-----
node_B_1_aggr0 1.49TB   74.12GB 95% online      1 node_B_1
raid4,

mirror

degraded
node_B_2_aggr0 1.49TB   74.12GB 95% online      1 node_B_2
raid4,

mirror

degraded
node_B_1_aggr1 2.99TB   2.88TB   3% online     15 node_B_1
```

```

raid_dp,

mirror

degraded
node_B_1_aggr2 2.99TB 2.91TB 3% online 14 node_B_1
raid_tec,

mirror

degraded
node_B_2_aggr1 2.95TB 2.80TB 5% online 37 node_B_2
raid_dp,

mirror

degraded
node_B_2_aggr2 2.99TB 2.87TB 4% online 35 node_B_2
raid_tec,

mirror

degraded
6 entries were displayed.

cluster_B::~>

```

2. Identify the failed remote plexes:

```
storage aggregate plex show
```

The following example calls out the plexes that are remote (not plex0) and have a status of "failed":

```
cluster_B::> storage aggregate plex show -fields aggregate,status,is-
online,Plex,pool
aggregate      plex  status      is-online pool
-----
node_B_1_aggr0 plex0 normal,active true      0
node_B_1_aggr0 plex4 failed,inactive false - <<<<---Plex at remote site
node_B_2_aggr0 plex0 normal,active true      0
node_B_2_aggr0 plex4 failed,inactive false - <<<<---Plex at remote site
node_B_1_aggr1 plex0 normal,active true      0
node_B_1_aggr1 plex4 failed,inactive false - <<<<---Plex at remote site
node_B_1_aggr2 plex0 normal,active true      0
node_B_1_aggr2 plex1 failed,inactive false - <<<<---Plex at remote site
node_B_2_aggr1 plex0 normal,active true      0
node_B_2_aggr1 plex4 failed,inactive false - <<<<---Plex at remote site
node_B_2_aggr2 plex0 normal,active true      0
node_B_2_aggr2 plex1 failed,inactive false - <<<<---Plex at remote site
node_A_1_aggr1 plex0 failed,inactive false -
node_A_1_aggr1 plex4 normal,active true      1
node_A_1_aggr2 plex0 failed,inactive false -
node_A_1_aggr2 plex1 normal,active true      1
node_A_2_aggr1 plex0 failed,inactive false -
node_A_2_aggr1 plex4 normal,active true      1
node_A_2_aggr2 plex0 failed,inactive false -
node_A_2_aggr2 plex1 normal,active true      1
20 entries were displayed.

cluster_B::>
```

3. Take offline each of the failed plexes, and then delete them:

a. Take offline the failed plexes:

```
storage aggregate plex offline -aggregate aggregate-name -plex plex-id
```

The following example shows the aggregate "node_B_2_aggr1/plex1" being taken offline:

```
cluster_B::> storage aggregate plex offline -aggregate node_B_1_aggr0
-plex plex4

Plex offline successful on plex: node_B_1_aggr0/plex4
```

b. Delete the failed plex:

```
storage aggregate plex delete -aggregate aggregate-name -plex plex-id
```

You can destroy the plex when prompted.

The following example shows the plex node_B_2_aggr1/plex1 being deleted.

```
cluster_B::> storage aggregate plex delete -aggregate node_B_1_aggr0
-plex plex4

Warning: Aggregate "node_B_1_aggr0" is being used for the local
management root
        volume or HA partner management root volume, or has been
marked as
        the aggregate to be used for the management root volume
after a
        reboot operation. Deleting plex "plex4" for this aggregate
could lead
        to unavailability of the root volume after a disaster
recovery
        procedure. Use the "storage aggregate show -fields
        has-mroot,has-partner-mroot,root" command to view such
aggregates.

Warning: Deleting plex "plex4" of mirrored aggregate "node_B_1_aggr0"
on node
        "node_B_1" in a MetroCluster configuration will disable its
synchronous disaster recovery protection. Are you sure you
want to
        destroy this plex? {y|n}: y
[Job 633] Job succeeded: DONE

cluster_B::>
```

You must repeat these steps for each of the failed plexes.

4. Confirm that the plexes have been removed:

```
storage aggregate plex show -fields aggregate,status,is-online,plex,pool
```

```
cluster_B::> storage aggregate plex show -fields aggregate,status,is-
online,Plex,pool
aggregate      plex  status          is-online pool
-----
node_B_1_aggr0 plex0 normal,active true      0
node_B_2_aggr0 plex0 normal,active true      0
node_B_1_aggr1 plex0 normal,active true      0
node_B_1_aggr2 plex0 normal,active true      0
node_B_2_aggr1 plex0 normal,active true      0
node_B_2_aggr2 plex0 normal,active true      0
node_A_1_aggr1 plex0 failed,inactive false   -
node_A_1_aggr1 plex4 normal,active true      1
node_A_1_aggr2 plex0 failed,inactive false   -
node_A_1_aggr2 plex1 normal,active true      1
node_A_2_aggr1 plex0 failed,inactive false   -
node_A_2_aggr1 plex4 normal,active true      1
node_A_2_aggr2 plex0 failed,inactive false   -
node_A_2_aggr2 plex1 normal,active true      1
14 entries were displayed.

cluster_B::>
```

5. Identify the switched-over aggregates:

```
storage aggregate show -is-home false
```

You can also use the `storage aggregate plex show -fields aggregate,status,is-online,plex,pool` command to identify plex 0 switched-over aggregates. They will have a status of "failed, inactive".

The following commands show four switched-over aggregates:

- `node_A_1_aggr1`
- `node_A_1_aggr2`
- `node_A_2_aggr1`
- `node_A_2_aggr2`


```

cluster_B::> storage aggregate show -is-home false

cluster_A Switched Over Aggregates:
Aggregate      Size Available Used% State   #Vols  Nodes      RAID
Status
-----
node_A_1_aggr1 2.12TB  1.88TB   11% online    91 node_B_1
raid_dp,

mirror

degraded
node_A_1_aggr2 2.89TB  2.64TB    9% online    90 node_B_1
raid_tec,

mirror

degraded
node_A_2_aggr1 2.12TB  1.86TB   12% online    91 node_B_2
raid_dp,

mirror

degraded
node_A_2_aggr2 2.89TB  2.64TB    9% online    90 node_B_2
raid_tec,

mirror

degraded
4 entries were displayed.

cluster_B::>

```

6. Identify switched-over plexes:

```
storage aggregate plex show -fields aggregate,status,is-online,Plex,pool
```

You want to identify the plexes with a status of "failed, inactive".

The following commands show four switched-over aggregates:

```
cluster_B::> storage aggregate plex show -fields aggregate,status,is-
online,Plex,pool
aggregate      plex  status          is-online pool
-----
node_B_1_aggr0 plex0 normal,active true      0
node_B_2_aggr0 plex0 normal,active true      0
node_B_1_aggr1 plex0 normal,active true      0
node_B_1_aggr2 plex0 normal,active true      0
node_B_2_aggr1 plex0 normal,active true      0
node_B_2_aggr2 plex0 normal,active true      0
node_A_1_aggr1 plex0 failed,inactive false - <<<<-- Switched over
aggr/Plex0
node_A_1_aggr1 plex4 normal,active true      1
node_A_1_aggr2 plex0 failed,inactive false - <<<<-- Switched over
aggr/Plex0
node_A_1_aggr2 plex1 normal,active true      1
node_A_2_aggr1 plex0 failed,inactive false - <<<<-- Switched over
aggr/Plex0
node_A_2_aggr1 plex4 normal,active true      1
node_A_2_aggr2 plex0 failed,inactive false - <<<<-- Switched over
aggr/Plex0
node_A_2_aggr2 plex1 normal,active true      1
14 entries were displayed.

cluster_B::>
```

7. Delete the failed plex:

```
storage aggregate plex delete -aggregate node_A_1_aggr1 -plex plex0
```

You can destroy the plex when prompted.

The following example shows the plex node_A_1_aggr1/plex0 being deleted:

```

cluster_B::> storage aggregate plex delete -aggregate node_A_1_aggr1
-plex plex0

Warning: Aggregate "node_A_1_aggr1" hosts MetroCluster metadata volume
"MDV_CRS_e8457659b8a711e78b3b00a0988fe74b_A". Deleting plex
"plex0"
      for this aggregate can lead to the failure of configuration
      replication across the two DR sites. Use the "volume show
-vserver
      <admin-vserver> -volume MDV_CRS*" command to verify the
location of
      such volumes.

Warning: Deleting plex "plex0" of mirrored aggregate "node_A_1_aggr1" on
node
      "node_A_1" in a MetroCluster configuration will disable its
      synchronous disaster recovery protection. Are you sure you want
to
      destroy this plex? {y|n}: y
[Job 639] Job succeeded: DONE

cluster_B::>

```

You must repeat these steps for each of the failed aggregates.

8. Verify that there are no failed plexes remaining on the surviving site.

The following output shows that all plexes are normal, active, and online.

```
cluster_B::> storage aggregate plex show -fields aggregate,status,is-
online,Plex,pool
aggregate      plex  status          is-online pool
-----
node_B_1_aggr0 plex0 normal,active true      0
node_B_2_aggr0 plex0 normal,active true      0
node_B_1_aggr1 plex0 normal,active true      0
node_B_2_aggr2 plex0 normal,active true      0
node_B_1_aggr1 plex0 normal,active true      0
node_B_2_aggr2 plex0 normal,active true      0
node_A_1_aggr1 plex4 normal,active true      1
node_A_1_aggr2 plex1 normal,active true      1
node_A_2_aggr1 plex4 normal,active true      1
node_A_2_aggr2 plex1 normal,active true      1
10 entries were displayed.

cluster_B::>
```

Performing aggregate healing and restoring mirrors (MetroCluster IP configurations)

After replacing hardware and assigning disks, in systems running ONTAP 9.5 or earlier you can perform the MetroCluster healing operations. In all versions of ONTAP, you must then confirm that aggregates are mirrored and, if necessary, restart mirroring.

About this task

Beginning with ONTAP 9.6, the healing operations are performed automatically when the disaster site nodes boot up. The healing commands are not required.

These steps are performed on the surviving cluster.

Steps

1. If you are using ONTAP 9.6 or later, you must verify that automatic healing completed successfully:
 - a. Confirm that the heal-aggr-auto and heal-root-aggr-auto operations completed:

```
metrocluster operation history show
```

The following output shows that the operations have completed successfully on cluster_A.

```
cluster_B::*> metrocluster operation history show
```

Operation Time	State	Start Time	End
-----	-----	-----	
heal-root-aggr-auto	successful	2/25/2019 06:45:58	
2/25/2019 06:46:02			
heal-aggr-auto	successful	2/25/2019 06:45:48	
2/25/2019 06:45:52			
.			
.			
.			

b. Confirm that the disaster site is ready for switchback:

```
metrocluster node show
```

The following output shows that the operations have completed successfully on cluster_A.

```
cluster_B::*> metrocluster node show
```

DR 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
switchback recovery				
		node_B_2	configured	enabled waiting for
switchback recovery				
4 entries were displayed.				

2. If you are using ONTAP 9.5 or earlier, you must perform aggregate healing:

a. Verify the state of the nodes:

```
metrocluster node show
```

The following output shows that switchover has completed, so healing can be performed.

```

cluster_B::> metrocluster node show
DR
Group Cluster Node Configuration State DR Mirroring Mode
-----
1 cluster_B
node_B_1 configured enabled switchover
completed
node_B_2 configured enabled switchover
completed
cluster_A
node_A_1 configured enabled waiting for
switchback recovery
node_A_2 configured enabled waiting for
switchback recovery
4 entries were displayed.

cluster_B::>

```

b. Perform the aggregates healing phase:

```
metrocluster heal -phase aggregates
```

The following output shows a typical aggregates healing operation.

```

cluster_B::*> metrocluster heal -phase aggregates
[Job 647] Job succeeded: Heal Aggregates is successful.

cluster_B::*> metrocluster operation show
Operation: heal-aggregates
State: successful
Start Time: 10/26/2017 12:01:15
End Time: 10/26/2017 12:01:17
Errors: -

cluster_B::*>

```

c. Verify that aggregate healing has completed and the disaster site is ready for switchback:

```
metrocluster node show
```

The following output shows that the "heal aggregates" phase has completed on cluster_A.

```
cluster_B::> metrocluster node show
```

DR Group	Cluster	Node	Configuration State	DR Mirroring Mode
1	cluster_A	node_A_1	configured	enabled heal
	aggregates completed	node_A_2	configured	enabled heal
	aggregates completed	cluster_B		
		node_B_1	configured	enabled waiting for
	switchback recovery	node_B_2	configured	enabled waiting for
	switchback recovery			

4 entries were displayed.

```
cluster_B::>
```

3. If disks have been replaced, you must mirror the local and switched-over aggregates:

a. Display the aggregates:

```
storage aggregate show
```

```
cluster_B::> storage aggregate show
```

cluster_B Aggregates:

Aggregate	Size	Available	Used%	State	#Vols	Nodes
node_B_1_aggr0	1.49TB	74.12GB	95%	online	1	node_B_1
raid4,						
normal						
node_B_2_aggr0	1.49TB	74.12GB	95%	online	1	node_B_2
raid4,						
normal						
node_B_1_aggr1	3.14TB	3.04TB	3%	online	15	node_B_1
raid_dp,						
normal						
node_B_1_aggr2	3.14TB	3.06TB	3%	online	14	node_B_1
raid_tec,						

```

normal
node_B_1_aggr1 3.14TB  2.99TB    5% online    37 node_B_2
raid_dp,

normal
node_B_1_aggr2 3.14TB  3.02TB    4% online    35 node_B_2
raid_tec,

normal

cluster_A Switched Over Aggregates:
Aggregate      Size Available Used% State  #Vols  Nodes
RAID Status
-----
node_A_1_aggr1 2.36TB  2.12TB   10% online    91 node_B_1
raid_dp,

normal
node_A_1_aggr2 3.14TB  2.90TB    8% online    90 node_B_1
raid_tec,

normal
node_A_2_aggr1 2.36TB  2.10TB   11% online    91 node_B_2
raid_dp,

normal
node_A_2_aggr2 3.14TB  2.89TB    8% online    90 node_B_2
raid_tec,

normal
12 entries were displayed.

cluster_B::>

```

b. Mirror the aggregate:

```
storage aggregate mirror -aggregate aggregate-name
```

The following output shows a typical mirroring operation.


```
cluster_B::> storage aggregate mirror -aggregate node_B_1_aggr1
```

Info: Disks would be added to aggregate "node_B_1_aggr1" on node "node_B_1" in the following manner:

Second Plex

	RAID Group rg0, 6 disks (block checksum, raid_dp)		
Size	Position	Disk	Type
	-----	-----	-----
	dparity	5.20.6	SSD
-	parity	5.20.14	SSD
-	data	5.21.1	SSD
894.0GB	data	5.21.3	SSD
894.0GB	data	5.22.3	SSD
894.0GB	data	5.21.13	SSD
894.0GB			

Aggregate capacity available for volume use would be 2.99TB.

Do you want to continue? {y|n}: y

- c. Repeat the previous step for each of the aggregates from the surviving site.
- d. Wait for the aggregates to resynchronize; you can check the status with the `storage aggregate show` command.

The following output shows that a number of aggregates are resynchronizing.

```
cluster_B::> storage aggregate show
```

cluster_B Aggregates:

Aggregate	Size	Available	Used%	State	#Vols	Nodes
RAID Status						
-----	-----	-----	-----	-----	-----	-----
node_B_1_aggr0	1.49TB	74.12GB	95%	online	1	node_B_1
raid4,						

```

mirrored,

normal
node_B_2_aggr0 1.49TB  74.12GB  95% online  1 node_B_2
raid4,

mirrored,

normal
node_B_1_aggr1 2.86TB  2.76TB   4% online  15 node_B_1
raid_dp,

resyncing
node_B_1_aggr2 2.89TB  2.81TB   3% online  14 node_B_1
raid_tec,

resyncing
node_B_2_aggr1 2.73TB  2.58TB   6% online  37 node_B_2
raid_dp,

resyncing
node_B-2_aggr2 2.83TB  2.71TB   4% online  35 node_B_2
raid_tec,

resyncing

cluster_A Switched Over Aggregates:
Aggregate      Size Available Used% State  #Vols  Nodes
RAID Status
-----
node_A_1_aggr1 1.86TB  1.62TB  13% online  91 node_B_1
raid_dp,

resyncing
node_A_1_aggr2 2.58TB  2.33TB  10% online  90 node_B_1
raid_tec,

resyncing
node_A_2_aggr1 1.79TB  1.53TB  14% online  91 node_B_2
raid_dp,

resyncing
node_A_2_aggr2 2.64TB  2.39TB   9% online  90 node_B_2
raid_tec,

```

```
resyncing
12 entries were displayed.
```

e. Confirm that all aggregates are online and have resynchronized:

```
storage aggregate plex show
```

The following output shows that all aggregates have resynchronized.

```
cluster_A::> storage aggregate plex show
()
```

Aggregate Plex	Is Online	Is Resyncing	Resyncing Percent	Status
node_B_1_aggr0 plex0	true	false		- normal,active
node_B_1_aggr0 plex8	true	false		- normal,active
node_B_2_aggr0 plex0	true	false		- normal,active
node_B_2_aggr0 plex8	true	false		- normal,active
node_B_1_aggr1 plex0	true	false		- normal,active
node_B_1_aggr1 plex9	true	false		- normal,active
node_B_1_aggr2 plex0	true	false		- normal,active
node_B_1_aggr2 plex5	true	false		- normal,active
node_B_2_aggr1 plex0	true	false		- normal,active
node_B_2_aggr1 plex9	true	false		- normal,active
node_B_2_aggr2 plex0	true	false		- normal,active
node_B_2_aggr2 plex5	true	false		- normal,active
node_A_1_aggr1 plex4	true	false		- normal,active
node_A_1_aggr1 plex8	true	false		- normal,active
node_A_1_aggr2 plex1	true	false		- normal,active
node_A_1_aggr2 plex5	true	false		- normal,active
node_A_2_aggr1 plex4	true	false		- normal,active
node_A_2_aggr1 plex8	true	false		- normal,active
node_A_2_aggr2 plex1	true	false		- normal,active
node_A_2_aggr2 plex5	true	false		- normal,active

```
20 entries were displayed.
```

4. On systems running ONTAP 9.5 and earlier, perform the root-aggregates healing phase:

```
metrocluster heal -phase root-aggregates
```

```
cluster_B::> metrocluster heal -phase root-aggregates
[Job 651] Job is queued: MetroCluster Heal Root Aggregates Job.Oct 26
13:05:00
[Job 651] Job succeeded: Heal Root Aggregates is successful.
```

5. Verify that the "heal roots" phase has completed and the disaster site is ready for switchback:

The following output shows that the "heal roots" phase has completed on cluster_A.

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

cluster_B::>
```

Proceed to verify the licenses on the replaced nodes.

[Verifying licenses on the replaced nodes](#)

Copyright information

Copyright © 2022 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.

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.