



Expanding a four-node MetroCluster FC configuration to an eight-node configuration

ONTAP MetroCluster

NetApp
June 14, 2021

Table of Contents

- Expanding a four-node MetroCluster FC configuration to an eight-node configuration 1
 - Determining the new cabling layout 2
 - Racking the new equipment 3
 - Verifying the health of the MetroCluster configuration 3
 - Checking for MetroCluster configuration errors with Config Advisor 4
 - Sending a custom AutoSupport message prior to adding nodes to the MetroCluster configuration 5
 - Recabling and zoning a switch fabric for the new nodes 5
 - Configuring ONTAP on the new controllers 7
 - Checking for MetroCluster configuration errors with Config Advisor 34
 - Sending a custom AutoSupport message after to adding nodes to the MetroCluster configuration 34
 - Verifying switchover, healing, and switchback 35

Expanding a four-node MetroCluster FC configuration to an eight-node configuration

Expanding a four-node MetroCluster FC configuration to an eight-node MetroCluster FC configuration involves adding two controllers to each cluster to form a second HA pair at each MetroCluster site, and then running the MetroCluster FC configuration operation.

- The nodes must be running ONTAP 9 in a MetroCluster FC configuration.

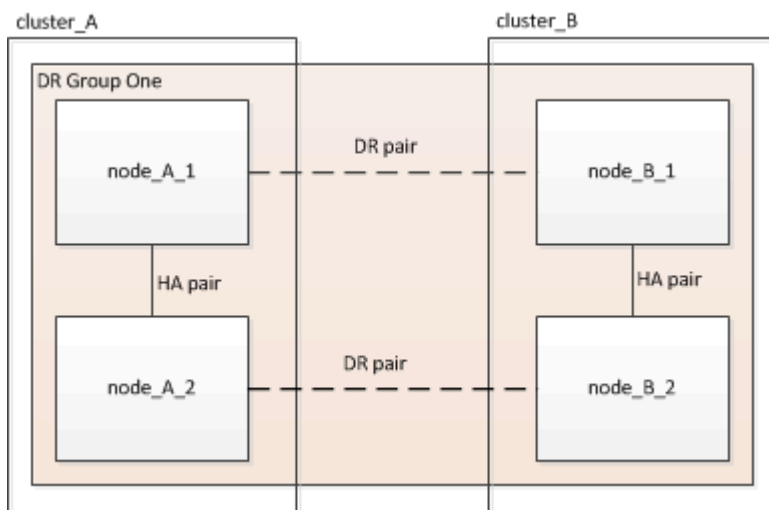
This procedure is not supported on earlier versions of ONTAP or in MetroCluster IP configurations.

- The existing MetroCluster FC configuration must be healthy.
- The equipment you are adding must be supported and meet all the requirements described in the *Fabric-attached MetroCluster Installation and Configuration Guide*.

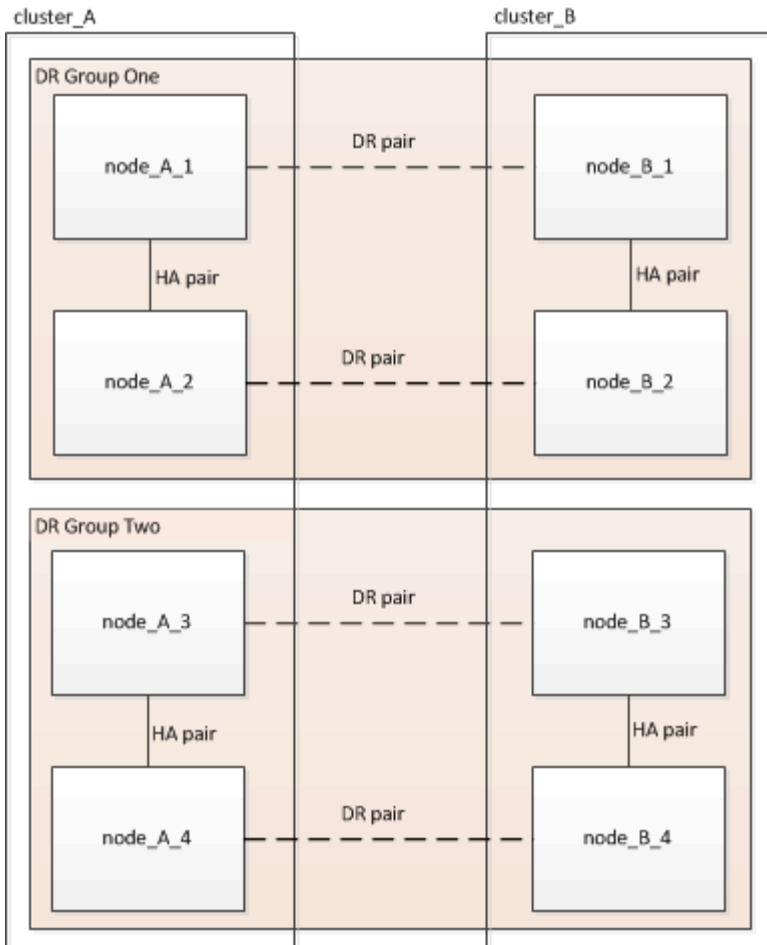
[Fabric-attached MetroCluster installation and configuration](#)

- You must have available FC switch ports to accommodate the new controllers and any new bridges.
- You need the admin password and access to an FTP or SCP server.
- This procedure applies only to MetroCluster FC configurations.
- This procedure is nondisruptive and takes approximately one day to complete (excluding rack and stack) when disks are zeroed.

Before performing this procedure, the MetroCluster FC configuration consists of four nodes, with one HA pair at each site:



At the conclusion of this procedure, the MetroCluster FC configuration consists of two HA pairs at each site:



Both sites must be expanded equally. A MetroCluster FC configuration cannot consist of an uneven number of nodes.

Determining the new cabling layout

You must determine the cabling for the new controller modules and any new disk shelves to the existing FC switches.

This task must be performed at each MetroCluster site.

Steps

1. Use the *Fabric-attached MetroCluster Installation and Configuration Guide* and create a cabling layout for your switch type, using the port usage for an eight-node MetroCluster configuration.

The FC switch port usage must match the usage described in the guide so that the Reference Configuration Files (RCFs) can be used.

[Fabric-attached MetroCluster installation and configuration](#)



If your environment cannot be cabled in such a way that RCF files can be used, you must manually configure the system according to instructions found in the *Fabric-attached MetroCluster Installation and Configuration Guide*. Do not use this procedure if the cabling cannot use RCF files.

Racking the new equipment

You must rack the equipment for the new nodes.

Steps

1. Use the MetroCluster Installation and Configuration guide and rack the new storage systems, disk shelves, and FC-to-SAS bridges.

[Fabric-attached MetroCluster installation and configuration](#)

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:

- 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 (*>).

- Perform the switchover operation with the -simulate parameter:

```
metrocluster switchover -simulate
```

- Return to the admin privilege level:

```
set -privilege admin
```

Checking for MetroCluster configuration errors with Config Advisor

You can go to the NetApp Support Site and download the Config Advisor tool to check for common configuration errors.

Config Advisor is a configuration validation and health check tool. You can deploy it at both secure sites and non-secure sites for data collection and system analysis.



Support for Config Advisor is limited, and available only online.

Steps

1. Go to the Config Advisor download page and download the tool.

[NetApp Downloads: Config Advisor](#)

2. Run Config Advisor, review the tool's output and follow the recommendations in the output to address any issues discovered.

Sending a custom AutoSupport message prior to adding nodes to the MetroCluster configuration

You should issue an AutoSupport message to notify NetApp technical support that maintenance is underway. Informing technical support that maintenance is underway prevents them from opening a case on the assumption that a disruption has occurred.

This task must be performed on each MetroCluster site.

Steps

1. Log in to the cluster at Site_A.
2. Invoke an AutoSupport message indicating the start of the maintenance:

```
system node autosupport invoke -node * -type all -message MAINT=maintenance-  
window-in-hours
```

`maintenance-window-in-hours` specifies the length of the maintenance window and can be a maximum of 72 hours. If the maintenance is completed before the time has elapsed, you can issue a command to indicating that the maintenance period has ended:

```
system node autosupport invoke -node * -type all -message MAINT=end
```

3. Repeat this step on the partner site.

Recabling and zoning a switch fabric for the new nodes

When adding nodes to the MetroCluster configuration, you must change the cabling and then run RCF files to redefine the zoning on the fabric.

This task must be performed on each switch fabric. It is done one fabric at a time.

Disconnecting the existing DR group from the fabric

You must disconnect the existing controller modules from the FC switches in the fabric.

This task must be performed at each MetroCluster site.

Steps

1. Disable the HBA ports that connect the existing controller modules to the switch fabric undergoing maintenance:

```
storage port disable -node node-name -port port-number
```

2. On the local FC switches, remove the cables from the ports for the existing controller module's HBA, FC-VI, and ATTO bridges.

You should label the cables for easy identification when you recable them. Only the ISL ports should remain cabled.

Applying the RCF files and recabling the switches

You must apply the RCF files to reconfigure your zoning to accommodate the new nodes.

Steps

1. Locate the RCF files for your configuration.

You must use the RCF files for an eight-node configuration and that match your switch model.

2. Apply the RCF files, following the directions on the download page, adjusting the ISL settings as needed.
3. Ensure that the switch configuration is saved.
4. Reboot the FC switches.
5. Cable both the pre-existing and the new FC-to-SAS bridges to the FC switches, using the cabling layout you created previously.

The FC switch port usage must match the MetroCluster eight-node usage described in the *Fabric-attached MetroCluster Installation and Configuration Guide* so that the Reference Configuration Files (RCFs) can be used.

[Fabric-attached MetroCluster installation and configuration](#)



If your environment cannot be cabled in such a way that RCF files can be used then contact technical support. Do NOT use this procedure if the cabling cannot use RCF files.

6. Verify that the ports are online by using the correct command for your switch.

Switch vendor	Command
Brocade	switchshow
Cisco	show interface brief

7. Cable the FC-VI ports from the existing and new controllers, using the cabling layout you created previously.

[Fabric-attached MetroCluster installation and configuration](#)

The FC switch port usage must match the MetroCluster eight-node usage described in the *Fabric-attached MetroCluster Installation and Configuration Guide* so that the Reference Configuration Files (RCFs) can be used.



If your environment cannot be cabled in such a way that RCF files can be used then contact technical support. Do NOT use this procedure if the cabling cannot use RCF files.

8. From the existing nodes, verify that the FC-VI ports are online:

```
metrocluster interconnect adapter show
```

```
metrocluster interconnect mirror show
```

9. Cable the HBA ports from the current and the new controllers.
10. On the existing controller modules, e-enable the ports connected to the switch fabric undergoing maintenance:

```
storage port enable -node node-name -port port-ID
```

11. Start the new controllers and boot them into Maintenance mode:

```
boot_ontap maint
```

12. Verify that only storage that will be used by the new DR group is visible to the new controller modules.

None of the storage that is used by the other DR group should be visible.

13. Return to the beginning of this process to recable the second switch fabric.

Configuring ONTAP on the new controllers

You must set up ONTAP on each new controller in the MetroCluster configuration, and then re-create the MetroCluster relationship between the two sites.

Clearing the configuration on a controller module

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

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` `bye`
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.

Assigning disk ownership in AFF systems

If you are using AFF systems in a configuration with mirrored aggregates and the nodes do not have the disks (SSDs) correctly assigned, you should assign half the disks on each shelf to one local node and the other half of the disks to its HA partner node. You should create a configuration in which each node has the same number of disks in its local and remote disk pools.

The storage controllers must be in Maintenance mode.

This does not apply to configurations which have unmirrored aggregates, an active/passive configuration, or that have an unequal number of disks in local and remote pools.

This task is not required if disks were correctly assigned when received from the factory.



Pool 0 always contains the disks that are found at the same site as the storage system that owns them, while Pool 1 always contains the disks that are remote to the storage system that owns them.

Steps

1. If you have not done so, boot each system into Maintenance mode.
2. Assign the disks to the nodes located at the first site (site A):

You should assign an equal number of disks to each pool.

- a. On the first node, systematically assign half the disks on each shelf to pool 0 and the other half to the HA partner's pool 0:

```
disk assign -disk disk-name -p pool -n number-of-disks
```

If storage controller Controller_A_1 has four shelves, each with 8 SSDs, you issue the following commands:

```
*> disk assign -shelf FC_switch_A_1:1-4.shelf1 -p 0 -n 4
*> disk assign -shelf FC_switch_A_1:1-4.shelf2 -p 0 -n 4

*> disk assign -shelf FC_switch_B_1:1-4.shelf1 -p 1 -n 4
*> disk assign -shelf FC_switch_B_1:1-4.shelf2 -p 1 -n 4
```

- b. Repeat the process for the second node at the local site, systematically assigning half the disks on each shelf to pool 1 and the other half to the HA partner's pool 1:

```
disk assign -disk disk-name -p pool
```

If storage controller Controller_A_1 has four shelves, each with 8 SSDs, you issue the following commands:

```
*> disk assign -shelf FC_switch_A_1:1-4.shelf3 -p 0 -n 4
*> disk assign -shelf FC_switch_B_1:1-4.shelf4 -p 1 -n 4

*> disk assign -shelf FC_switch_A_1:1-4.shelf3 -p 0 -n 4
*> disk assign -shelf FC_switch_B_1:1-4.shelf4 -p 1 -n 4
```

3. Assign the disks to the nodes located at the second site (site B):

You should assign an equal number of disks to each pool.

- a. On the first node at the remote site, systematically assign half the disks on each shelf to pool 0 and the other half to the HA partner's pool 0:

```
disk assign -disk disk-name -p pool
```

If storage controller Controller_B_1 has four shelves, each with 8 SSDs, you issue the following commands:

```
*> disk assign -shelf FC_switch_B_1:1-5.shelf1 -p 0 -n 4
*> disk assign -shelf FC_switch_B_1:1-5.shelf2 -p 0 -n 4

*> disk assign -shelf FC_switch_A_1:1-5.shelf1 -p 1 -n 4
*> disk assign -shelf FC_switch_A_1:1-5.shelf2 -p 1 -n 4
```

- b. Repeat the process for the second node at the remote site, systematically assigning half the disks on each shelf to pool 1 and the other half to the HA partner's pool 1:

```
disk assign -disk disk-name -p pool
```

If storage controller Controller_B_2 has four shelves, each with 8 SSDs, you issue the following commands:

```
*> disk assign -shelf FC_switch_B_1:1-5.shelf3 -p 0 -n 4
*> disk assign -shelf FC_switch_B_1:1-5.shelf4 -p 0 -n 4

*> disk assign -shelf FC_switch_A_1:1-5.shelf3 -p 1 -n 4
*> disk assign -shelf FC_switch_A_1:1-5.shelf4 -p 1 -n 4
```

4. Confirm the disk assignments:

```
storage show disk
```

5. Exit Maintenance mode:

```
halt
```

6. Display the boot menu:

```
boot_ontap menu
```

7. On each node, select option 4 to initialize all disks.

Assigning disk ownership in non-AFF systems

If the MetroCluster nodes do not have the disks correctly assigned, or if you are using DS460C disk shelves in your configuration, you must assign disks to each of the nodes in the MetroCluster configuration on a shelf-by-shelf basis. You will create a configuration in which each node has the same number of disks in its local and remote disk pools.

The storage controllers must be in Maintenance mode.

If your configuration does not include DS460C disk shelves, this task is not required if disks were correctly assigned when received from the factory.



Pool 0 always contains the disks that are found at the same site as the storage system that owns them.

Pool 1 always contains the disks that are remote to the storage system that owns them.

If your configuration includes DS460C disk shelves, you should manually assign the disks using the following guidelines for each 12-disk drawer:

Assign these disks in the drawer...	To this node and pool...
0 - 2	Local node's pool 0
3 - 5	HA partner node's pool 0
6 - 8	DR partner of the local node's pool 1
9 - 11	DR partner of the HA partner's pool 1

This disk assignment pattern ensures that an aggregate is minimally affected in case a drawer goes offline.

Steps

1. If you have not done so, boot each system into Maintenance mode.
2. Assign the disk shelves to the nodes located at the first site (site A):

Disk shelves at the same site as the node are assigned to pool 0 and disk shelves located at the partner site are assigned to pool 1.

You should assign an equal number of shelves to each pool.

- a. On the first node, systematically assign the local disk shelves to pool 0 and the remote disk shelves to pool 1:

```
disk assign -shelf local-switch-name:shelf-name.port -p pool
```

If storage controller Controller_A_1 has four shelves, you issue the following commands:

```
*> disk assign -shelf FC_switch_A_1:1-4.shelf1 -p 0
*> disk assign -shelf FC_switch_A_1:1-4.shelf2 -p 0

*> disk assign -shelf FC_switch_B_1:1-4.shelf1 -p 1
*> disk assign -shelf FC_switch_B_1:1-4.shelf2 -p 1
```

- b. Repeat the process for the second node at the local site, systematically assigning the local disk shelves to pool 0 and the remote disk shelves to pool 1:

`disk assign -shelf local-switch-name:shelf-name.port -p pool`

If storage controller Controller_A_2 has four shelves, you issue the following commands:

```
*> disk assign -shelf FC_switch_A_1:1-4.shelf3 -p 0
*> disk assign -shelf FC_switch_B_1:1-4.shelf4 -p 1

*> disk assign -shelf FC_switch_A_1:1-4.shelf3 -p 0
*> disk assign -shelf FC_switch_B_1:1-4.shelf4 -p 1
```

3. Assign the disk shelves to the nodes located at the second site (site B):

Disk shelves at the same site as the node are assigned to pool 0 and disk shelves located at the partner site are assigned to pool 1.

You should assign an equal number of shelves to each pool.

- a. On the first node at the remote site, systematically assign its local disk shelves to pool 0 and its remote disk shelves to pool 1:

`disk assign -shelf local-switch-namesshelf-name -p pool`

If storage controller Controller_B_1 has four shelves, you issue the following commands:

```
*> disk assign -shelf FC_switch_B_1:1-5.shelf1 -p 0
*> disk assign -shelf FC_switch_B_1:1-5.shelf2 -p 0

*> disk assign -shelf FC_switch_A_1:1-5.shelf1 -p 1
*> disk assign -shelf FC_switch_A_1:1-5.shelf2 -p 1
```

- b. Repeat the process for the second node at the remote site, systematically assigning its local disk shelves to pool 0 and its remote disk shelves to pool 1:

`disk assign -shelf shelf-name -p pool`

If storage controller Controller_B_2 has four shelves, you issue the following commands:

```
*> disk assign -shelf FC_switch_B_1:1-5.shelf3 -p 0
*> disk assign -shelf FC_switch_B_1:1-5.shelf4 -p 0

*> disk assign -shelf FC_switch_A_1:1-5.shelf3 -p 1
*> disk assign -shelf FC_switch_A_1:1-5.shelf4 -p 1
```

4. Confirm the shelf assignments:

```
storage show shelf
```

5. Exit Maintenance mode:

```
halt
```

6. Display the boot menu:

```
boot_ontap menu
```

7. On each node, select option **4** to initialize all disks.

Verifying 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** so they boot up properly.

The system must be in Maintenance mode.

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 HA state for all components should be **mcc**.

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

```
ha-config modify controller mcc
```

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

```
ha-config modify chassis mcc
```

4. Repeat these steps on the other replacement node.

Booting the new controllers and joining them to the cluster

To join the new controllers to the cluster, you must boot each new controller module and use the ONTAP cluster setup wizard to identify the cluster will join.

You must have cabled the MetroCluster configuration.

You must not have configured the Service Processor prior to performing this task.

This task must be performed on each of the new controllers at both clusters in the MetroCluster configuration.

Steps

1. If you have not already done so, power up each node and let them boot completely.

If the system is in Maintenance mode, issue the **halt** command to exit Maintenance mode, and then issue the following command from the LOADER prompt:

```
boot_ontap
```

The controller module enters the node setup wizard.

The output should be similar to the following:

```
Welcome to node setup

You can enter the following commands at any time:
  "help" or "?" - if you want to have a question clarified,
  "back" - if you want to change previously answered questions, and
  "exit" or "quit" - if you want to quit the setup wizard.
                Any changes you made before quitting will be saved.

To accept a default or omit a question, do not enter a value.
.
.
.
```

2. Enable the AutoSupport tool by following the directions provided by the system.
3. Respond to the prompts to configure the node management interface.

The prompts are similar to the following:

```
Enter the node management interface port: [e0M]:
Enter the node management interface IP address: 10.228.160.229
Enter the node management interface netmask: 225.225.252.0
Enter the node management interface default gateway: 10.228.160.1
```

4. Confirm that nodes are configured in high-availability mode:

```
storage failover show -fields mode
```

If not, you must issue the following command on each node, and then reboot the node:

```
storage failover modify -mode ha -node localhost
```

This command configures high availability mode but does not enable storage failover. Storage failover is automatically enabled when you issue the `metrocluster configure` command later in the configuration process.

5. Confirm that you have four ports configured as cluster interconnects:

`network port show`

The following example shows output for two controllers in cluster_A. If it is a two-node MetroCluster configuration, the output shows only one node.

```
cluster_A::> network port show

(Mbps)
Node   Port           IPspace      Broadcast Domain Link  MTU  Admin/Oper
-----
node_A_1
  **e0a      Cluster      Cluster      up    1500
  auto/1000
  e0b        Cluster      Cluster      up    1500
  auto/1000**
  e0c        Default      Default      up    1500  auto/1000
  e0d        Default      Default      up    1500  auto/1000
  e0e        Default      Default      up    1500  auto/1000
  e0f        Default      Default      up    1500  auto/1000
  e0g        Default      Default      up    1500  auto/1000
node_A_2
  **e0a      Cluster      Cluster      up    1500
  auto/1000
  e0b        Cluster      Cluster      up    1500
  auto/1000**
  e0c        Default      Default      up    1500  auto/1000
  e0d        Default      Default      up    1500  auto/1000
  e0e        Default      Default      up    1500  auto/1000
  e0f        Default      Default      up    1500  auto/1000
  e0g        Default      Default      up    1500  auto/1000
14 entries were displayed.
```

6. Because you are using the CLI to set up the cluster, exit the Node Setup wizard:

`exit`

7. Log in to the admin account by using the `admin` user name.
8. Start the Cluster Setup wizard, and then join the existing cluster: `cluster setup`


```
::> cluster setup
```

Welcome to the cluster setup wizard.

You can enter the following commands at any time:

"help" or "?" - if you want to have a question clarified,
"back" - if you want to change previously answered questions, and
"exit" or "quit" - if you want to quit the cluster setup wizard.
Any changes you made before quitting will be saved.

You can return to cluster setup at any time by typing "cluster setup".
To accept a default or omit a question, do not enter a value.

Do you want to create a new cluster or join an existing cluster?
{create, join}:`join`

9. After you complete the **Cluster Setup** wizard and it exits, verify that the cluster is active and the node is healthy:

cluster show

The following example shows a cluster in which the first node (cluster1-01) is healthy and eligible to participate:

```
cluster_A::> cluster show
Node           Health Eligibility
-----
node_A_1       true   true
node_A_2       true   true
node_A_3       true   true
```

If it becomes necessary to change any of the settings you entered for the admin SVM or node SVM, you can access the **Cluster Setup** wizard by using the **cluster setup command**.

Configuring the clusters into a MetroCluster configuration

You must peer the clusters, mirror the root aggregates, create a mirrored data aggregate, and then issue the command to implement the MetroCluster operations.

Configuring intercluster LIFs

You must create intercluster LIFs on ports used for communication between the MetroCluster partner clusters. You can use dedicated ports or ports that also have data traffic.

Configuring intercluster LIFs on dedicated ports

You can configure intercluster LIFs on dedicated ports. Doing so typically increases the available bandwidth for replication traffic.

Steps

1. List the ports in the cluster:

```
network port show
```

For complete command syntax, see the man page.

The following example shows the network ports in cluster01:

```
cluster01::> network port show
```

							Speed
(Mbps)							
Node	Port	IPspace	Broadcast Domain	Link	MTU	Admin/Oper	
-----	-----	-----	-----	-----	-----	-----	
cluster01-01							
	e0a	Cluster	Cluster	up	1500	auto/1000	
	e0b	Cluster	Cluster	up	1500	auto/1000	
	e0c	Default	Default	up	1500	auto/1000	
	e0d	Default	Default	up	1500	auto/1000	
	e0e	Default	Default	up	1500	auto/1000	
	e0f	Default	Default	up	1500	auto/1000	
cluster01-02							
	e0a	Cluster	Cluster	up	1500	auto/1000	
	e0b	Cluster	Cluster	up	1500	auto/1000	
	e0c	Default	Default	up	1500	auto/1000	
	e0d	Default	Default	up	1500	auto/1000	
	e0e	Default	Default	up	1500	auto/1000	
	e0f	Default	Default	up	1500	auto/1000	

2. Determine which ports are available to dedicate to intercluster communication:

```
network interface show -fields home-port,curr-port
```

For complete command syntax, see the man page.

The following example shows that ports e0e and e0f have not been assigned LIFs:

```

cluster01::> network interface show -fields home-port,curr-port
vserver lif                home-port curr-port
-----
Cluster cluster01-01_clus1  e0a       e0a
Cluster cluster01-01_clus2  e0b       e0b
Cluster cluster01-02_clus1  e0a       e0a
Cluster cluster01-02_clus2  e0b       e0b
cluster01
      cluster_mgmt           e0c       e0c
cluster01
      cluster01-01_mgmt1     e0c       e0c
cluster01
      cluster01-02_mgmt1     e0c       e0c

```

3. Create a failover group for the dedicated ports:

```
network interface failover-groups create -vserver system_SVM -failover-group failover_group -targets physical_or_logical_ports
```

The following example assigns ports e0e and e0f to the failover group intercluster01 on the system SVMcluster01:

```

cluster01::> network interface failover-groups create -vserver cluster01
-failover-group
intercluster01 -targets
cluster01-01:e0e,cluster01-01:e0f,cluster01-02:e0e,cluster01-02:e0f

```

4. Verify that the failover group was created:

```
network interface failover-groups show
```

For complete command syntax, see the man page.

```

cluster01::> network interface failover-groups show

```

Vserver	Group	Failover Targets
Cluster	Cluster	cluster01-01:e0a, cluster01-01:e0b, cluster01-02:e0a, cluster01-02:e0b
cluster01	Default	cluster01-01:e0c, cluster01-01:e0d, cluster01-02:e0c, cluster01-02:e0d, cluster01-01:e0e, cluster01-01:e0f cluster01-02:e0e, cluster01-02:e0f
	intercluster01	cluster01-01:e0e, cluster01-01:e0f cluster01-02:e0e, cluster01-02:e0f

5. Create intercluster LIFs on the system SVM and assign them to the failover group.

ONTAP version	Command
9.6 and later	<code>network interface create -vserver system_SVM -lif LIF_name -service-policy default-intercluster -home-node node -home-port port -address port_IP -netmask netmask -failover-group failover_group</code>
9.5 and earlier	<code>network interface create -vserver system_SVM -lif LIF_name -role intercluster -home-node node -home-port port -address port_IP -netmask netmask -failover-group failover_group</code>

For complete command syntax, see the man page.

The following example creates intercluster LIFs `cluster01_icl01` and `cluster01_icl02` in the failover group `intercluster01`:

```

cluster01::> network interface create -vserver cluster01 -lif
cluster01_icl01 -service-
policy default-intercluster -home-node cluster01-01 -home-port e0e
-address 192.168.1.201
-netmask 255.255.255.0 -failover-group intercluster01

cluster01::> network interface create -vserver cluster01 -lif
cluster01_icl02 -service-
policy default-intercluster -home-node cluster01-02 -home-port e0e
-address 192.168.1.202
-netmask 255.255.255.0 -failover-group intercluster01

```

6. Verify that the intercluster LIFs were created:

In ONTAP 9.6 and later:

```
network interface show -service-policy default-intercluster
```

In ONTAP 9.5 and earlier:

```
network interface show -role intercluster
```

For complete command syntax, see the man page.

```

cluster01::> network interface show -service-policy default-intercluster

```

Current Is	Logical	Status	Network	Current
Vserver	Interface	Admin/Oper	Address/Mask	Node
Home				Port
cluster01	cluster01_icl01	up/up	192.168.1.201/24	cluster01-01 e0e
true				
	cluster01_icl02	up/up	192.168.1.202/24	cluster01-02 e0f
true				

7. Verify that the intercluster LIFs are redundant:

In ONTAP 9.6 and later:

```
network interface show -service-policy default-intercluster -failover
```

In ONTAP 9.5 and earlier:

```
network interface show -role intercluster -failover
```

For complete command syntax, see the man page.

The following example shows that the intercluster LIFs `cluster01_icl01` and `cluster01_icl02` on the SVMe0e port will fail over to the e0f port.

```
cluster01::> network interface show -service-policy default-intercluster
-failover
          Logical          Home          Failover          Failover
Vserver  Interface          Node:Port        Policy            Group
-----  -
cluster01
          cluster01_icl01 cluster01-01:e0e  local-only
intercluster01
          Failover Targets: cluster01-01:e0e,
                           cluster01-01:e0f
          cluster01_icl02 cluster01-02:e0e  local-only
intercluster01
          Failover Targets: cluster01-02:e0e,
                           cluster01-02:e0f
```

Configuring intercluster LIFs on shared data ports

You can configure intercluster LIFs on ports shared with the data network. Doing so reduces the number of ports you need for intercluster networking.

Steps

1. List the ports in the cluster:

```
network port show
```

For complete command syntax, see the man page.

The following example shows the network ports in cluster01:

```
cluster01::> network port show
```

(Mbps)							Speed
Node	Port	IPspace	Broadcast Domain	Link	MTU	Admin/Oper	

cluster01-01							
	e0a	Cluster	Cluster	up	1500	auto/1000	
	e0b	Cluster	Cluster	up	1500	auto/1000	
	e0c	Default	Default	up	1500	auto/1000	
	e0d	Default	Default	up	1500	auto/1000	
cluster01-02							
	e0a	Cluster	Cluster	up	1500	auto/1000	
	e0b	Cluster	Cluster	up	1500	auto/1000	
	e0c	Default	Default	up	1500	auto/1000	
	e0d	Default	Default	up	1500	auto/1000	

2. Create intercluster LIFs on the system SVM:

In ONTAP 9.6 and later:

```
network interface create -vserver system_SVM -lif LIF_name -service-policy default-intercluster -home-node node -home-port port -address port_IP -netmask netmask
```

In ONTAP 9.5 and earlier:

```
network interface create -vserver system_SVM -lif LIF_name -role intercluster -home-node node -home-port port -address port_IP -netmask *netmask
```

For complete command syntax, see the man page.

The following example creates intercluster LIFs `cluster01_icl01` and `cluster01_icl02`:

```

cluster01::> network interface create -vserver cluster01 -lif
cluster01_icl01 -service-
policy default-intercluster -home-node cluster01-01 -home-port e0c
-address 192.168.1.201
-netmask 255.255.255.0

cluster01::> network interface create -vserver cluster01 -lif
cluster01_icl02 -service-
policy default-intercluster -home-node cluster01-02 -home-port e0c
-address 192.168.1.202
-netmask 255.255.255.0

```

3. Verify that the intercluster LIFs were created:

In ONTAP 9.6 and later:

```
network interface show -service-policy default-intercluster
```

In ONTAP 9.5 and earlier:

```
network interface show -role intercluster
```

For complete command syntax, see the man page.

```

cluster01::> network interface show -service-policy default-intercluster

```

Current Is	Logical	Status	Network	Current
Vserver	Interface	Admin/Oper	Address/Mask	Node
Home				Port
cluster01	cluster01_icl01	up/up	192.168.1.201/24	cluster01-01 e0c
true				
	cluster01_icl02	up/up	192.168.1.202/24	cluster01-02 e0c
true				

4. Verify that the intercluster LIFs are redundant:

In ONTAP 9.6 and later:

```
network interface show -service-policy default-intercluster -failover
```


In ONTAP 9.6 and later:

In ONTAP 9.5 and earlier:

```
network interface show -role intercluster -failover
```

For complete command syntax, see the man page.

The following example shows that the intercluster LIFs `cluster01_icl01` and `cluster01_icl02` on the `e0c` port will fail over to the `e0d` port.

```
cluster01::> network interface show -service-policy default-intercluster
-failover
      Logical          Home          Failover          Failover
Vserver Interface      Node:Port        Policy            Group
-----
cluster01
      cluster01_icl01 cluster01-01:e0c  local-only
192.168.1.201/24
                                Failover Targets: cluster01-01:e0c,
                                                cluster01-01:e0d
      cluster01_icl02 cluster01-02:e0c  local-only
192.168.1.201/24
                                Failover Targets: cluster01-02:e0c,
                                                cluster01-02:e0d
```

Mirroring the root aggregates

You must mirror the root aggregates to provide data protection.

By default, the root aggregate is created as RAID-DP type aggregate. You can change the root aggregate from RAID-DP to RAID4 type aggregate. The following command modifies the root aggregate for RAID4 type aggregate:

```
storage aggregate modify -aggregate aggr_name -raidtype raid4
```



On non-ADP systems, the RAID type of the aggregate can be modified from the default RAID-DP to RAID4 before or after the aggregate is mirrored.

Steps

1. Mirror the root aggregate:

```
storage aggregate mirror aggr_name
```

The following command mirrors the root aggregate for controller_A_1:

```
controller_A_1::> storage aggregate mirror aggr0_controller_A_1
```

This mirrors the aggregate, so it consists of a local plex and a remote plex located at the remote MetroCluster site.

2. Repeat the previous step for each node in the MetroCluster configuration.

Implementing the MetroCluster configuration

You must run the `metrocluster configure -refresh true` command to start data protection on the nodes that you have added to a MetroCluster configuration.

You issue the `metrocluster configure -refresh true` command once, on one of the newly added nodes, to refresh the MetroCluster configuration. You do not need to issue the command on each of the sites or nodes.

The `metrocluster configure -refresh true` command automatically pairs the two nodes with the lowest system IDs in each of the two clusters as disaster recovery (DR) partners. In a four-node MetroCluster configuration, there are two DR partner pairs. The second DR pair is created from the two nodes with higher system IDs.

Steps

1. Refresh the MetroCluster configuration:
 - a. Enter advanced privilege mode:
`set -privilege advanced`
 - b. Refresh the MetroCluster configuration on one of the new nodes:
`metrocluster configure -refresh true`

The following example shows the MetroCluster configuration refreshed on both DR groups:

```
controller_A_2::*> metrocluster configure -refresh true  
[Job 726] Job succeeded: Configure is successful.
```

```
controller_A_4::*> metrocluster configure -refresh true  
[Job 740] Job succeeded: Configure is successful.
```

- c. Return to admin privilege mode:
`set -privilege admin`
2. Verify the networking status on site A:

```
network port show
```

The following example shows the network port usage on a four-node MetroCluster configuration:

```
cluster_A::> network port show
```

Node	Port	IPspace	Broadcast Domain	Link	MTU	Speed (Mbps) Admin/Oper

controller_A_1						
	e0a	Cluster	Cluster	up	9000	auto/1000
	e0b	Cluster	Cluster	up	9000	auto/1000
	e0c	Default	Default	up	1500	auto/1000
	e0d	Default	Default	up	1500	auto/1000
	e0e	Default	Default	up	1500	auto/1000
	e0f	Default	Default	up	1500	auto/1000
	e0g	Default	Default	up	1500	auto/1000
controller_A_2						
	e0a	Cluster	Cluster	up	9000	auto/1000
	e0b	Cluster	Cluster	up	9000	auto/1000
	e0c	Default	Default	up	1500	auto/1000
	e0d	Default	Default	up	1500	auto/1000
	e0e	Default	Default	up	1500	auto/1000
	e0f	Default	Default	up	1500	auto/1000
	e0g	Default	Default	up	1500	auto/1000

14 entries were displayed.

3. Verify the MetroCluster configuration from both sites in the MetroCluster configuration:

a. Verify the configuration from site A:

metrocluster show

```
cluster_A::> metrocluster show
```

Configuration: IP fabric

Cluster	Entry Name	State

Local: cluster_A	Configuration state	configured
	Mode	normal
Remote: cluster_B	Configuration state	configured
	Mode	normal

b. Verify the configuration from site B:

metrocluster show

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

```
Configuration: IP fabric
```

Cluster	Entry Name	State
Local: cluster_B	Configuration state	configured
	Mode	normal
Remote: cluster_A	Configuration state	configured
	Mode	normal

Creating a mirrored data aggregate on each node

You must create a mirrored data aggregate on each node in the DR group.

- You should know what drives will be used in the new aggregate.
- If you have multiple drive types in your system (heterogeneous storage), you should understand how you can ensure that the correct drive type is selected.
- Drives are owned by a specific node; when you create an aggregate, all drives in that aggregate must be owned by the same node, which becomes the home node for that aggregate.

In systems using ADP, aggregates are created using partitions in which each drive is partitioned into P1, P2 and P3 partitions.

- Aggregate names should conform to the naming scheme you determined when you planned your MetroCluster configuration.

Disk and aggregate management

1. Display a list of available spares:

```
storage disk show -spare -owner node_name
```

2. Create the aggregate by using the `storage aggregate create -mirror true` command.

If you are logged in to the cluster on the cluster management interface, you can create an aggregate on any node in the cluster. To ensure that the aggregate is created on a specific node, use the `-node` parameter or specify drives that are owned by that node.

You can specify the following options:

- Aggregate's home node (that is, the node that owns the aggregate in normal operation)
- List of specific drives that are to be added to the aggregate
- Number of drives to include



In the minimum supported configuration, in which a limited number of drives are available, you must use the `force-small-aggregate` option to allow the creation of a three disk RAID-DP aggregate.

- Checksum style to use for the aggregate
- Type of drives to use
- Size of drives to use
- Drive speed to use
- RAID type for RAID groups on the aggregate
- Maximum number of drives that can be included in a RAID group
- Whether drives with different RPM are allowed For more information about these options, see the [storage aggregate create](#) man page.

The following command creates a mirrored aggregate with 10 disks:

+

```
cluster_A::> storage aggregate create aggr1_node_A_1 -diskcount 10 -node
node_A_1 -mirror true
[Job 15] Job is queued: Create aggr1_node_A_1.
[Job 15] The job is starting.
[Job 15] Job succeeded: DONE
```

1. Verify the RAID group and drives of your new aggregate:

storage aggregate show-status -aggregate aggregate-name

Configuring FC-to-SAS bridges for health monitoring

- Third-party SNMP monitoring tools are not supported for FibreBridge bridges.
- Starting with ONTAP 9.8, FC-to-SAS bridges are monitored via in-band connections by default, and additional configuration is not required.



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

Steps

1. From the ONTAP cluster prompt, add the bridge to health monitoring:
 - a. Add the bridge, using the command for your version of ONTAP:

ONTAP version	Command
9.5 and later	storage bridge add -address 0.0.0.0 -managed-by in-band -name bridge-name
9.4 and earlier	storage bridge add -address bridge-ip-address -name bridge-name

b. Verify that the bridge has been added and is properly configured:

```
storage bridge show
```

It might take as long as 15 minutes to reflect all data because of the polling interval. The ONTAP health monitor can contact and monitor the bridge if the value in the **Status** column is **ok**, and other information, such as the worldwide name (WWN), is displayed.

The following example shows that the FC-to-SAS bridges are configured:

```
controller_A_1::> storage bridge show

Bridge              Symbolic Name  Is Monitored  Monitor Status
Vendor Model        Bridge WWN
-----
-----
ATTO_10.10.20.10    atto01        true          ok             Atto
FibreBridge 7500N    20000010867038c0
ATTO_10.10.20.11    atto02        true          ok             Atto
FibreBridge 7500N    20000010867033c0
ATTO_10.10.20.12    atto03        true          ok             Atto
FibreBridge 7500N    20000010867030c0
ATTO_10.10.20.13    atto04        true          ok             Atto
FibreBridge 7500N    2000001086703b80

4 entries were displayed

controller_A_1::>
```

Moving a metadata volume in MetroCluster configurations

You can move a metadata volume from one aggregate to another aggregate in a MetroCluster configuration. You might want to move a metadata volume when the source aggregate is decommissioned or unmirrored, or for other reasons that make the aggregate ineligible.

- You must have cluster administrator privileges to perform this task.
- The target aggregate must be mirrored and should not be in the degraded state.
- The available space in the target aggregate must be larger than the metadata volume that you are moving.

Steps

1. Set the privilege level to advanced:

```
set -privilege advanced
```

2. Identify the metadata volume that should be moved:

```
volume show MDV_CRS*
```

```

Cluster_A::*> volume show MDV_CRS*
Vserver   Volume                Aggregate             State    Type    Size
Available Used%
-----
Cluster_A
          MDV_CRS_14c00d4ac9f311e7922800a0984395f1_A
                Node_A_1_aggr1
                        online    RW      10GB
9.50GB    5%
Cluster_A
          MDV_CRS_14c00d4ac9f311e7922800a0984395f1_B
                Node_A_2_aggr1
                        online    RW      10GB
9.50GB    5%
Cluster_A
          MDV_CRS_15035e66c9f311e7902700a098439625_A
                Node_B_1_aggr1
                        -        RW      -
-        -
Cluster_A
          MDV_CRS_15035e66c9f311e7902700a098439625_B
                Node_B_2_aggr1
                        -        RW      -
-        -
4 entries were displayed.

Cluster_A::>

```

3. Identify an eligible target aggregate:

metrocluster check config-replication show-aggregate-eligibility

The following command identifies the aggregates in cluster_A that are eligible to host metadata volumes:

```
Cluster_A::*> metrocluster check config-replication show-aggregate-eligibility
```

```
Aggregate Hosted Config Replication Vols Host Addl Vols Comments
-----
-----
Node_A_1_aggr0 - false Root Aggregate
Node_A_2_aggr0 - false Root Aggregate
Node_A_1_aggr1 MDV_CRS_1bc7134a5ddf11e3b63f123478563412_A true -
Node_A_2_aggr1 MDV_CRS_1bc7134a5ddf11e3b63f123478563412_B true -
Node_A_1_aggr2 - true
Node_A_2_aggr2 - true
Node_A_1_Aggr3 - false Unable to determine available space of aggregate
Node_A_1_aggr5 - false Unable to determine mirror configuration
Node_A_2_aggr6 - false Mirror configuration does not match requirement
Node_B_1_aggr4 - false NonLocal Aggregate
```



In the previous example, Node_A_1_aggr2 and Node_A_2_aggr2 are eligible.

4. Start the volume move operation:

```
volume move start -vserver svm_name -volume metadata_volume_name -destination -aggregate destination_aggregate_name
```

The following command moves metadata volume MDV_CRS_14c00d4ac9f311e7922800a0984395f1 from aggregate Node_A_1_aggr1 to aggregate Node_A_1_aggr2:

```
Cluster_A::*> volume move start -vserver svm_cluster_A -volume
MDV_CRS_14c00d4ac9f311e7922800a0984395f1
-destination-aggregate aggr_cluster_A_02_01

Warning: You are about to modify the system volume
"MDV_CRS_9da04864ca6011e7b82e0050568be9fe_A". This may cause
severe
performance or stability problems. Do not proceed unless
directed to
do so by support. Do you want to proceed? {y|n}: y
[Job 109] Job is queued: Move
"MDV_CRS_9da04864ca6011e7b82e0050568be9fe_A" in Vserver
"svm_cluster_A" to aggregate "aggr_cluster_A_02_01".
Use the "volume move show -vserver svm_cluster_A -volume
MDV_CRS_9da04864ca6011e7b82e0050568be9fe_A" command to view the status
of this operation.
```


5. Verify the state of the volume move operation:

```
volume move show -volume vol_constituent_name
```

6. Return to the admin privilege level:

```
set -privilege admin
```

Checking the MetroCluster configuration

You can check that the components and relationships in the MetroCluster configuration are working correctly. You should do a check after initial configuration and after making any changes to the MetroCluster configuration. You should also do a check before a negotiated (planned) switchover or a switchback operation.

If the `metrocluster check run` command is issued twice within a short time on either or both clusters, a conflict can occur and the command might not collect all data. Subsequent `metrocluster check show` commands do not show the expected output.

Steps

1. Check the configuration:

```
metrocluster check run
```

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

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

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

Component	Result
nodes	ok
lifs	ok
config-replication	ok
aggregates	ok
clusters	ok
connections	ok

6 entries were displayed.

2. Display more detailed results from the most recent `metrocluster check run` command:

```
metrocluster check aggregate show
```

```
metrocluster check cluster show
```

```
metrocluster check config-replication show
```

```
metrocluster check lif show
```

```
metrocluster check node show
```

The `metrocluster check show` commands show the results of the most recent `metrocluster check run` command. You should always run the `metrocluster check run` command prior to using the `metrocluster check show` commands so that the information displayed is current.

The following example shows the `metrocluster check aggregate show` command output for a healthy four-node MetroCluster configuration:

```
cluster_A::> metrocluster check aggregate show

Last Checked On: 8/5/2014 00:42:58

Node          Aggregate          Check
Result
-----
controller_A_1 controller_A_1_aggr0
ok
ok
ok
controller_A_1_aggr1
ok
ok
ok
controller_A_1_aggr2
ok
ok
ok
```

```

controller_A_2      controller_A_2_aggr0
ok
ok
ok
ok
controller_A_2_aggr1
ok
ok
ok
ok
controller_A_2_aggr2
ok
ok
ok
ok
18 entries were displayed.

```

The following example shows the `metrocluster check cluster show` command output for a healthy four-node MetroCluster configuration. It indicates that the clusters are ready to perform a negotiated switchover if necessary.

Last Checked On: 9/13/2017 20:47:04

Cluster	Check	Result
mccint-fas9000-0102	negotiated-switchover-ready	not-applicable
	switchback-ready	not-applicable
	job-schedules	ok
	licenses	ok
	periodic-check-enabled	ok
mccint-fas9000-0304	negotiated-switchover-ready	not-applicable
	switchback-ready	not-applicable
	job-schedules	ok
	licenses	ok
	periodic-check-enabled	ok

10 entries were displayed.

Checking for MetroCluster configuration errors with Config Advisor

You can go to the NetApp Support Site and download the Config Advisor tool to check for common configuration errors.

Config Advisor is a configuration validation and health check tool. You can deploy it at both secure sites and non-secure sites for data collection and system analysis.



Support for Config Advisor is limited, and available only online.

Steps

1. Go to the Config Advisor download page and download the tool.

[NetApp Downloads: Config Advisor](#)

2. Run Config Advisor, review the tool's output and follow the recommendations in the output to address any issues discovered.

Sending a custom AutoSupport message after to adding nodes to the MetroCluster configuration

You should issue an AutoSupport message to notify NetApp technical support that maintenance is complete.

This task must be performed on each MetroCluster site.

Steps

1. Log in to the cluster at Site_A.
2. Invoke an AutoSupport message indicating the end of the maintenance:

```
system node autosupport invoke -node * -type all -message MAINT=end
```

3. Repeat this step on the partner site.

Verifying switchover, healing, and switchback

You should verify the switchover, healing, and switchback operations of the MetroCluster configuration.

Steps

1. Use the procedures for negotiated switchover, healing, and switchback that are mentioned in the *MetroCluster Management and Disaster Recovery Guide*.

[MetroCluster management and disaster recovery](#)

Copyright Information

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

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

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

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

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

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

Trademark Information

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