Recover from a disaster

ONTAP MetroCluster

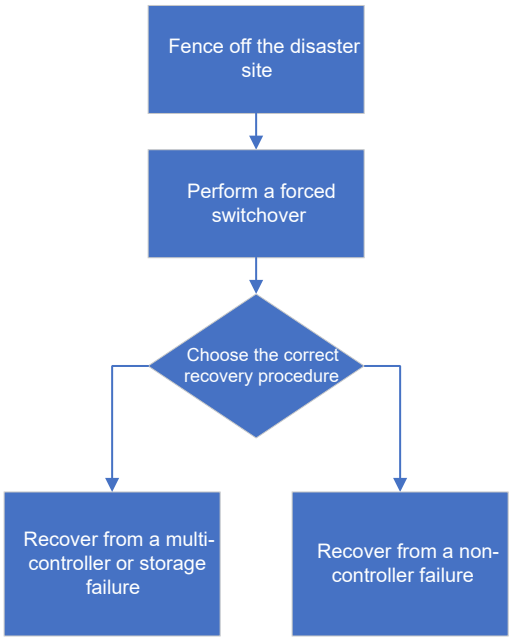
NetApp
August 29, 2025

Table of Contents

Recover from a disaster

Workflow for disaster recovery

Use the workflow to perform disaster recovery.



Performing a forced switchover after a disaster

If a disaster has occurred, there are steps you must perform on both the disaster cluster and the surviving cluster after the switchover to ensure safe and continued data service.

Determining if a disaster has occurred is done by:

- An administrator
- The MetroCluster Tiebreaker software, if it is configured
- The ONTAP Mediator software, if it is configured

Fencing off the disaster site

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

Step

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

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

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

Performing a forced switchover

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

Before you begin

- At least one of the surviving site nodes must be up and running before you perform the switchover.
- All previous configuration changes must be complete before performing a switchback operation.

  This is to avoid competition with the negotiated switchover or switchback operation.

SnapMirror and SnapVault configurations are deleted automatically.

About this task

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

Steps

1. Perform the switchover by running the following command at the surviving site:

   ```
   metrocluster switchover -forced-on-disaster true
   ```

   The operation can take a period of minutes to complete. You can verify progress using the `metrocluster operation show` command.

2. Answer `y` when prompted to continue with the switchover.
3. Verify that the switchover was completed successfully by running the `metrocluster operation show` command.

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

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

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

After you finish

SnapMirror relationships need to be reestablished after switchover.

Output for the storage aggregate plex show command is indeterminate after a MetroCluster switchover

When you run the `storage aggregate plex show` command after a MetroCluster switchover, the status of plex0 of the switched over root aggregate is indeterminate and is displayed as failed. During this time, the switched over root is not updated. The actual status of this plex can only be determined after the MetroCluster healing phase.

Accessing volumes in NVFAIL state after a switchover

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

Before you begin

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

About this task

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

Step

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

After you finish

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

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

Related information

Monitoring and protecting database validity by using NVFAIL

Choosing the correct recovery procedure

After a failure in a MetroCluster configuration, you must select the correct recovery procedure. Use the following table and examples to select the appropriate recovery procedure.

This information in this table assumes that the installation or transition is complete, meaning that the `metrocluster configure` command ran successfully.

Scope of failures at disaster site	Procedure

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

Controller module failure scenarios during MetroCluster installation

Responding to a controller module failure during the MetroCluster configuration procedure depends on whether the `metrocluster configure` command successfully completed.

- If the `metrocluster configure` command was not yet run, or failed, you must restart the MetroCluster software configuration procedure from the beginning with a replacement controller module.

  You must be sure to perform the steps in Restoring system defaults on a controller module on each controller (including the replacement controller) to verify that the previous configuration is removed.

- If the `metrocluster configure` command successfully completed and then the controller module failed, use the previous table to determine the correct recovery procedure.

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

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

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

Controller module failure scenarios in eight-node MetroCluster configurations

Failure scenarios:

- Single controller module failures in a single DR group
- Two controller module failures in a single DR group
- Single controller module failures in separate DR groups
- Three controller module failures spread across the DR groups

Single controller module failures in a single DR group

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

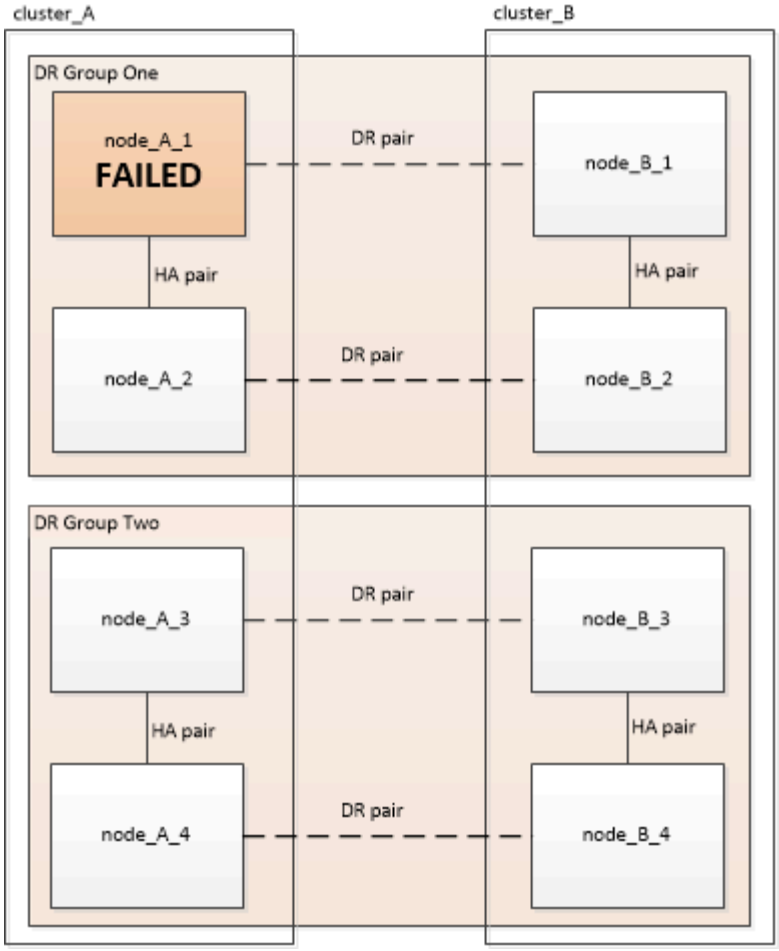
- If no storage requires replacement, you can use the controller module FRU replacement procedure for the platform model.

  ONTAP Hardware Systems Documentation

- If storage requires replacement, you can use the multi-controller module recovery procedure.

  Recovering from a multi-controller or storage failure

  This scenario applies to four-node MetroCluster configurations also.

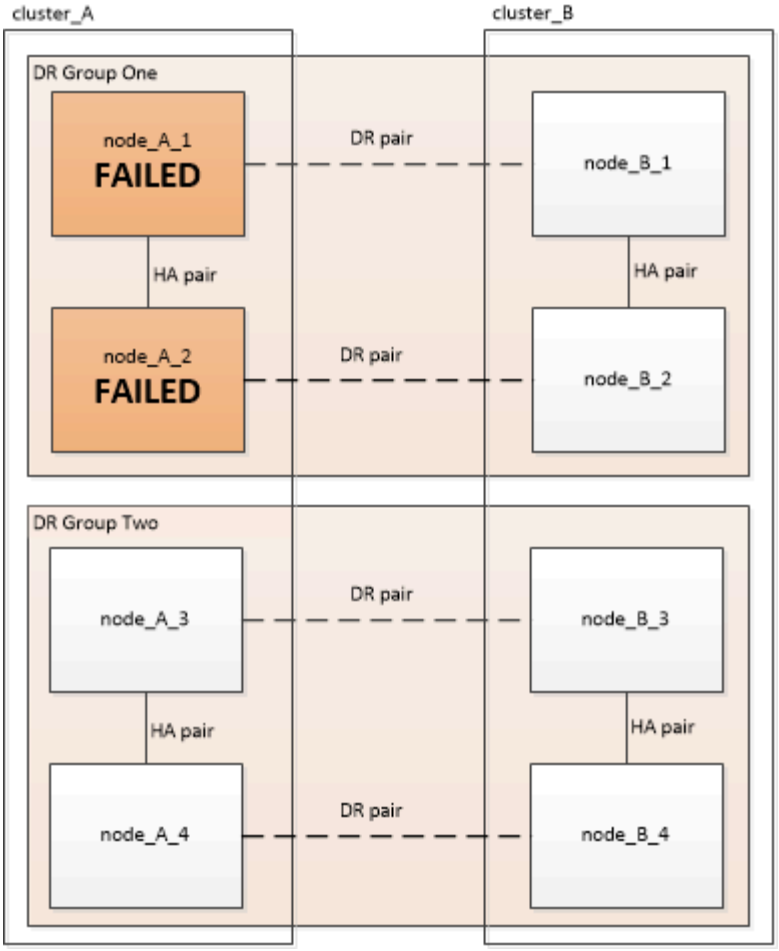


Two controller module failures in a single DR group

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

Recovering from a multi-controller or storage failure

This scenario applies to four-node MetroCluster configurations also.



Single controller module failures in separate DR groups

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

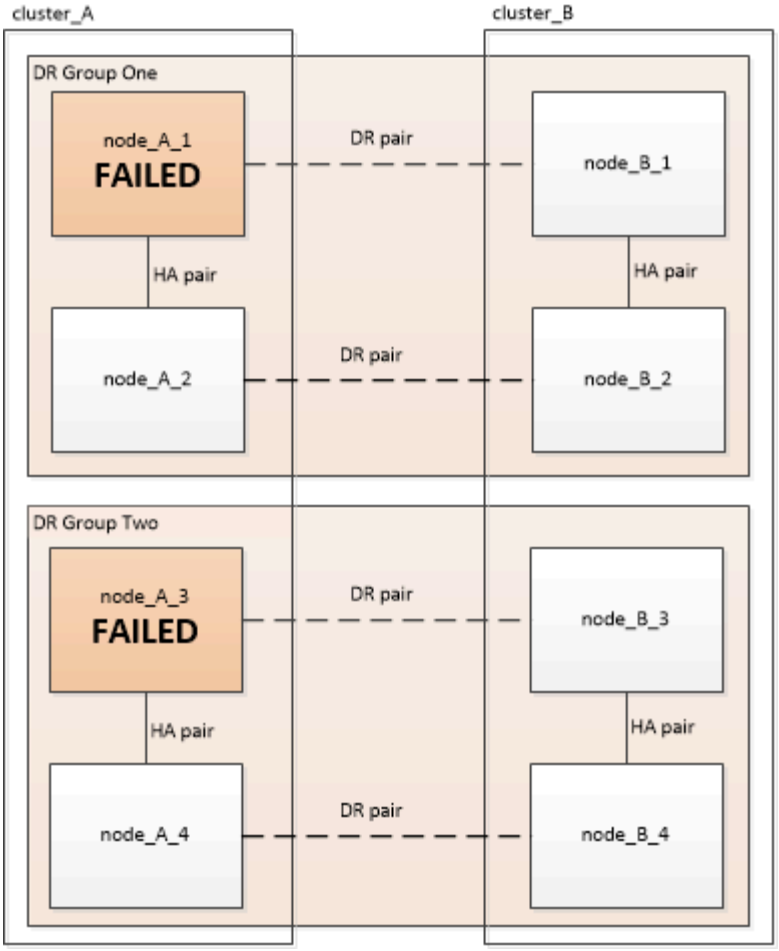
- If no storage requires replacement, you can use the controller module FRU replacement procedure for the platform model.

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

  ONTAP Hardware Systems Documentation

- If storage requires replacement, you can use the multi-controller module recovery procedure.

  Recovering from a multi-controller or storage failure



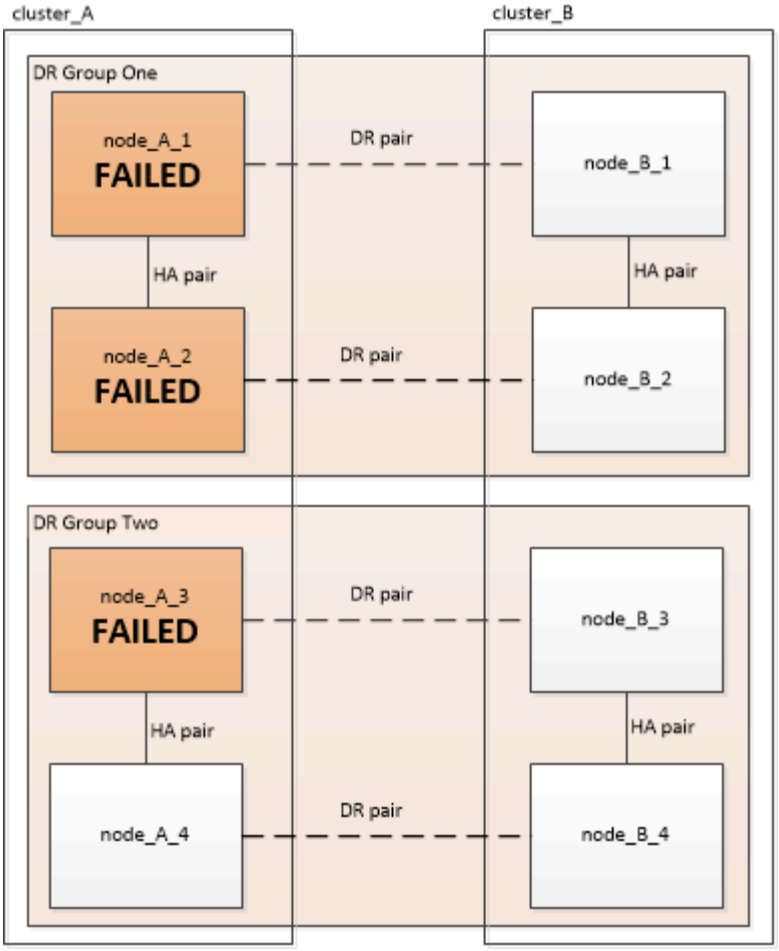
Three controller module failures spread across the DR groups

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

Recovering from a multi-controller or storage failure

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

ONTAP Hardware Systems Documentation



Controller module failure scenarios in two-node MetroCluster configurations

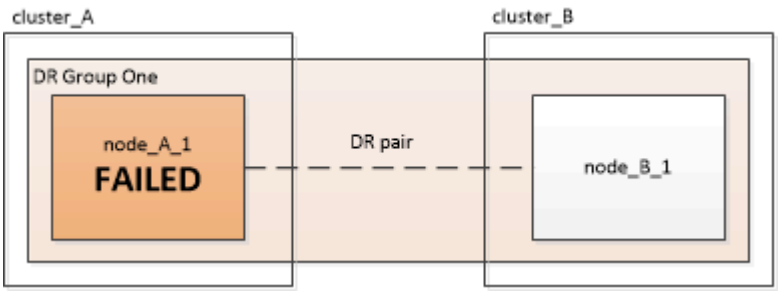
The procedure you use depends on the extent of the failure.

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

  ONTAP Hardware Systems Documentation

- If storage requires replacement, you can use the multi-controller module recovery procedure.

  Recovering from a multi-controller or storage failure



Recover from a multi-controller or storage failure

Recovering from a multi-controller or storage failure

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

Verify that you have checked and performed the following tasks before using this procedure:

- Review the available recovery procedures before deciding to use this procedure.

  Choosing the correct recovery procedure

- Confirm that console logging is enabled on your devices.

  Enable console logging

- Ensure that the disaster site is fenced off.

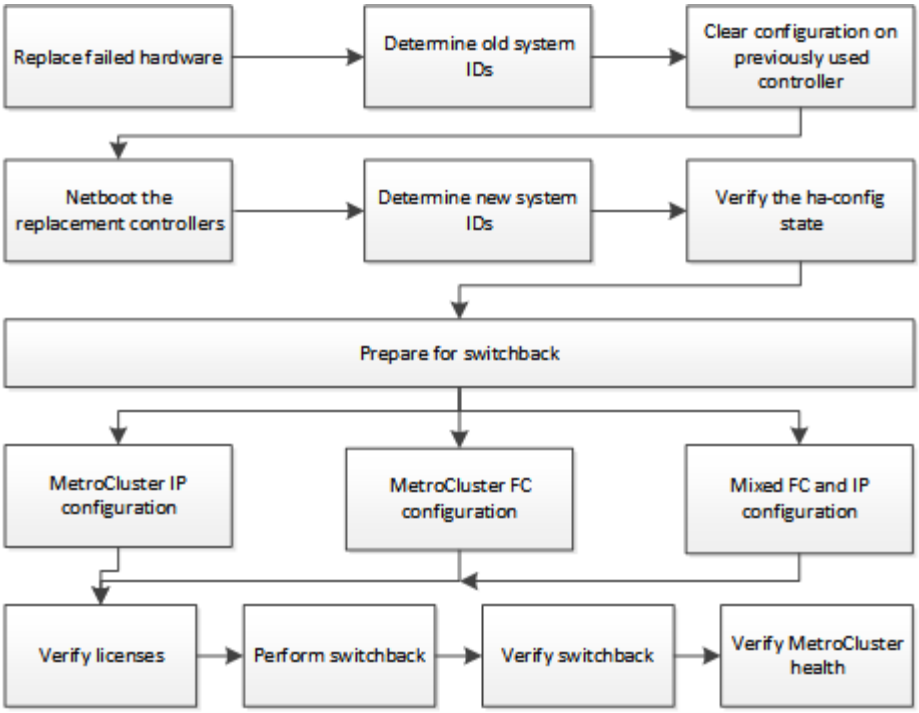
  Fencing off the disaster site.

- Verify that switchover was performed.

  Performing a forced switchover.

- Verify that the replacement drives and the controller modules are new and must not have been assigned ownership previously.
- The examples in this procedure show two or four-node configurations. If you have an eight-node configuration (two DR groups), you must take into account any failures and perform the required recovery task on the additional controller modules.

This procedure uses the following workflow:



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

Enable console logging

Enable console logging on your devices before proceeding to replace hardware and boot new controllers.

NetApp strongly recommends that you enable console logging on the devices that you are using and take the following actions when performing this procedure:

- Leave AutoSupport enabled during maintenance.
- Trigger a maintenance AutoSupport message before and after maintenance to disable case creation for the duration of the maintenance activity.

  See the Knowledge Base article How to suppress automatic case creation during scheduled maintenance windows.

- Enable session logging for any CLI session. For instructions on how to enable session logging, review the "Logging Session Output" section in the Knowledge Base article How to configure PuTTY for optimal connectivity to ONTAP systems.

Replace hardware and boot new controllers

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

Replace hardware at the disaster site

Before you begin

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

Steps

1. Replace the components as necessary.

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

If you are replacing…	Perform these steps…	Using these guides…
FC switches in a MetroCluster FC configuration	a. Install the new switches. b. Cable the ISL links. Do not power on the FC switches at this time.	Maintain MetroCluster Components
IP switches in a MetroCluster IP configuration	a. Install the new switches. b. Cable the ISL links. Do not power on the IP switches at this time.	MetroCluster IP installation and configuration: Differences among the ONTAP MetroCluster configurations
Disk shelves	a. Install the disk shelves and disks. ◦ Disk shelf stacks should be the same configuration as at the surviving site. ◦ Disks can be the same size or larger, but must be of the same type (SAS or SATA). b. Cable the disk shelves to adjacent shelves within the stack and to the FC-to-SAS bridge. Do not power on the disk shelves at this time.	ONTAP Hardware Systems Documentation
SAS cables	a. Install the new cables. Do not power on the disk shelves at this time.	ONTAP Hardware Systems Documentation

FC-to-SAS bridges in a MetroCluster FC configuration	a. Install the FC-to-SAS bridges. b. Cable the FC-to-SAS bridges. Cable them to the FC switches or to the controller modules, depending on your MetroCluster configuration type. Do not power on the FC-to-SAS bridges at this time.	Fabric-attached MetroCluster installation and configuration Stretch MetroCluster installation and configuration

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

2. Verify that all components are cabled correctly for your configuration.
   - MetroCluster IP configuration
   - MetroCluster fabric-attached configuration

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

After you have replaced all hardware at the disaster site, you must determine the system IDs of the replaced controller modules. You need the old system IDs when you reassign disks to the new controller modules. If the

systems are AFF A220, AFF A250, AFF A400, AFF A800, FAS2750, FAS500f, FAS8300, or FAS8700 models, you must also determine the VLAN IDs used by the MetroCluster IP interfaces.

Before you begin

All equipment at the disaster site must be powered off.

About this task

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

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

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

- Site A is the disaster site.
- node_A_1 has failed and is being completely replaced.
- node_A_2 has failed and is being completely replaced.

  node _A_2 is present in a four-node MetroCluster configuration only.

- Site B is the surviving site.
- node_B_1 is healthy.
- node_B_2 is healthy.

  node_B_2 is present in a four-node MetroCluster configuration only.

The controller modules have the following original system IDs:

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

Steps

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

Number of nodes in MetroCluster configuration	Use this command

Four or eight	`metrocluster node show -fields node-systemid,ha-partner-systemid,dr-partner-systemid,dr-auxiliary-systemid`
Two	`metrocluster node show -fields node-systemid,dr-partner-systemid`

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

- Node_A_1: 4068741258
- Node_A_2: 4068741260

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

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

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

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

- Node_A_1: 4068741258

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

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

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

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

   ```
   storage iscsi-initiator show -node * -label mediator
   ```

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

   ```
   metrocluster interconnect show
   ```

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

   In this example, the VLAN IDs are 120 and 130:

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

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

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

About this task

This procedure is only required on MetroCluster IP configurations.

Steps

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

   ```
   set -privilege advanced
   ```

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

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

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

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

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

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

3. Return to the admin privilege level:

   ```
   set -privilege admin
   ```

Clear the configuration on a controller module

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

Steps

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

   ```
   halt
   ```

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

   ```
   set-defaults
   ```

3. Save the environment:

   ```
   saveenv
   ```

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

   ```
   boot_ontap menu
   ```

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

   ```
   wipeconfig
   ```

   Respond `yes` to the confirmation prompt.

   The node reboots and the boot menu is displayed again.

6. At the boot menu, select option **5** to boot the system into Maintenance mode.

   Respond `yes` to the confirmation prompt.

Netboot the new controller modules

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

Before you begin

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

  NetApp Support

Steps

1. Access the NetApp Support Site to download the files used for performing the netboot of the system.
2. Download the appropriate ONTAP software from the software download section of the NetApp Support Site and store the ontap-version_image.tgz file on a web-accessible directory.
3. Go to the web-accessible directory and verify that the files you need are available.

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

4. At the LOADER prompt, configure the netboot connection for a management LIF:
   - If IP addressing is DHCP, configure the automatic connection:

     ```
     ifconfig e0M -auto
     ```

   - If IP addressing is static, configure the manual connection:

     ```
     ifconfig e0M -addr=ip_addr -mask=netmask -gw=gateway
     ```

5. Perform the netboot.

- If the platform is an 80xx series system, use this command:

  ```
  netboot http://web_server_ip/path_to_web-accessible_directory/netboot/kernel
  ```

- If the platform is any other system, use the following command:

  ```
  netboot http://web_server_ip/path_to_web-accessible_directory/ontap-version_image.tgz
  ```

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

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

7. If you are prompted to continue the procedure, enter `y`, and when prompted for the package, enter the URL of the image file: `http://web_server_ip/path_to_web-accessible_directory/ontap-version_image.tgz`

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

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

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

9. Reboot by entering `y` when you see a prompt similar to the following:

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

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

Determine the system IDs of the replacement controller modules

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

About this task

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

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

- Site A is the disaster site.
- node_A_1 has been replaced.
- node_A_2 has been replaced.

  Present only in four-node MetroCluster configurations.

- Site B is the surviving site.
- node_B_1 is healthy.
- node_B_2 is healthy.

  Present only in four-node MetroCluster configurations.

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

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

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

In the preceding example:

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

Steps

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

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

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

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

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

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

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

Verify the ha-config state of components

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

Before you begin

The system must be in Maintenance mode.

About this task

This task must be performed on each new controller module.

Steps

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

   `ha-config show`

   The correct HA state depends on your MetroCluster configuration.

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

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

Number of controllers in the MetroCluster configuration	Command

Eight- or four-node MetroCluster FC configuration	`ha-config modify controller mcc`
Two-node MetroCluster FC configuration	`ha-config modify controller mcc-2n`
MetroCluster IP configuration	`ha-config modify controller mccip`

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

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

4. Repeat these steps on the other replacement node.

Determine if end-to-end encryption was enabled on the original systems

You should verify if the original systems were configured for end-to-end encryption.

Step

1. Run the following command from the surviving site:

   ```
   metrocluster node show -fields is-encryption-enabled
   ```

   If encryption is enabled, the following output is displayed:

   ```
   1 cluster_A node_A_1 true
   1 cluster_A node_A_2 true
   1 cluster_B node_B_1 true
   1 cluster_B node_B_2 true
   4 entries were displayed.
   ```

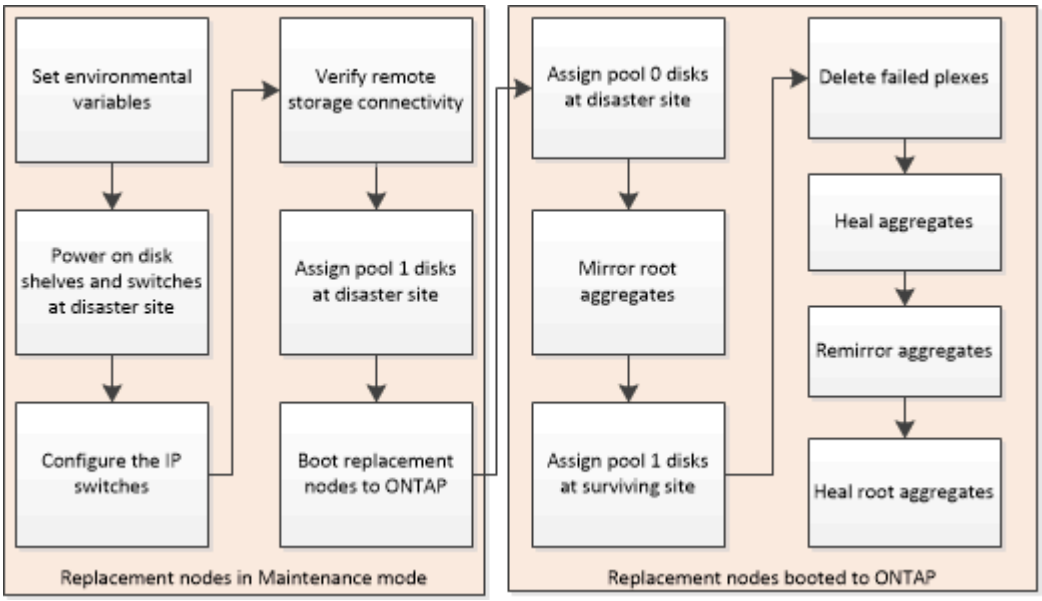
   Refer to Configure end-to-end encryption for supported systems.

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.
- Certain platforms use a VLAN for the MetroCluster IP interface. By default, each of the two ports use a different VLAN: 10 and 20.

  If supported, you can also specify a different (non-default) VLAN higher than 100 (between 101 and 4095) using the `vlan-id` parameter.

  The following platforms do **not** support the `vlan-id` parameter:

  - FAS8200 and AFF A300
  - AFF A320
  - FAS9000 and AFF A700
  - AFF C800, ASA C800, AFF A800 and ASA A800

    All other platforms support the `vlan-id` parameter.

- 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 MetroCluster bootarg values you set depend on whether your new system uses shared cluster/HA ports or shared MetroCluster/HA ports. Use the following information to determine the ports for your system.

Shared cluster/HA ports

The systems listed in the following table use shared cluster/HA ports:

AFF and ASA systems	FAS systems
• AFF A20 • AFF A30 • AFF C30 • AFF A50 • AFF C60 • AFF C80 • AFF A70 • AFF A90 • AFF A1K	• FAS50 • FAS70 • FAS90

Shared MetroCluster/HA ports

The systems listed in the following table use shared MetroCluster/HA ports:

AFF and ASA systems	FAS systems
• AFF A150, ASA A150 • AFF A220 • AFF C250, ASA C250 • AFF A250, ASA A250 • AFF A300 • AFF A320 • AFF C400, ASA C400 • AFF A400, ASA A400 • AFF A700 • AFF C800, ASA C800 • AFF A800, ASA A800 • AFF A900, ASA A900	• FAS2750 • FAS500f • FAS8200 • FAS8300 • FAS8700 • FAS9000 • FAS9500

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 command output varies depending on whether your platform model uses shared cluster/HA ports or shared MetroCluster/HA ports.

Systems using shared cluster/HA ports

```
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    172.17.26.10    DR Partner
completed
                 Home Port: e5a
                      172.17.26.13    172.17.26.11    DR Auxiliary
completed
                 Home Port: e5b
                      172.17.27.13    172.17.27.10    DR Partner
completed
                 Home Port: e5b
                      172.17.27.13    172.17.27.11    DR Auxiliary
completed
              node_B_2
                 Home Port: e5a
                      172.17.26.12    172.17.26.11    DR Partner
completed
                 Home Port: e5a
                      172.17.26.12    172.17.26.10    DR Auxiliary
completed
                 Home Port: e5b
                      172.17.27.12    172.17.27.11    DR Partner
completed
                 Home Port: e5b
                      172.17.27.12    172.17.27.10    DR Auxiliary
completed
12 entries were displayed.
```

Systems using shared MetroCluster/HA ports

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 can vary depending on the platform type.

```
cluster_B::*> metrocluster configuration-settings connection show
DR                    Source          Destination
```

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

2. If 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 to determine the VLAN ID if the platform models support 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. At the `LOADER` prompt for each of the disaster site nodes, set the bootarg value depending on whether your platform model uses shared cluster/HA ports or shared MetroCluster/HA ports:

   - If the interfaces are using the default VLANs, or the platform model does not use a VLAN ID (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).

Systems using shared cluster/HA ports

Set the following bootarg:

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

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

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
172.17.26.10/23,0,0,172.17.26.13,172.17.26.12,120

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

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

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

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

Systems using shared MetroCluster/HA ports

Set the following bootarg:

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

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

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. 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

5. 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
```

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

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

7. 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
      ```

8. If the original systems were configured for end-to-end encryption, at each of the replacement nodes' LOADER prompt, set the following bootarg:

   ```
   setenv bootarg.mccip.encryption_enabled 1
   ```

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
```

Determine the system 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	`(IP_Switch_A_1)> enable`
	`(IP_switch_A_1)# configure`
	`(IP_switch_A_1)(Config)# interface 0/13-0/16`
	`(IP_switch_A_1)(Interface 0/13-0/16 )# no shutdown`
	`(IP_switch_A_1)(Interface 0/13-0/16 )# exit`
	`(IP_switch_A_1)(Config)# exit`

Cisco	`IP_switch_A_1# conf t`
	`IP_switch_A_1(config)# int eth1/15-eth1/20`
	`IP_switch_A_1(config)# no shutdown`
	`IP_switch_A_1(config)# copy running startup`
	`IP_switch_A_1(config)# show interface brief`

   b. Enable the ISL interfaces on the partner switch:

```
no shutdown
```

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

Switch vendor	Commands

Broadcom	`(IP_Switch_A_2)> enable`
	`(IP_switch_A_2)# configure`
	`(IP_switch_A_2)(Config)# interface 0/13-0/16`
	`(IP_switch_A_2)(Interface 0/13-0/16 )# no shutdown`
	`(IP_switch_A_2)(Interface 0/13-0/16 )# exit`
	`(IP_switch_A_2)(Config)# exit`

Cisco	`IP_switch_A_2# conf t`
	`IP_switch_A_2(config)# int eth1/15-eth1/20`
	`IP_switch_A_2(config)# no shutdown`
	`IP_switch_A_2(config)# copy running startup`
	`IP_switch_A_2(config)# show interface brief`

c. Verify that the interfaces are enabled:

```
show interface brief
```

The following example shows the output for a Cisco switch.

```
IP_switch_A_2(config)# show interface brief

--------------------------------------------------------
Port VRF Status IP Address Speed MTU
--------------------------------------------------------
mt0 -- up 10.10.99.10 100 1500
--------------------------------------------------------
Ethernet    VLAN Type Mode    Status Reason Speed   Port
Interface                                       Ch
#
--------------------------------------------------------
.
.
.
Eth1/15    10   eth   access  up     none   40G(D)  --
Eth1/16    10   eth   access  up     none   40G(D)  --
Eth1/17    10   eth   access  down   none   auto(D) --
Eth1/18    10   eth   access  down   none   auto(D) --
Eth1/19    10   eth   access  down   none   auto(D) --
Eth1/20    10   eth   access  down   none   auto(D) --
.
.
.
IP_switch_A_2#
```

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 Replace hardware and boot new controllers.

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)	`disk reassign -s old-system-ID -d new-system-ID -r dr-partner-system-ID`
Yes (ONTAP 9.7.x and earlier)	`disk reassign -s old-system-ID -d new-system-ID -p old-partner-system-ID`
No	`disk reassign -s old-system-ID -d new-system-ID`

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

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

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

2. Destroy the contents of the mailbox disks:

   ```
   mailbox destroy local
   ```

   You can proceed with the destroy operation when prompted.

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

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

3. If disks have been replaced, there will be failed local plexes that must be deleted.
   a. Display the aggregate status:

      ```
      aggr status
      ```

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

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

b. Delete the failed plex:

`aggr destroy plex-id`

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

4. Halt the node to display the LOADER prompt:

   `halt`

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

Booting to ONTAP on replacement controller modules in MetroCluster IP configurations

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

About this task

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

Steps

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

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

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

      `printenv`

   c. Set the value to the correct partner system ID:

      `setenv partner-sysid partner-sysID`

   d. Boot the node:

      `boot_ontap`

   e. Repeat these substeps on the other node, if necessary.
6. Confirm that the replacement nodes at the disaster site are ready for switchback:

   `metrocluster node show`

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

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

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

cluster_B::>
```

7. Verify the MetroCluster connection configuration settings:

   `metrocluster configuration-settings connection show`

   The configuration state should indicate completed.

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

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

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

cluster_B::*>
```

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

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
                  Usable     Disk      Container           Container
Disk              Size       Shelf Bay Type    Type        Name
Owner
---------------- ---------- ----- --- ------- ----------- ---------
--------
node_A_1:0n.12   1.75TB     0     12  SSD-NVM shared      aggr0
node_A_1
node_A_1:0n.13   1.75TB     0     13  SSD-NVM shared      aggr0
node_A_1
node_A_1:0n.14   1.75TB     0     14  SSD-NVM shared      aggr0
node_A_1
node_A_1:0n.15   1.75TB     0     15  SSD-NVM shared      aggr0
node_A_1
node_A_1:0n.16   1.75TB     0     16  SSD-NVM shared      aggr0
node_A_1
node_A_1:0n.17   1.75TB     0     17  SSD-NVM shared      aggr0
node_A_1
node_A_1:0n.18   1.75TB     0     18  SSD-NVM shared      aggr0
node_A_1
node_A_1:0n.19   1.75TB     0     19  SSD-NVM shared      -
node_A_1
node_A_2:0n.0    1.75TB     0     0   SSD-NVM shared
aggr0_node_A_2_0 node_A_2
node_A_2:0n.1    1.75TB     0     1   SSD-NVM shared
aggr0_node_A_2_0 node_A_2
node_A_2:0n.2    1.75TB     0     2   SSD-NVM shared
aggr0_node_A_2_0 node_A_2
node_A_2:0n.3    1.75TB     0     3   SSD-NVM shared
aggr0_node_A_2_0 node_A_2
node_A_2:0n.4    1.75TB     0     4   SSD-NVM shared
aggr0_node_A_2_0 node_A_2
node_A_2:0n.5    1.75TB     0     5   SSD-NVM shared
aggr0_node_A_2_0 node_A_2
node_A_2:0n.6    1.75TB     0     6   SSD-NVM shared
aggr0_node_A_2_0 node_A_2
node_A_2:0n.7    1.75TB     0     7   SSD-NVM shared      -
node_A_2
node_A_2:0n.24   -          0     24  SSD-NVM unassigned  -         -
node_A_2:0n.25   -          0     25  SSD-NVM unassigned  -         -
```

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

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

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

About this task

For ADP systems, the drives are assigned automatically.

Steps

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

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

   This command assigns the newly added (and unowned) drives on the disaster site. You should assign the same number and size (or larger) of drives that the node had prior to the disaster. The `storage disk assign` man page contains 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                     State          Start Time       End
Time
----------------------------- -------------- ----------------
----------------
heal-root-aggr-auto           successful      2/25/2019 06:45:58

2/25/2019 06:46:02
heal-aggr-auto                successful     2/25/2019 06:45:48

2/25/2019 06:45:52
.
.
.
```

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

```
metrocluster node show
```

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

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

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

a. Verify the state of the nodes:

```
metrocluster node show
```

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

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

cluster_B::>
```

b. Perform the aggregates healing phase:

```
metrocluster heal -phase aggregates
```

The following output shows a typical aggregates healing operation.

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

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

cluster_B::*>
```

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

cluster_B::>
```

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

      `storage aggregate show`

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

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

normal
node_B_1_aggr1 3.14TB  3.04TB    3% online     15 node_B_1
raid_dp,

normal
node_B_1_aggr2 3.14TB  3.06TB    3% online     14 node_B_1
raid_tec,
```

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

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

normal

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

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

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

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

normal
12 entries were displayed.

cluster_B::>
```

b. Mirror the aggregate:

   `storage aggregate mirror -aggregate aggregate-name`

   The following output shows a typical mirroring operation.

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

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

      Second Plex

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

      Aggregate capacity available for volume use would be 2.99TB.

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

c. Repeat the previous step for each of the aggregates from the surviving site.

d. Wait for the aggregates to resynchronize; you can check the status with the `storage aggregate show` command.

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

```
cluster_B::> storage aggregate show

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

```
mirrored,

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

mirrored,

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

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

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

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

resyncing

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

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

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

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

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

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

```
storage aggregate plex show
```

The following output shows that all aggregates have resynchronized.

```
cluster_A::> storage aggregate plex show
  ()
                    Is      Is         Resyncing
Aggregate Plex      Online  Resyncing    Percent Status
--------- --------- ------- ---------- --------- ---------------
node_B_1_aggr0 plex0 true    false              - normal,active
node_B_1_aggr0 plex8 true    false              - normal,active
node_B_2_aggr0 plex0 true    false              - normal,active
node_B_2_aggr0 plex8 true    false              - normal,active
node_B_1_aggr1 plex0 true    false              - normal,active
node_B_1_aggr1 plex9 true    false              - normal,active
node_B_1_aggr2 plex0 true    false              - normal,active
node_B_1_aggr2 plex5 true    false              - normal,active
node_B_2_aggr1 plex0 true    false              - normal,active
node_B_2_aggr1 plex9 true    false              - normal,active
node_B_2_aggr2 plex0 true    false              - normal,active
node_B_2_aggr2 plex5 true    false              - normal,active
node_A_1_aggr1 plex4 true    false              - normal,active
node_A_1_aggr1 plex8 true    false              - normal,active
node_A_1_aggr2 plex1 true    false              - normal,active
node_A_1_aggr2 plex5 true    false              - normal,active
node_A_2_aggr1 plex4 true    false              - normal,active
node_A_2_aggr1 plex8 true    false              - normal,active
node_A_2_aggr2 plex1 true    false              - normal,active
node_A_2_aggr2 plex5 true    false              - normal,active
20 entries were displayed.
```

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

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

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

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

cluster_B::>
```

Proceed to verify the licenses on the replaced nodes.

Verifying licenses on the replaced nodes

Prepare for switchback in a MetroCluster FC configuration

Verifying port configuration (MetroCluster FC configurations only)

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

About this task

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

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

Steps

1. In Maintenance mode, restore the FC port configuration:

   `ucadmin modify -m fc -t initiatoradapter_name`

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

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

If the FC port configuration is…	Then…
The same for both ports	Answer “y” when prompted by the system, because modifying one port in a port pair also modifies the other port.
Different	a. Answer “n” when prompted by the system. b. Restore the FC port configuration: `ucadmin modify -m fc -t initiator|targetadapter_name`

3. Exit Maintenance mode:

   `halt`

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

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

   `boot_ontap maint`

5. Verify the values of the variables:

   `ucadmin show`

6. Exit Maintenance mode and display the LOADER prompt:

   `halt`

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

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

Steps

1. Power on the FC-to-SAS bridges.
2. Set the IP address on the Ethernet ports by using the `set IPAddress port ipaddress` command.
   - `port` can be either "MP1" or "MP2".
   - `ipaddress` can be an IP address in the format xxx.xxx.xxx.xxx.

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

```
Ready.
set IPAddress MP1 10.10.10.55

Ready. *
```

3. Set the IP subnet mask on the Ethernet ports by using the `set IPSubnetMask port mask` command.
   - `port` can be "MP1" or "MP2".
   - `mask` can be a subnet mask in the format xxx.xxx.xxx.xxx.

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

```
Ready.
set IPSubnetMask MP1 255.255.255.0

Ready. *
```

4. Set the speed on the Ethernet ports by using the `set EthernetSpeed port speed` command.
   - `port` can be "MP1" or "MP2".
   - `speed` can be "100" or "1000".

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

```
Ready.
set EthernetSpeed MP1 1000

Ready. *
```

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

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

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

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

6. After the FC-to-SAS bridge reboots, log in again.
7. Set the speed on the FC ports by using the `set fcdatarate port speed` command.
   - `port` can be "1" or "2".
   - `speed` can be "2 Gb", "4 Gb", "8 Gb", or "16 Gb", depending on your model bridge.

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

   ```
   Ready.
   set fcdatarate 1 8Gb

   Ready. *
   ```

8. Set the topology on the FC ports by using the `set FCConnMode port mode` command.
   - `port` can be "1" or "2".
   - `mode` can be "ptp", "loop", "ptp-loop", or "auto".

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

   ```
   Ready.
   set FCConnMode 1 ptp

   Ready. *
   ```

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

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

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

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

```
Ready.
set snmp enabled

Ready. *
saveconfiguration

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

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

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

Configuring the FC switches (MetroCluster FC configurations only)

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

Related tasks

Port assignments for FC switches

Configuring a Brocade FC switch after site disaster

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

About this task

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

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

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

- Port assignments for FC switches

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

Steps

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

      `firmwareShow`

   c. Configure the new switch as described in the following topics, skipping the steps for configuring zoning on the switch.

      Fabric-attached MetroCluster installation and configuration

      Stretch MetroCluster installation and configuration

   d. Disable the switch persistently:

      `switchcfgpersistentdisable`

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

      The following example shows the command on BrocadeSwitchA:

      ```
      BrocadeSwitchA:admin> switchcfgpersistentdisable
      ```

      The following example shows the command on BrocadeSwitchB:

      ```
      BrocadeSwitchA:admin> switchcfgpersistentdisable
      ```

2. Complete configuration of the new switch:
   a. Enable the ISLs on the surviving site:

      `portcfgpersistentenable port-number`

      ```
      FC_switch_B_1:admin> portcfgpersistentenable 10
      FC_switch_B_1:admin> portcfgpersistentenable 11
      ```

   b. Enable the ISLs on the replacement switches:

      `portcfgpersistentenable port-number`

      ```
      FC_switch_A_1:admin> portcfgpersistentenable 10
      FC_switch_A_1:admin> portcfgpersistentenable 11
      ```

   c. On the replacement switch (FC_switch_A_1 in this example) verify that the ISL's are online:

```
switchshow
```

```
FC_switch_A_1:admin> switchshow
switchName: FC_switch_A_1
switchType: 71.2
switchState:Online
switchMode: Native
switchRole: Principal
switchDomain:       4
switchId:   fffc03
switchWwn:  10:00:00:05:33:8c:2e:9a
zoning:             OFF
switchBeacon:       OFF

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

3. Persistently enable the switch:

   ```
   switchcfgpersistentenable
   ```

4. Verify that the ports are online:

   ```
   switchshow
   ```

Configuring a Cisco FC switch after site disaster

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

About this task

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

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

Steps

1. Configure the switch:
   a. Refer to Fabric-attached MetroCluster installation and configuration
   b. Follow the steps for configuring the switch in Configuring the Cisco FC switches section, *except* for the "Configuring zoning on a Cisco FC switch" section:

      Zoning is configured later in this procedure.

2. On the healthy switch (in this example, FC_switch_B_1), enable the ISL ports.

   The following example shows the commands to enable the ports:

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

3. Verify that the ISL ports are up by using the show interface brief command.
4. Retrieve the zoning information from the fabric.

   The following example shows the commands to distribute the zoning configuration:

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

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

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

```
show zone
```

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

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

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

6. Determine the worldwide names (WWNs) of the switches in the switch fabric.

   In this example, the two switch WWNs are as follows:

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

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

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

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

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

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

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

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

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

   The following example shows the removal of these interfaces:

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

8. Add the ports of the new switch to the zones.

   The following example assumes that the cabling on the replacement switch is the same as on the old switch:

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

9. Verify that the zoning is properly configured: `show zone`

   The following example output shows the three zones:

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

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

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

Verifying the storage configuration

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

Steps

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

   The surviving site and disaster site should have the same number of disk shelf stacks, disk shelves, and disks. In a bridge-attached or fabric-attached MetroCluster configuration, the sites should have the same number of FC-to-SAS bridges.

2. Confirm that all disks that have been replaced at the disaster site are unowned:

   `run local disk show-n`

   Disks should appear as being unowned.

3. If no disks were replaced, confirm that all disks are present:

   ```
   disk show
   ```

Powering on the equipment at the disaster site

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

Before you begin

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

Fabric-attached MetroCluster installation and configuration

Stretch MetroCluster installation and configuration

About this task

The examples in this procedure assume the following:

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

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

Steps

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

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

2. On the surviving site, disable disk autoassignment:

   ```
   storage disk option modify -autoassign off *
   ```

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

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

   ```
   storage disk option show
   ```

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

 cluster_B::>
```

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

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

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

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

Assigning ownership for replaced drives

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

Before you begin

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

The drives shelves and drives arrangement must meet the requirements in Required MetroCluster IP component and naming conventions section of the MetroCluster IP installation and configuration.

About this task

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

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

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

- Site A is the disaster site.
  - node_A_1 has been replaced.
  - node_A_2 has been replaced.

Present only in four-node MetroCluster configurations.

- Site B is the surviving site.
  - node_B_1 is healthy.
  - node_B_2 is healthy.

    Present only in four-node MetroCluster configurations.

The controller modules have the following original system IDs:

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

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

- The old-count-of-disks must be at least the same number of disks for each node that were present before the disaster.

  If a lower number of disks is specified or present, the healing operations might not be completed due to insufficient space.

- The new plexes to be created are remote plexes belonging to the surviving site (node_B_x pool1) and local plexes belonging to the disaster site (node_B_x pool0).
- The total number of required drives should not include the root aggr disks.

  If n disks are assigned to pool1 of the surviving site, then n-3 disks should be assigned to the disaster site with the assumption that the root aggregate uses three disks.

- None of the disks can be assigned to a pool that is different from the one to which all other disks on the same stack are assigned.
- Disks belonging to the surviving site are assigned to pool 1 and disks belonging to the disaster site are assigned to pool 0.

Steps

1. Assign the new, unowned drives based on whether you have a four-node or two-node MetroCluster configuration:
   - For four-node MetroCluster configurations, assign the new, unowned disks to the appropriate disk pools by using the following series of commands on the replacement nodes:
     i. Systematically assign the replaced disks for each node to their respective disk pools:

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

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

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

The following example shows the commands with the system IDs:

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

ii. Confirm the ownership of the disks:

```
storage disk show -fields owner, pool
```

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

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

     ```
     run local storage show shelf
     ```

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

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

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

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

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

```
storage disk option modify -autoassign on *
```

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

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

   `storage disk option show`

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

 cluster_B::>
```

Related information

Disk and aggregate management

How MetroCluster configurations use SyncMirror to provide data redundancy

Performing aggregate healing and restoring mirrors (MetroCluster FC configurations)

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

Steps

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

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

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

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

   `storage aggregate show -node local`

If the aggregate shows this state…	Then…
resyncing	No action is required. Let the aggregate complete resyncing.

mirror degraded	Proceed to If one or more plexes remain offline, additional steps are required to rebuild the mirror.
mirrored, normal	No action is required.
unknown, offline	The root aggregate shows this state if all the disks on the disaster sites were replaced.

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

 Aggregate     Size Available Used% State   #Vols  Nodes      RAID
Status
 --------- -------- --------- ----- ------- ------ ----------
------------
 node_B_1_aggr1
           227.1GB   11.00GB   95% online       1 node_B_1   raid_dp,
                                                 resyncing
 NodeA_1_aggr2
           430.3GB   28.02GB   93% online       2 node_B_1   raid_dp,
                                                 mirror
                                                 degraded
 node_B_1_aggr3
           812.8GB   85.37GB   89% online       5 node_B_1   raid_dp,
                                                 mirrored,
                                                 normal
 3 entries were displayed.

cluster_B::>
```

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

Node	State
node_B_1_aggr1	resyncing
node_B_1_aggr2	mirror degraded
node_B_1_aggr3	mirrored, normal

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

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

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

```
storage aggregate show -r -aggregate node_B_1_aggr2
```

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

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

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

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

 cluster_B::>
```

b. Delete the failed plex:

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

c. Reestablish the mirror:

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

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

```
storage aggregate show
```

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

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

About this task

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

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

- Site A is the disaster site.
  - node_A_1 has been replaced.
  - node_A_2 has been replaced.

    Present only in four-node MetroCluster configurations.
- Site B is the surviving site.
  - node_B_1 is healthy.
  - node_B_2 is healthy.

    Present only in four-node MetroCluster configurations.

The old and new system IDs were identified in Replace hardware and boot new controllers.

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

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

Two	node_A_1	4068741258	1574774970

Steps

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

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

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

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

   `disk show`

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

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

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

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

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

3. View the aggregate status:

   ```
   aggr status
   ```

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

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

4. Delete the contents of the mailbox disks:

   ```
   mailbox destroy local
   ```

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

   ```
   aggr online aggr_name
   ```

6. Halt the node to display the LOADER prompt:

   ```
   halt
   ```

Booting the new controller modules (MetroCluster FC configurations)

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

About this task

This task begins with the nodes showing the LOADER prompt.

Steps

1. Display the boot menu:

   ```
   boot_ontap menu
   ```

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

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

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

   > If you did not clear the NVRAM contents of a used replacement controller, then you might see a panic with the following message: `PANIC: NVRAM contents are invalid…` If this occurs, repeat From the boot menu, select option 6, **Update flash from backup config**. to boot the system to the ONTAP prompt. You then need to Reset the boot recovery and rdb_corrupt bootargs

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

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

   c. Assign unowned disks to pool0 on the local node
5. If you have a four-node configuration, repeat the previous steps on the other node at the disaster site.
6. Refresh the MetroCluster configuration:
   a. Enter advanced privilege mode:

      ```
      set -privilege advanced
      ```

   b. Refresh the configuration:

```
metrocluster configure -refresh true
```

c. Return to admin privilege mode:

```
set -privilege admin
```

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

```
metrocluster node show
```

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

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

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

cluster_B::>
```

What to do next

Proceed to Complete the disaster recovery process.

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
```

Preparing for switchback in a mixed configuration (recovery during transition)

You must perform certain tasks in order to prepare the mixed MetroCluster IP and FC configuration for the switchback operation. This procedure only applies to configurations that encountered a failure during the MetroCluster FC to IP transition process.

About this task

This procedure should only be used when performing recovery on a system that was in mid-transition when the failure occurred.

In this scenario, the MetroCluster is a mixed configuration:

- One DR group consists of fabric-attached MetroCluster FC nodes.

  You must perform the MetroCluster FC recovery steps on these nodes.

- One DR group consists of MetroCluster IP nodes.

  You must perform the MetroCluster IP recovery steps on these nodes.

Steps

Perform the steps in the following order.

1. Prepare the FC nodes for switchback by performing the following tasks in order:
   a. Verifying port configuration (MetroCluster FC configurations only)
   b. Configuring the FC-to-SAS bridges (MetroCluster FC configurations only)
   c. Configuring the FC switches (MetroCluster FC configurations only)
   d. Verifying the storage configuration (only perform these steps on replaced drives on the MetroCluster FC nodes)
   e. Powering on the equipment at the disaster site (only perform these steps on replaced drives on the MetroCluster FC nodes)
   f. Assigning ownership for replaced drives (only perform these steps on replaced drives on the MetroCluster FC nodes)
   g. Perform the steps in Reassigning disk ownership for root aggregates to replacement controller modules (MetroCluster FC configurations), up to and including the step to issue the mailbox destroy command.

h. Destroy the local plex (plex 0) of the root aggregate:

```
aggr destroy plex-id
```

i. If the root aggr is not online, bring it online.

2. Boot the MetroCluster FC nodes.

You must perform these steps on both of the MetroCluster FC nodes.

a. Display the boot menu:

```
boot_ontap menu
```

b. From the boot menu, select option 6, **Update flash from backup config**.

c. Respond `y` to the following prompt:

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

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

If you did not clear the NVRAM contents of a used replacement controller, then you might see a panic with the following message: `PANIC: NVRAM contents are invalid...` If this occurs, repeat these substeps to boot the system to the ONTAP prompt. You then need to Reset the boot recovery and rdb_corrupt bootargs

3. Mirror the root aggregate on plex 0:

You must perform these steps on both of the MetroCluster FC nodes.

a. Assign three pool0 disks to the new controller module.

b. Mirror the root aggregate pool1 plex:

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

c. Assign unowned disks to pool0 on the local node

4. Return to Maintenance mode.

You must perform these steps on both of the MetroCluster FC nodes.

a. Halt the node:

```
halt
```

b. Boot the node to Maintenance:

```
mode:boot_ontap maint
```

5. Delete the contents of the mailbox disks:

```
mailbox destroy local
```

You must perform these steps on both of the MetroCluster FC nodes.

6. Halt the nodes:
   `halt`
7. After the nodes boot up, verify the status of the node:

   `metrocluster node show`

```
siteA::*> metrocluster node show
DR                               Configuration  DR
Group Cluster Node               State          Mirroring Mode
----- ------- ------------------ -------------- ---------
--------------------
1     siteA
              wmc66-a1           configured     enabled   waiting for
switchback recovery
              wmc66-a2           configured     enabled   waiting for
switchback recovery
      siteB
              wmc66-b1           configured     enabled   switchover
completed
              wmc66-b2           configured     enabled   switchover
completed
2     siteA
              wmc55-a1           -              -         -
              wmc55-a2           unreachable    -         -
      siteB
              wmc55-b1           configured     enabled   switchover
completed
              wmc55-b2           configured
```

8. Prepare the MetroCluster IP nodes for switchback by performing the tasks in Preparing for switchback in a MetroCluster IP configuration up to and including Deleting failed plexes owned by the surviving site (MetroCluster IP configurations).
9. On the MetroCluster FC nodes, perform the steps in Performing aggregate healing and restoring mirrors (MetroCluster FC configurations).
10. On the MetroCluster IP nodes, perform the steps in Performing aggregate healing and restoring mirrors (MetroCluster IP configurations).
11. Proceed through the remaining tasks of the recovery process beginning with Reestablishing object stores for FabricPool configurations.

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
```

Completing recovery

Perform the required tasks to complete the recovery from a multi-controller or storage failure.

Reestablishing object stores for FabricPool configurations

If one of the object stores in a FabricPool mirror was co-located with the MetroCluster disaster site and was destroyed, you must reestablish the object store and the FabricPool mirror.

About this task

- If the object-stores are remote and a MetroCluster site is destroyed, you do not need to rebuild the object store, and the original object store configurations as well as cold data contents are retained.
- For more information about FabricPool configurations, see the Disk and aggregates management.

Step

1. Follow the procedure "Replacing a FabricPool mirror on a MetroCluster configuration" in the Disk and aggregates management.

Verifying licenses on the replaced nodes

You must install new licenses for the replacement nodes if the impaired nodes were using ONTAP features that require a standard (node-locked) license. For features with standard licenses, each node in the cluster should have its own key for the feature.

About this task

Until you install license keys, features requiring standard licenses continue to be available to the replacement node. However, if the impaired node was the only node in the cluster with a license for the feature, no configuration changes to the feature are allowed. Also, using unlicensed features on the node might put you out of compliance with your license agreement, so you should install the replacement license key or keys on

the replacement node as soon as possible.

The licenses keys must be in the 28-character format.

You have a 90-day grace period in which to install the license keys. After the grace period, all old licenses are invalidated. After a valid license key is installed, you have 24 hours to install all of the keys before the grace period ends.

> If all nodes at a site have been replaced (a single node in the case of a two-node MetroCluster configuration), license keys must be installed on the replacement node or nodes prior to switchback.

Steps

1. Identify the licenses on the node:

   ```
   license show
   ```

   The following example displays the information about licenses in the system:

   ```
   cluster_B::>  license show
            (system license show)

   Serial Number: 1-80-00050
   Owner: site1-01
   Package           Type       Description             Expiration
   -------           -------    -------------           -----------
   Base              license    Cluster Base License        -
   NFS               site       NFS License                 -
   CIFS              site       CIFS License                -
   iSCSI             site       iSCSI License               -
   FCP               site       FCP License                 -
   FlexClone         site       FlexClone License           -

   6 entries were displayed.
   ```

2. Verify that the licenses are good for the node after switchback:

   ```
   metrocluster check license show
   ```

   The following example displays the licenses that are good for the node:

```
cluster_B::> metrocluster check license show

Cluster           Check                              Result
-------           -------                            -------------
Cluster_B         negotiated-switchover-ready        not-applicable
NFS               switchback-ready                   not-applicable
CIFS              job-schedules                      ok
iSCSI             licenses                           ok
FCP               periodic-check-enabled             ok
```

3. If you need new license keys, obtain replacement license keys on the NetApp Support Site in the My Support section under Software licenses.

> The new license keys that you require are automatically generated and sent to the email address on file. If you fail to receive the email with the license keys within 30 days, refer to the *"Who to contact if I have issues with my Licenses?"* section in the Knowledge Base article Post Motherboard Replacement Process to update Licensing on a AFF/FAS system.

4. Install each license key:

```
system license add -license-code license-key, license-key…+
```

5. Remove the old licenses, if desired:
   a. Check for unused licenses:

```
license clean-up -unused -simulate
```

   b. If the list looks correct, remove the unused licenses:

```
license clean-up -unused
```

Restoring key management

If data volumes are encrypted, you must restore key management. If the root volume is encrypted, you must recover key management.

Steps

1. If data volumes are encrypted, restore the keys using the correct command for your key management configuration.

If you are using…	Use this command…
Onboard key management	`security key-manager onboard sync` For more information, see Restoring onboard key management encryption keys.

External key management	`security key-manager key query -node node-name` For more information, see Restoring external key management encryption keys.

2. If the root volume is encrypted, use the procedure in Recovering key management if the root volume is encrypted.

Performing a switchback

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

Before you begin

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

Steps

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

   `metrocluster node show`

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

```
cluster_B::>  metrocluster node show

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

2. Confirm that resynchronization is complete on all SVMs:

   ```
   metrocluster vserver show
   ```

3. Verify that any automatic LIF migrations being performed by the healing operations have been successfully completed:

   ```
   metrocluster check lif show
   ```

4. Perform the switchback by running the `metrocluster switchback` command from any node in the surviving cluster.
5. Check the progress of the switchback operation:

   ```
   metrocluster show
   ```

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

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

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

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

   If a switchback takes a long time to finish, you can check on the status of in-progress baselines by using the the following command at the advanced privilege level:

   ```
   metrocluster config-replication resync-status show
   ```

6. Reestablish any SnapMirror or SnapVault configurations.

   In ONTAP 8.3, you need to manually reestablish a lost SnapMirror configuration after a MetroCluster

switchback operation. In ONTAP 9.0 and later, the relationship is reestablished automatically.

Verifying a successful switchback

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

Steps

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

   ```
   storage aggregate show
   ```

   In the following example, aggr_b2 on node B2 has switched back:

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

   mirrored,

   normal

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

   If the disaster site included unmirrored aggregates and the unmirrored aggregates are no longer present, the aggregate might 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, reference the Knowledge Base article How to remove stale unmirrored aggregate entries in a MetroCluster following disaster where storage was lost.

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

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

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

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

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

   `vserver show -subtype sync-source`

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

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

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

After you finish

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

Mirroring the root aggregates of the replacement nodes

If disks were replaced, you must mirror the root aggregates of the new nodes on the disaster site.

Steps

1. On the disaster site, identify the aggregates which are not mirrored:

   `storage aggregate show`

```
cluster_A::> storage aggregate show

Aggregate     Size Available Used% State   #Vols  Nodes            RAID
Status
--------- -------- --------- ----- ------- ------ ----------------
------------
node_A_1_aggr0
            1.49TB   74.12GB   95% online       1 node_A_1
raid4,

normal
node_A_2_aggr0
            1.49TB   74.12GB   95% online       1 node_A_2
raid4,

normal
node_A_1_aggr1
            1.49TB   74.12GB   95% online       1 node_A_1         raid
4, normal

mirrored
node_A_2_aggr1
            1.49TB   74.12GB   95% online       1 node_A_2         raid
4, normal

mirrored
4 entries were displayed.

cluster_A::>
```

2. Mirror one of the root aggregates:

   `storage aggregate mirror -aggregate root-aggregate`

   The following example shows how the command selects disks and prompts for confirmation when mirroring the aggregate.

```
cluster_A::> storage aggregate mirror -aggregate node_A_2_aggr0

Info: Disks would be added to aggregate "node_A_2_aggr0" on node
"node_A_2" in
      the following manner:

      Second Plex

        RAID Group rg0, 3 disks (block checksum, raid4)
          Position   Disk                      Type
Size
          ---------- ------------------------- ----------
---------------
          parity     2.10.0                    SSD
-
          data       1.11.19                   SSD
894.0GB
          data       2.10.2                    SSD
894.0GB

      Aggregate capacity available for volume use would be 1.49TB.

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

cluster_A::>
```

3. Verify that mirroring of the root aggregate is complete:

   `storage aggregate show`

   The following example shows that the root aggregates are mirrored.

```
cluster_A::> storage aggregate show

Aggregate     Size Available Used% State   #Vols  Nodes       RAID
Status
--------- -------- --------- ----- ------- ------ ----------
------------
node_A_1_aggr0
            1.49TB   74.12GB   95% online       1 node_A_1    raid4,
                                                               mirrored,
                                                               normal
node_A_2_aggr0
            2.24TB   838.5GB   63% online       1 node_A_2    raid4,
                                                               mirrored,
                                                               normal
node_A_1_aggr1
            1.49TB   74.12GB   95% online       1 node_A_1    raid4,
                                                               mirrored,
                                                               normal
node_A_2_aggr1
            1.49TB   74.12GB   95% online       1 node_A_2    raid4
                                                               mirrored,
                                                               normal
4 entries were displayed.

cluster_A::>
```

4. Repeat these steps for the other root aggregates.

   Any root aggregate that does not have a status of mirrored must be mirrored.

Reconfiguring ONTAP Mediator (MetroCluster IP configurations)

If you have a MetroCluster IP configuration that was configured with ONTAP Mediator, you must remove and reconfigure the association with ONTAP Mediator.

Before you begin

- You must have the IP address and username and password for ONTAP Mediator.
- ONTAP Mediator must be configured and operating on the Linux host.

Steps

1. Remove the existing ONTAP Mediator configuration:

   ```
   metrocluster configuration-settings mediator remove
   ```

2. Reconfigure the ONTAP Mediator configuration:

   ```
   metrocluster configuration-settings mediator add -mediator-address mediator-
   ```

```
IP-address
```

Verifying the health of the MetroCluster configuration

You should check the health of the MetroCluster configuration to verify proper operation.

Steps

1. Check that the MetroCluster is configured and in normal mode on each cluster:

   ```
   metrocluster show
   ```

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

2. Check that mirroring is enabled on each node:

   ```
   metrocluster node show
   ```

   ```
   cluster_A::> metrocluster node show
   DR                           Configuration  DR
   Group Cluster Node           State          Mirroring Mode
   ----- ------- -------------- -------------- ---------
   --------------------
   1     cluster_A
                 node_A_1       configured     enabled   normal
         cluster_B
                 node_B_1       configured     enabled   normal
   2 entries were displayed.
   ```

3. Check that the MetroCluster components are healthy:

   ```
   metrocluster check run
   ```

```
cluster_A::> metrocluster check run

Last Checked On: 10/1/2014 16:03:37

Component           Result
------------------- ---------
nodes               ok
lifs                ok
config-replication  ok
aggregates          ok
4 entries were displayed.

Command completed. Use the `metrocluster check show -instance` command
or sub-commands in `metrocluster check` directory for detailed results.
To check if the nodes are ready to do a switchover or switchback
operation, run `metrocluster switchover -simulate` or `metrocluster
switchback -simulate`, respectively.
```

4. Check that there are no health alerts:

   `system health alert show`

5. Simulate a switchover operation:
   a. From any node's prompt, change to the advanced privilege level:

      `set -privilege advanced`

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

   b. Perform the switchover operation with the `-simulate` parameter:

      `metrocluster switchover -simulate`

   c. Return to the admin privilege level:

      `set -privilege admin`

6. For MetroCluster IP configurations using ONTAP Mediator, confirm that ONTAP Mediator is up and operating.
   a. Check that the Mediator disks are visible to the system:

      `storage failover mailbox-disk show`

      The following example shows that the mailbox disks have been recognized.

```
node_A_1::*> storage failover mailbox-disk show
                 Mailbox
Node             Owner     Disk    Name        Disk UUID
-------------     ------   -----   -----        ----------------
sti113-vsim-ucs626g
.
.
     local     0m.i2.3L26
7BBA77C9:AD702D14:831B3E7E:0B0730EE:00000000:00000000:00000000:000000
00:00000000:00000000
     local     0m.i2.3L27
928F79AE:631EA9F9:4DCB5DE6:3402AC48:00000000:00000000:00000000:000000
00:00000000:00000000
     local     0m.i1.0L60
B7BCDB3C:297A4459:318C2748:181565A3:00000000:00000000:00000000:000000
00:00000000:00000000
.
.
.
     partner   0m.i1.0L14
EA71F260:D4DD5F22:E3422387:61D475B2:00000000:00000000:00000000:000000
00:00000000:00000000
     partner   0m.i2.3L64
4460F436:AAE5AB9E:D1ED414E:ABF811F7:00000000:00000000:00000000:000000
00:00000000:00000000
28 entries were displayed.
```

b. Change to the advanced privilege level:

   ```
   set -privilege advanced
   ```

c. Check that the mailbox LUNs are visible to the system:

   ```
   storage iscsi-initiator show
   ```

   The output will show the presence of the mailbox LUNs:

```
Node    Type        Label      Target Portal     Target Name
Admin/Op
----    ----        --------   ---------    ---------
--------------------------------       --------
.
.
.
.node_A_1
               mailbox
                     mediator 172.16.254.1    iqn.2012-
05.local:mailbox.target.db5f02d6-e3d3    up/up
.
.
.
17 entries were displayed.
```

d. Return to the administrative privilege level:

```
set -privilege admin
```

Recovering from a non-controller failure

After the equipment at the disaster site has undergone any required maintenance or replacement, but no controllers were replaced, you can begin the process of returning the MetroCluster configuration to a fully redundant state. This includes healing the configuration (first the data aggregates and then the root aggregates) and performing the switchback operation.

Before you begin

- All MetroCluster hardware in the disaster cluster must be functional.
- The overall MetroCluster configuration must be in switchover.
- In a fabric-attached MetroCluster configuration, the ISL must be up and operating between the MetroCluster sites.

Enable console logging

NetApp strongly recommends that you enable console logging on the devices that you are using and take the following actions when performing this procedure:

- Leave AutoSupport enabled during maintenance.
- Trigger a maintenance AutoSupport message before and after maintenance to disable case creation for the duration of the maintenance activity.

  See the Knowledge Base article How to suppress automatic case creation during scheduled maintenance windows.

- Enable session logging for any CLI session. For instructions on how to enable session logging, review the "Logging Session Output" section in the Knowledge Base article How to configure PuTTY for optimal connectivity to ONTAP systems.

Healing the configuration in a MetroCluster configuration

In MetroCluster FC configurations, you perform the healing operations in a specific order to restore MetroCluster functionality following a switchover.

In MetroCluster IP configurations, healing operations should start automatically following a switchover. If they do not, you can perform the healing operations manually.

Before you begin

- Switchover must have been performed and the surviving site must be serving data.
- Nodes on the disaster site must be halted or remain powered off.

  They must not be fully booted during the healing process.

- Storage at the disaster site must be accessible (shelves are powered up, functional, and accessible).
- In fabric-attached MetroCluster configurations, inter-switch links (ISLs) must be up and operating.
- In four-node MetroCluster configurations, nodes in the surviving site must not be in HA failover state (all nodes must be up and running for each HA pair).

About this task

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

Healing the data aggregates

You must heal the data aggregates after repairing and replacing any hardware on the disaster site. This process resynchronizes the data aggregates and prepares the (now repaired) disaster site for normal operation. You must heal the data aggregates prior to healing the root aggregates.

About this task

The following example shows a forced switchover, where you bring the switched-over aggregate online. All configuration updates in the remote cluster successfully replicate to the local cluster. You power up the storage on the disaster site as part of this procedure, but you do not and must not power up the controller modules on the disaster site.

Steps

1. Verify that switchover was completed:

   ```
   metrocluster operation show
   ```

   ```
   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 following command from the surviving cluster:

   `metrocluster heal -phase aggregates`

   ```
   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:

   `metrocluster operation show`

   ```
   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:

   `storage aggregate show` command.

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

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

Healing the root aggregates after a disaster

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

Before you begin

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

Steps

1. Switch back the mirrored aggregates:

   `metrocluster heal -phase root-aggregates`

   ```
   mcc1A::> metrocluster heal -phase root-aggregates
   [Job 137] Job succeeded: Heal Root Aggregates is successful
   ```

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

2. Ensure that the heal operation is complete by running the following command on the destination cluster:

   `metrocluster operation show`

   ```
   mcc1A::> metrocluster operation show
     Operation: heal-root-aggregates
         State: successful
    Start Time: 7/29/2014 20:54:41
      End Time: 7/29/2014 20:54:42
        Errors: -
   ```

Verifying that your system is ready for a switchback

If your system is already in the switchover state, you can use the `-simulate` option to preview the results of a switchback operation.

Steps

1. Power up each controller module on the disaster site.

   If the nodes are powered off:

   Power on the nodes.

   If the nodes are at the LOADER prompt:

   Run the command: `boot_ontap`

2. After the node boot completes, verify that the root aggregates are mirrored.

If a plex fails:

a. Destroy the failed plex:

```
storage aggregate plex delete -aggregate <aggregate_name> -plex
<plex_name>
```

b. Reestablish the mirror relationship by recreating the mirror:

```
storage aggregate mirror -aggregate <aggregate-name>
```

If a plex is offline:

Online the plex:

```
storage aggregate plex online -aggregate <aggregate_name> -plex <plex_name>
```

If both plexes are present:

Resynchronization starts automatically.

3. Simulate the switchback operation:

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

   ```
   set -privilege advanced
   ```

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

   b. Perform the switchback operation with the `-simulate` parameter:

   ```
   metrocluster switchback -simulate
   ```

   c. Return to the admin privilege level:

   ```
   set -privilege admin
   ```

4. Review the output that is returned.

   The output shows whether the switchback operation would run into errors.

Example of verification results

The following example shows the successful verification of a switchback operation:

```
cluster4::*> metrocluster switchback -simulate
  (metrocluster switchback)
[Job 130] Setting up the nodes and cluster components for the switchback
operation...DBG:backup_api.c:327:backup_nso_sb_vetocheck : MetroCluster
Switch Back
[Job 130] Job succeeded: Switchback simulation is successful.

cluster4::*> metrocluster op show
  (metrocluster operation show)
  Operation: switchback-simulate
      State: successful
 Start Time: 5/15/2014 16:14:34
   End Time: 5/15/2014 16:15:04
     Errors: -

cluster4::*> job show -name Me*
                            Owning
Job ID Name                 Vserver    Node           State
------ -------------------- ---------- -------------- ----------
130    MetroCluster Switchback
                            cluster4
                                       cluster4-01
                                                      Success
       Description: MetroCluster Switchback Job - Simulation
```

Performing a switchback

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

Before you begin

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

Steps

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

   ```
   metrocluster node show
   ```

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

```
cluster_B::>  metrocluster node show

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

2. Confirm that resynchronization is complete on all SVMs:

   ```
   metrocluster vserver show
   ```

3. Verify that any automatic LIF migrations being performed by the healing operations have been successfully completed:

   ```
   metrocluster check lif show
   ```

4. Perform the switchback by running the following command from any node in the surviving cluster.

   ```
   metrocluster switchback
   ```

5. Check the progress of the switchback operation:

   ```
   metrocluster show
   ```

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

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

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

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

If a switchback takes a long time to finish, you can check on the status of in-progress baselines by using the following command at the advanced privilege level.

```
metrocluster config-replication resync-status show
```

6. Reestablish any SnapMirror or SnapVault configurations.

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

Verifying a successful switchback

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

Steps

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

   ```
   storage aggregate show
   ```

   In the following example, aggr_b2 on node B2 has switched back:

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

mirrored,

normal

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

If the disaster site included unmirrored aggregates and the unmirrored aggregates are no longer present, the aggregate might 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 and reference the Knowledge Base article How to remove stale unmirrored aggregate entries in a MetroCluster following disaster where storage was lost.

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

   `vserver show -subtype sync-destination`

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

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

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

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

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

4. Confirm that the switchback operations succeeded:

```
metrocluster operation show
```

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

After you finish

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

Deleting stale aggregate listings after switchback

In some circumstances after switchback, you might notice the presence of *stale* aggregates. Stale aggregates are aggregates that have been removed from ONTAP, but whose information remains recorded on disk. Stale aggregates are displayed with the `nodeshell aggr status -r` command but not with the `storage aggregate show` command. You can delete these records so that they no longer appear.

About this task

Stale aggregates can occur if you relocated aggregates while the MetroCluster configuration was in switchover. For example:

1. Site A switches over to Site B.
2. You delete the mirroring for an aggregate and relocate the aggregate from node_B_1 to node_B_2 for load balancing.
3. You perform aggregate healing.

At this point a stale aggregate appears on node_B_1, even though the actual aggregate has been deleted from that node. This aggregate appears in the output from the `nodeshell aggr status -r` command. It does not appear in the output of the `storage aggregate show` command.

1. Compare the output of the following commands:

   ```
   storage aggregate show

   run local aggr status -r
   ```

   Stale aggregates appear in the `run local aggr status -r` output but not in the `storage aggregate show` output. For example, the following aggregate might appear in the `run local aggr status -r` output:

   ```
   Aggregate aggr05 (failed, raid_dp, partial) (block checksums)
   Plex /aggr05/plex0 (offline, failed, inactive)
     RAID group /myaggr/plex0/rg0 (partial, block checksums)

    RAID Disk Device  HA  SHELF BAY CHAN Pool Type  RPM  Used (MB/blks)
   Phys (MB/blks)
    --------- ------  ------------ ---- ---- ----  ----- --------------
   --------------
    dparity   FAILED          N/A                        82/ -
    parity    0b.5    0b    -   -   SA:A   0 VMDISK  N/A 82/169472
   88/182040
    data      FAILED          N/A                        82/ -
    data      FAILED          N/A                        82/ -
    data      FAILED          N/A                        82/ -
    data      FAILED          N/A                        82/ -
    data      FAILED          N/A                        82/ -
    data      FAILED          N/A                        82/ -
    Raid group is missing 7 disks.
   ```

2. Remove the stale aggregate:
   a. From either node's prompt, change to the advanced privilege level:

      ```
      set -privilege advanced
      ```

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

   b. Remove the stale aggregate:

      ```
      aggregate remove-stale-record -aggregate aggregate_name
      ```

   c. Return to the admin privilege level:

      ```
      set -privilege admin
      ```

3. Confirm that the stale aggregate record was removed:

   ```
   run local aggr status -r
   ```