Technical Report

Continuous FlexPod Operations

David Klem, Dave Eckert, Craig Chadwell, NetApp
January 2014 | TR-4257

TABLE OF CONTENTS

1 Introduction 4

1.1 General Considerations 4

2 Business Challenges 5

3 Addressing These Business Challenges 6

3.1 Upgrade Clustered Data ONTAP 7

3.2 Upgrade PCIe Cards in FAS 18

3.3 Grow a Cluster 22

3.4 Add Clustered Data ONTAP FAS to 7-Mode FlexPod 36

3.5 Add Shelf/Shelves to Existing Storage Controller 59

3.6 Permanently Removing a Controller from a Clustered Data ONTAP Cluster 67

Appendix 72

Aggregate Relocation-Based Controller Upgrade 72

Workload Rebalancing 72

Creating a Switchless Cluster 82

LIST OF TABLES

Table 1) Nondisruptive operations per protocol. 5

Table 2) Business challenges. 6

Table 3) List of variables. 7

Table 4) Data ONTAP upgrade considerations. 11

Table 5) PCIe card upgrade—dependencies. 18

Table 6) List of variables. 22

Table 7) Joining FAS to clustered Data ONTAP considerations. 27

Table 8) Configuration variables. 36

Table 9) Upgrade procedure. 42

Table 10) Cisco Nexus 5596 cluster network switch configuration prerequisites. 42

Table 11) List of variables. 59

Table 12) List of variables. 67

Table 13) Workload rebalancing procedures. 73

LIST OF FIGURES

Figure 1) Before Data ONTAP upgrade 8

Figure 2) During Data ONTAP upgrade 1. 9

Figure 3) During Data ONTAP upgrade 2. 10

Figure 4) After Data ONTAP upgrade....11
Figure 5) Adding an HA pair to a FlexPod system....25
Figure 6) Ethernet before. 26
Figure 7) Ethernet during 1....26
Figure 8) Ethernet during 2....27
Figure 9) After. 27
Figure 10) Single HA pair in 7-Mode. 40
Figure 11) Data ONTAP 7-Mode HA pair and clustered Data ONTAP HA pair. 41
Figure 12) Data ONTAP 7-Mode and clustered Data ONTAP HA pair online. 41
Figure 13) Add shelf/shelves to existing storage controller. 60
Figure 14) MPHA SAS cabling. 60
Figure 15) ACP cabling. 61
Figure 16) MPHA SAS cabling. 61
Figure 17) ACP cabling. 62
Figure 18) MPHA SAS cabling. 62
Figure 19) ACP cabling. 63
Figure 20) Before. 68
Figure 21) During 1. 68
Figure 22) During 2. 69
Figure 23) After. 69
Figure 24) Before. 77
Figure 25) After. 78
Figure 26) Verify FC port configuration. 80
Figure 27) Switchless cluster configuration. 83

1 Introduction

Industry trends show a clear transformation in the data center from traditional silo-based architectures into shared infrastructure and cloud computing. Enterprise customers are moving toward this model to increase agility and reduce costs. Because companies must address resistance to change in both their organizational and technical IT models, achieving this transformation can seem daunting and complex. To accelerate the process and simplify the evolution to a shared cloud infrastructure, Cisco and NetApp have developed a solution called FlexPod® Datacenter.

FlexPod Datacenter is a predesigned, best practice data center architecture that is built on Cisco UCS®, the Cisco Nexus® family of switches, and NetApp® fabric-attached storage (FAS) systems. FlexPod Datacenter is a suitable platform for running a variety of virtualization hypervisors as well as bare metal operating systems and enterprise workloads. It delivers a baseline configuration and also has the flexibility to be sized and optimized to accommodate many different use cases and requirements.

Although FlexPod Datacenter helps to simplify the transition to a shared infrastructure, because multiple applications are consolidated, many other considerations must be taken into account. For instance, the importance of nondisruptive operations (NDO) within FlexPod Datacenter cannot be understated. With a traditional silo-based infrastructure that contained a single application, scheduling downtime to upgrade the hardware or software within the infrastructure, although not ideal, was still manageable. Now with multiple applications sharing that same infrastructure, it becomes almost impossible to coordinate scheduled downtime between all of the applications.

Continuous FlexPod Operations and clustered Data ONTAP® solve this problem by providing true NDO within the FlexPod Datacenter environment. Hardware and software can be upgraded, resources added and removed, and application workloads moved across hardware with no downtime. This document covers several use cases around storage NDO within a FlexPod Datacenter environment and describes in detail how to perform these tasks.

1.1 General Considerations

Each solution to the challenges detailed below is designed to be performed without a planned or unplanned disruption in service. Most application workloads hosted on a FlexPod unit should be able to implement the solutions to the given business challenges. Before engaging in any of the upgrade procedures mentioned below, NetApp highly recommends that the administrator consult with any application-specific upgrade guides for further guidance.

This document makes the following assumptions:

- The FlexPod configuration undergoing upgrades has been configured and cabled to best practice designs. For details and the official definition of FlexPod, refer to TR-4036: FlexPod Technical Specifications.
- The NetApp FAS controllers in the FlexPod configuration are running clustered Data ONTAP version 8.1.2 to 8.2.
- The administrator performing the upgrade procedures has insight into the network and compute layers of the FlexPod infrastructure.
- All hardware and software must be supported at every stage of the upgrade process to prevent disruption in service. Check the NetApp Interoperability Matrix to confirm hardware and software interoperability support.

Note: Not all protocols allow for true nondisruptive operations. Review Table 1 to see which protocols are supported nondisruptively with each clustered Data ONTAP feature. Contact your NetApp sales or support representative for more details.

Table 1) Nondisruptive operations per protocol.

Operation	SMB 1.0	SMB 2.x/3.0	SMB 3.0 (Hyper-V or SQL Server)	NFS	SAN
LIF migrate	Possibly disruptive	✔	✔	✔	N/A
Volume move	✔	✔	✔	✔	✔
Aggregate relocate	Disruptive	Disruptive	✔	✔	✔
Storage failover	Disruptive	Disruptive	✔	✔	✔

Many of the preupgrade tasks help to confirm that the non storage components of the FlexPod configuration are also configured according to NetApp best practices. Reviewing the configuration of components in the FlexPod unit prior to engaging in the upgrade procedures makes sure that all best practices have been followed during configuration. This allows the administrator an opportunity to correct any configurations that are out of line with best practices prior to engaging in activities that could cause a disruption if best practices were not followed.

2 Business Challenges

Business operations should never feel the impact of hardware or infrastructure software changes within the data center. However, there are many events during the lifecycle of data center infrastructure that require hardware or software changes. These changes keep the company's IT infrastructure aligned with the business needs. The FlexPod architecture and operational best practices allow for nondisruptive operations for various events that take place during the life of the equipment.

Business Challenges
Equipment end of lease/end of life
Increasing business demands
Feature enhancements
Company compliance requirements

A FlexPod unit that utilizes NetApp FAS unified storage arrays running clustered Data ONTAP allows these challenges to be addressed as part of the following scenarios.

- **The FAS controllers within a FlexPod unit have reached the end of their lease period or the end of their useful life.** As a result, the FlexPod administrator needs to replace the existing controllers without disrupting active workloads being hosted on those controllers.
- **Business demands have outgrown the compute capacity of the existing FlexPod storage controllers.** As a result, the FlexPod administrator needs to increase the compute capacity of the storage environment. In this case, the business has two options: replace the controllers and/or add new controller/shelves to the existing cluster and migrate workloads onto the new resources. Both of these tasks can be nondisruptively performed. Additionally, the migration from a switchless to a switched cluster should be considered if the FlexPod unit is currently running a two-node switchless cluster to allow for expansion.
- **Business demands have outgrown the storage performance/capacity of an existing FlexPod unit.** As a result, the business has three options: grow the existing storage capacity/performance and

rebalance workloads, install new storage shelves and rebalance workloads, and/or utilize flash technology to increase existing storage performance.

- **The disks and shelves within a FlexPod unit have reached the end of their lease period or the end of their useful life.** As a result the FlexPod administrator must replace the existing shelves with newer models without disrupting active workloads that are being hosted on those controllers.
- **The FAS controllers within a FlexPod unit have an older or noncompliant operating system installed.** If the FlexPod unit contains storage controllers whose operating system is a major version/minor version/patch behind, the administrator might need to upgrade the version of Data ONTAP to enable features, fix bugs affecting workload performance, and/or bring the FlexPod unit into compliance with security standards.

3 Addressing These Business Challenges

FlexPod provides specific procedures to address the business challenges in outlined in Table 2. In some cases multiple procedures might be required.

Table 2) Business challenges.

Business Challenges	Nondisruptive Procedures That Can Address Challenges
End-of-lease/end-of-life equipment	• Data ONTAP software upgrade • Add shelf/shelves • Add controllers to an existing cluster • Retire controllers from a cluster
Increasing business demands	• Add FAS to an existing cluster • Add clustered Data ONTAP FAS to 7-Mode FlexPod • Add adapters • Add shelf/shelves
Feature enhancements	• Data ONTAP software upgrade • Add adapters (similar to upgrading Data ONTAP) • Add clustered Data ONTAP FAS to 7-Mode FlexPod
Company compliance requirements	• Data ONTAP software upgrade

End-of-Life Equipment

When the storage components of a FlexPod unit have reached their end of life or end of lease, both the controllers and shelves might need to be upgraded and the workloads migrated to the new equipment.

Alternatively, the administrator could choose to temporarily expand the cluster to include the new controllers and shelves, migrate the workloads, and then retire the old controllers and shelves. In this case, make sure that both the old and the new controllers are running the same software version of Data ONTAP.

In either case, after the workloads have been migrated from the old equipment to the new, the old equipment can be retired and nondisruptively disconnected.

The administrator can determine the appropriate upgrade path based on information such as the equipment retirement and return timeline and the space/power/network availability within the data center.

Add Hardware

To deal with increased business demands that can only be resolved by adding more storage hardware, FlexPod offers several procedures to accommodate specific business demands. These demands might affect the storage in the following ways:

- Increased CPU demand on storage controllers
- Increased IOPS demand on disks
- Increased capacity demand on disks

To resolve the increased CPU demand on the storage controllers, an administrator can add new controllers with shelves and rebalance the workload or upgrade to new controllers. FlexPod facilitates the first scenario by allowing the addition of storage nodes to an existing cluster and, in some cases, adding a new clustered Data ONTAP FAS cluster to a FlexPod unit that might already have an existing pair of controllers running in Data ONTAP 7-Mode.

To provide more IOPS, the business may add new shelves or Flash Cache™ adapters to their storage controllers. When adding more shelves, they might choose to expand existing aggregates and allow the workload to rebalance across the new disks. This can include the creation of Flash Pool™ aggregates in cases where having an automatically tiered SSD solution would be valuable. Alternatively, the business might choose to create new aggregates and rebalance manually by migrating workloads to the new aggregates.

To provide more storage capacity, the business must add new shelves. After, the administrator has two options: add the disks to the existing workloads to accommodate the space requirements or create new aggregates and migrate workloads as necessary.

Feature Enhancements

Feature enhancements including Flash Cache, previously unavailable features prior to clustered Data ONTAP 8.2 such as SnapVault® or storage QoS, and features that might not be available to an HA pair running Data ONTAP in 7-Mode are resolved by upgrading the hardware or Data ONTAP version of existing storage controllers. Another option would be to add an additional storage cluster to the environment that has the desired hardware/software versions to support the desired features. The administrator should choose the upgrade method based on what best serves the needs of the business and can be accommodated by the specific workloads under consideration.

The following sections cover the in-depth technical procedures required for nondisruptively addressing each of these business challenges.

3.1 Upgrade Clustered Data ONTAP

Table 3) List of variables.

Configuration Variable Name	Description
<<var_pkg_location>>	The path to the Data ONTAP upgrade package to be downloaded. This can be an HTTP, FTP, or TFTP URL or the path of a mounted file system.
<<var_node01>>	The first controller in an HA pair of controllers in a cluster.
<<var_node02>>	The second controller in an HA pair of controllers in a cluster.

This section provides the preupgrade considerations, upgrade process, and postupgrade verification guidance required to make sure that the controller's operating system upgrade is nondisruptively performed.

Scenario

A FAS controller is nondisruptively upgraded from clustered Data ONTAP 8.1.2 to clustered Data ONTAP 8.2. This upgrade procedure could be applied to any HA pair of controllers within a Data ONTAP cluster.

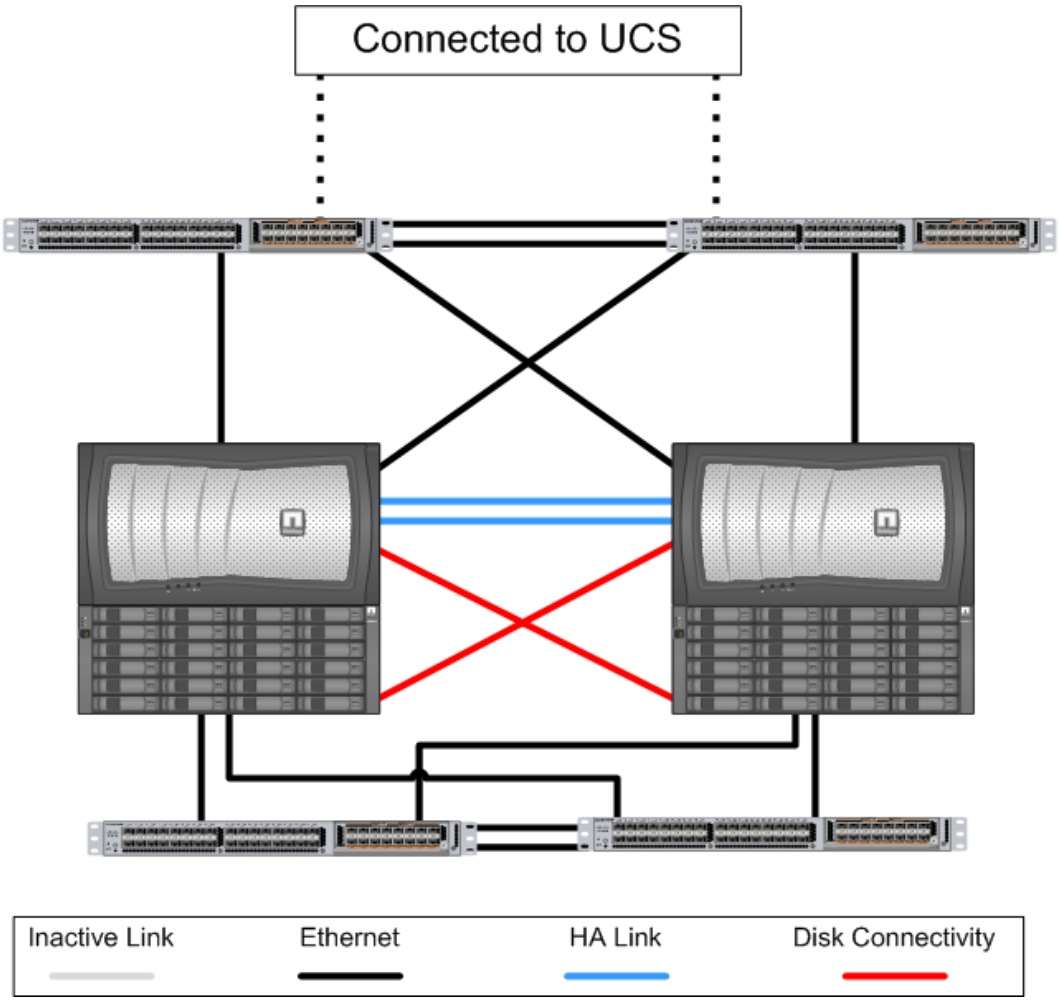
Note: This document assumes that the FlexPod system has been properly cabled and configured according the FlexPod technical specifications. See TR-4036: FlexPod Technical Specifications for more details.

Overview

Before

In a standard FlexPod configuration, all links between the compute, network, and storage layer are active by leveraging technologies such as Cisco® Virtual Port Channels (vPCs) within the networking layer and interface groups (ifgrps) on the storage controllers. Redundant high-availability (HA) links are used between the storage controllers for any storage failover events, planned or unplanned.

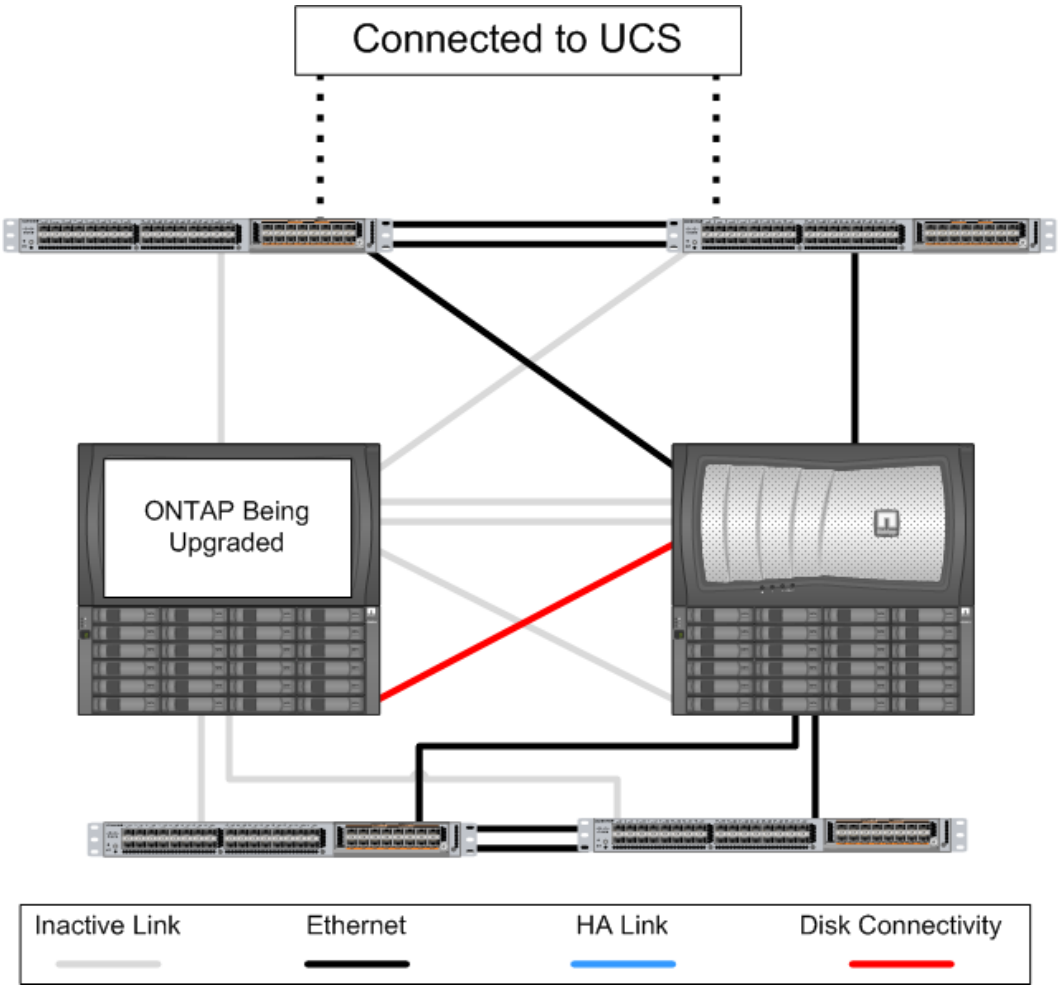
Figure 1) Before Data ONTAP upgrade.



During

During a Data ONTAP upgrade, the controllers are individually upgraded, one at a time. One controller performs a takeover to handle the active duties of both controllers in the HA pair. The remaining controller, now with no active traffic, can then be nondisruptively upgraded.

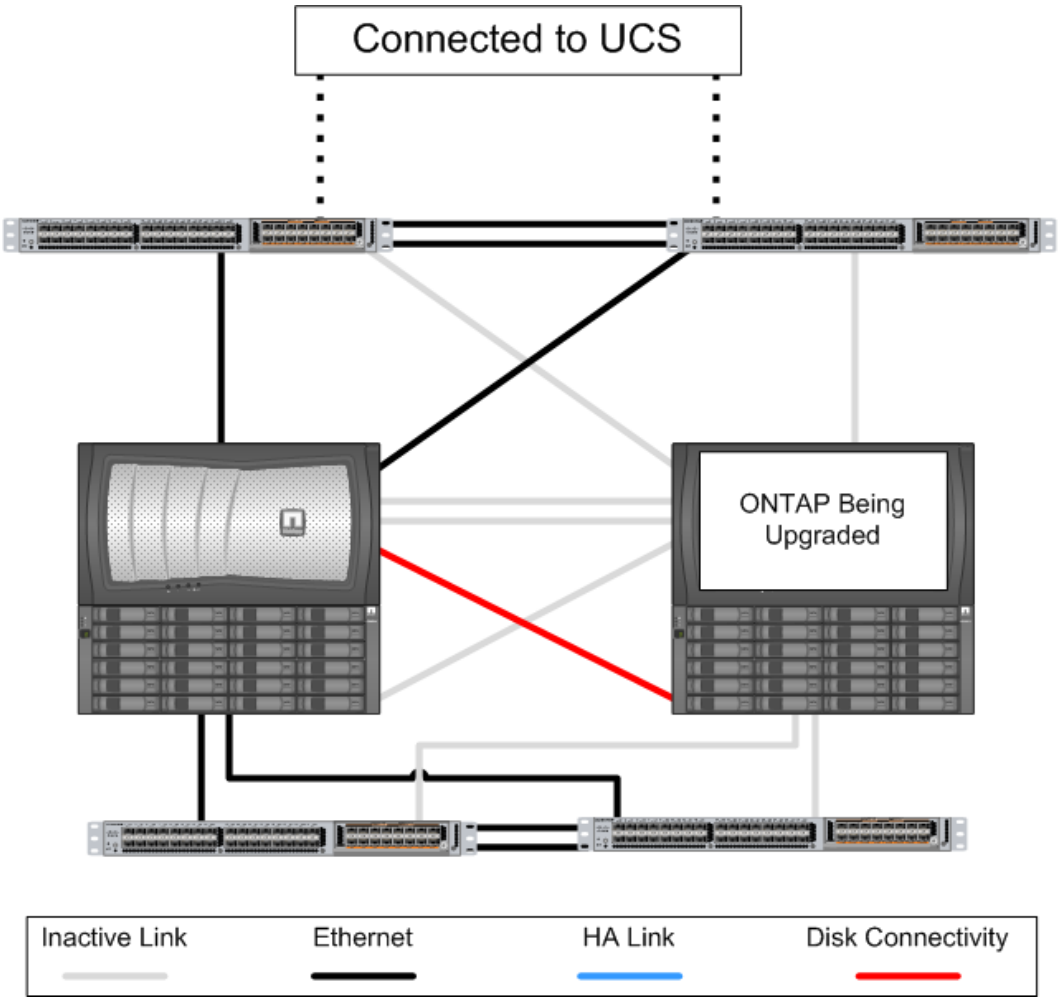
Figure 2) During Data ONTAP upgrade 1.



After both nodes are back online, the entire process is then repeated in the reverse direction, with the newly upgraded controller performing a takeover for its partner. That controller is then upgraded to the desired clustered Data ONTAP version using the same procedure.

Note: Due to the nature of the Data ONTAP software upgrade process, a single point of failure will be exposed during the procedure due to a single storage controller taking over the active connections for the entire system during a takeover, to upgrade the partner controller.

Figure 3) During Data ONTAP upgrade 2.



After

After the upgrade process, the partner performs a giveback, making them both active, and connectivity to the controller and full redundancy in the environment are restored.

Figure 4) After Data ONTAP upgrade.

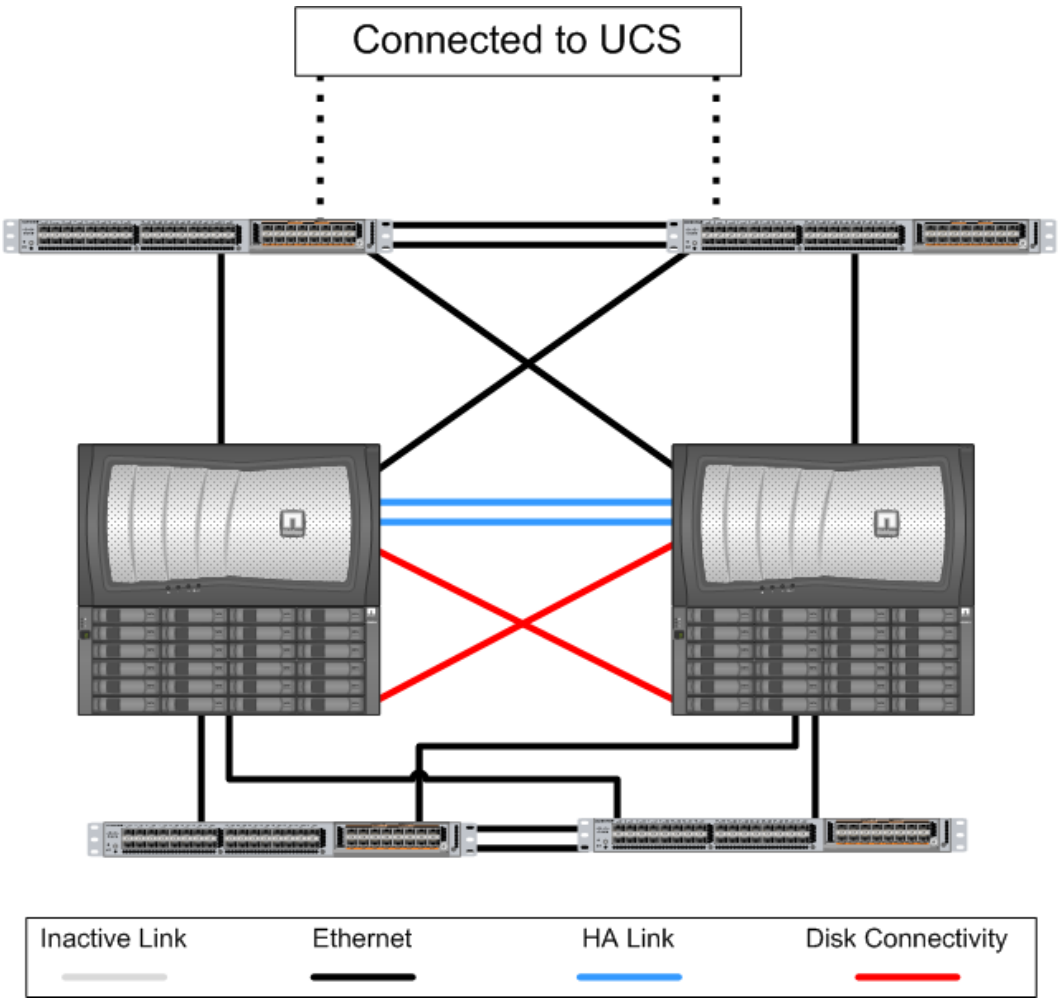


Table 4) Data ONTAP upgrade considerations.

Considerations	Description
The total number of storage objects owned by the nodes in an HA pair should not exceed the maximum number approved for nondisruptive upgrade.	See pages 16 and 24 of Clustered Data ONTAP 8.2 Upgrade and Revert/Downgrade Guide.
Verify that the controllers, shelves, and storage hosts are supported by the new Data ONTAP version.	Check NetApp Hardware Universe.
Make sure that the combined CPU utilization, data network utilization, or cluster network utilization for both nodes does not exceed 100%.	See pages 16 and 24 of Clustered Data ONTAP 8.2 Upgrade and Revert/Downgrade Guide.
Make sure that all NAS LIFs have a failover group and failover policy and that failover is enabled. Make sure that the failover groups assigned to NAS LIFs have member ports from multiple nodes.	To check the failover parameters of NAS LIFs from CLI, run `network interface show -data-protocol nfs -fields failover-group, failover-policy, use-failover-group` and `network interface show -data-protocol cifs -fields failover-group,`

Considerations	Description
	`failover-policy, use-failover-group` To check the member ports of a failover group, run `network interface failover-groups show`
Make sure that the correct VLANs on the switches are trunked across all ports, port channels, or virtual port channels that map to failover groups on the storage controllers.	To check proper VLAN trunking, on each switch, for each interface, run `show running-config interface [interface-name]` Compare the interfaces that correspond to ports in the same failover group.
Make sure that all Vservers (also known as storage virtual machine [SVM]) using SAN protocols have SAN LIFs on multiple nodes.	To check whether a Vserver contains SAN LIFs from multiple nodes from CLI, run `network interface show -data-protocol fcp -fields curr-port, curr-node`.
Make sure that all zones on both switches contain WWPNs or IQNs from multiple nodes.	To check SAN zoning redundancy, on both switches run the `show zone` command. Confirm that each zone contains at least one WWPN/IQN from at least two nodes in the cluster.
Provide a connection to the controllers through serial connection, remote LAN module, or baseboard management controller. SSH and OCSM will not work for the NDU process.	
Synchronize date and time.	To confirm that the time for all nodes is synchronized, run the `node date show` command. Run the `node date modify` command to adjust time/date. If your cluster has access to an NTP server, use `system services ntp server modify` to adjust NTP server settings.
Remove all failed disks.	1. Run `storage disk show -broken` to discover and failed disks. Note the disk names. 2. Run `storage disk set-led -disk [disk_name]` to light the fault LED indicator. 3. Remove the disk from the shelf. For more information, see Removing a failed disk.

Make sure that the cluster is in quorum.	1. Use `set -privilege advanced` to enable advanced commands.
	2. Run the following commands:
	`cluster ring show -unitname mgmt`
	`cluster ring show -unitname vldb`
	`cluster ring show -unitname vifmgr`
	`cluster ring show -unitname bcomd`
	For each command above, make sure that the Epoch and Epoch DB columns match for all nodes. If out of quorum, contact NetApp support.
	For more information, see Verifying that the cluster is in quorum.

Make sure that SnapMirror® operations	For details, see pages 35–36 of Clustered Data ONTAP 8.2

Considerations	Description
are suspended.	Upgrade and Revert/Downgrade Guide.
Make sure that no jobs are running.	1. Run `job show` to determine what jobs are running or queued. 2. Run `job delete *` to delete jobs related to aggregates, volumes, or Snapshot™ copies. For details, see page 37 of Clustered Data ONTAP 8.2 Upgrade and Revert/Downgrade Guide.
Make note of the owners and status of aggregates owned by the controllers being upgraded.	Run `storage aggregate show`. Make note of the Nodes column.
Make sure that all protocols used on the cluster support nondisruptive upgrade.	See section 1.1, “General Considerations.”

Note: Refer to the Cluster upgrade checklist.

Upgrade Data ONTAP

To upgrade Data ONTAP, complete the following steps:

1. Download the clustered Data ONTAP 8.2 image to each node in the cluster.

```
system node image update -node * -package <<var_pkg_location>> -replace-package true -background
true
```

2. Verify the software image has been downloaded and installed on each node.

```
system node image show-update-progress -node *
```

Example output of successful upgrade:

```
There is no update/install in progress
Status of most recent operation:
 Run Status: Exited
 Exit Status: Success
 Phase: Run Script
 Exit Message: Installation complete. image2 updated on node node0.
There is no update/install in progress
Status of most recent operation:
 Run Status: Exited
 Exit Status: Success
 Phase: Run Script
 Exit Message: Installation complete. image2 updated on node node1.
2 entries were acted on
```

3. Verify that the correct images are installed on each node.

```
system node image show
```

Example output:

```
                  Is        Is                   Install
Node      Image   Default   Current    Version    Date
--------  ------- -------   -------    --------   -------------------
node1
          image1    true    true       8.1.2    10/25/2012 12:37:36
          image2    false   false      8.2P4    10/22/2013 13:52:22
node2
          image1    true    true       8.1.2    10/25/2012 12:41:16
          image2    false   false      8.2P4    10/22/2013 13:55:22
4 entries were displayed.
```

4. Set the new image as the default image for the node being upgraded.

During Node 1 Upgrade	During Node 2 Upgrade
`system node image modify -node <<var_node01>> -image image2 -isdefault true` **Note:** If your environment lists image1 as the most recent package, use that instead.	`system node image modify -node <<var_node02>> -image image2 -isdefault true` **Note:** If your environment lists image1 as the most recent package, use that instead.

5. Verify that the correct image has been set as the default image for the node being upgraded.

```
system node image show
```

Example output:

```
                  Is        Is                   Install
Node      Image   Default   Current    Version    Date
--------  ------- -------   -------    --------   -------------------
node1
          image1    false    true       8.1.2    10/25/2012 12:37:36
          image2    true     false      8.2P4    10/22/2013 13:52:22
node2
          image1    true    true       8.1.2    10/25/2012 12:41:16
          image2    false   false      8.2P4    10/22/2013 13:55:22
4 entries were displayed.
```

6. Make sure that high availability (HA) is enabled and possible on both controllers.

```
storage failover show
```

Example output:

```
Node              Partner         Possible  State
----------------- --------------- -------- ------------------------------------
node1             node2            true       Connected to node2
node2             node1            true       Connected to node1
2 entries were displayed.
```

Note: During the upgrade of the second node, the following might be displayed, which is expected behavior.

```
Node              Partner         Possible  State
----------------- --------------- -------- ------------------------------------
node1             node2            false      Connected to node2,
                                              takeover is not possible. The
                                              version of software running on each
                                              node of the SFO pair is incompatible.
                                              NVRAM log not synchronized
node2             node1            false      Connected to node1,
                                              takeover is not possible. The
                                              version of software running on each
                                              node of the SFO pair is incompatible.
                                              NVRAM log not synchronized

2 entries were displayed.
```

7. If the cluster only has two nodes, skip to step 15. If the node is part of a cluster that contains more than two nodes, make sure that epsilon is not located on the next node that will be shut down by following the next steps.
8. Enter advanced privileges mode.

```
set -privilege advanced
```

Example output:

```
Warning: These advanced commands are potentially dangerous; use them only when directed to do so
by NetApp personnel.
```

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

9. Check cluster health.

```
cluster show
```

Example output:

```
Node           Health         Eligibility           Epsilon
------------   -------        ------------          ------------
node0           true           true                 false
node1           true           true                 true
node2           true           true                 false
node3           true           true                 false
4 entries were displayed.
```

10. Remove and reassign epsilon.

```
cluster modify -node <<var_controller_name>> -epsilon false
cluster modify -node <<var_other_node>> -epsilon true
```

11. Return to the administration console mode.

```
set -privilege admin
```

12. Disable automatic giveback on both nodes.

```
storage failover modify -node <<var_node01>> -auto-giveback false
storage failover modify -node <<var_node02>> -auto-giveback false
```

13. Verify automatic giveback is disabled for both nodes.

```
storage failover show -auto-giveback false
```

Example output:

```
                              Takeover
Node              Partner        Possible    State
--------------- -------------- -------- -------------------------------------
node0             node1             true       Connected to node1
node1             node0             true       Connected to node0
2 entries were displayed.
```

14. Migrate LIFs away from the given node to make sure that there's no disruption in service during the takeover process.

During Node 1 Upgrade	During Node 2 Upgrade
`network interface migrate-all -node <<var_node01>>`	`network interface migrate-all -node <<var_node02>>`

15. Verify that LIFs have been migrated away from the node being upgraded. No NAS LIFs should display on the node being upgraded in the "Current Node" column.

During Node 1 Upgrade	During Node 2 Upgrade
`network interface show -data-protocol nfs|cifs -role data -curr-node <<var_node01>>`	`network interface show -data-protocol nfs|cifs -role data -curr-node <<var_node02>>`

Example output:

```
            Logical     Status     Network           Current      Current  Is
Vserver     Interface   Admin/Oper Address/Mask      Node           Port     Home
----------- ---------- ---------- ------------------ ------------- ------- ----
vs0
               lif1        up/up      192.0.2.130/24        node1          e0b     true
               lif2        up/up      192.0.2.131/24        node1          e0b     false
               lif3        up/up      192.0.2.132/24        node1          e0b     true
vs1
               lif1        up/up      192.0.2.133/24        node1          e0b     false
               lif2        up/up      192.0.2.134/24        node1          e0b     true
```

```
5 entries were displayed.
```

16. Trigger AutoSupport™ notification.

```
system node autosupport invoke -node <<var_node01>> -type all -message "starting_NDU"
```

17. Initiate a takeover of the node being upgraded.

During Node 1 Upgrade	During Node 2 Upgrade
`storage failover takeover -ofnode <<var_node01>>`	`storage failover takeover -ofnode <<var_node02>> -options allow-version-mismatch`

18. Verify takeover.

```
storage failover show
```

When executing this command, the following outputs might be displayed and are normal:

```
                              Takeover
Node            Partner         Possible     State Description
--------------- --------------- -------- -------------------------------------
node0            node1          false         Pending
node1            node0          false         Takeover scheduled in 163 seconds
2 entries were displayed
```

```
                              Takeover
Node            Partner         Possible     State Description
--------------- --------------- -------- -------------------------------------
node0            node1          -             Unknown
node1            node0           false        In takeover
2 entries were displayed
```

After `storage failover show` displays the following, the upgrade has been completed, and the node has rebooted. Proceed to the next step.

```
                              Takeover
Node            Partner         Possible     State Description
--------------- --------------- -------- -------------------------------------
node0            node1          -             Waiting for giveback (HA mailboxes)
node1            node0           false        In takeover
2 entries were displayed
```

19. Wait eight minutes to make sure that multipathing is stabilized and the clients have recovered.
20. Initiate a giveback to return the upgraded controller to an operational state and reestablish control over its disks.

During Node 1 Upgrade	During Node 2 Upgrade
`storage failover giveback -ofnode <<var_node01>>`	`storage failover giveback -fromnode <<var_node01>> -override-vetoes true -required-partner-waiting false`

21. Wait while the aggregates are given back and watch for errors.

```
storage failover show-giveback
```

Note: If any errors occur, refer to the clustered Data ONTAP Upgrade Guide.

Example output after the giveback process has completed:

```
               Partner
Node           Aggregate         Giveback Status
-------------- ----------------- -------------------------------------------
node0            -                 No aggregates to give back
node1            -                 No aggregates to give back
```

22. Verify all aggregates have been given back and are online.

During Node 1 Upgrade	During Node 2 Upgrade
`storage aggregate show -node <<var_node01>>`	`storage aggregate show -node <<var_node02>>`

Example output:

```
                            Takeover
Node             Partner         Possible    State
---------------- --------------- -------- ------------------------------------
node0            node1             true       Connected to node1
node1            node0             true       Connected to node0
2 entries were displayed.
```

23. If required, revert the LIFs to the upgraded node.

```
network interface revert *
```

24. Verify all LIFs have reverted. The Status Admin/Oper column should display “up/up,” and the “Is Home” column should display “true” for all LIFs that have migrated back successfully.

```
network interface show
```

Example output:

```
            Logical     Status     Network           Current      Current  Is
Vserver     Interface   Admin/Oper Address/Mask      Node           Port     Home
----------- ---------- ---------- ------------------ ------------- ------- ----
vs0
               lif1        up/up      192.0.2.130/24         node1          e0b     true
               lif2        up/up      192.0.2.131/24         node0          e0b     true
               lif3        up/up      192.0.2.132/24         node1          e0b     true
vs1
               lif1        up/up      192.0.2.133/24         node0          e0b     true
               lif2        up/up      192.0.2.134/24         node1          e0b     true
5 entries were displayed.
```

25. Repeat steps 3 through 23 for each node that is to be upgraded.

Postupgrade Procedure

1. Make sure that both controllers are running the proper Data ONTAP version. The “`Is Current`” column should display `true` for the 8.2P4 version for both nodes.

```
node image show
```

Example output:

```
                Is      Is                        Install
Node     Image  Default Current Version             Date
-------- ------- ------- ------- ------------------------- -------------------
node1
         image1  true    true    8.2P4             10/14/2013 09:28:16
         image2  false   false   8.1.2             12/3/2007 12:52:26
node-2
         image1  true    true    8.2P4             10/14/2013 09:30:40
         image2  false   false   8.1.2             12/3/2007 18:51:50
```

2. Reenable HA automatic giveback.

```
storage failover modify -node * -auto-giveback true
```

3. Make sure that failover reversion has happened successfully.
   - Aggregates:

```
aggregate show -fields is-home
```

Note: The `Is-Home` column should display true for all aggregates.

– LIFs:

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

Note: The `curr-port` and `home-port` columns should match, and the `curr-node` and `home-node` columns should match for each LIF.

4. If the cluster uses SnapMirror technology and relationships, those relationships will need to be updated. See Upgrading older style SnapMirror relationships for details.

Note: SnapMirror has undergone significant changes between clustered Data ONTAP 8.1.2 and clustered Data ONTAP 8.2. If the cluster uses SnapMirror technology, SnapMirror relationships must be updated to align with the new operating model before they can be used.

3.2 Upgrade PCIe Cards in FAS

To increase fault tolerance or performance or to enable new features on the FAS platform, upgrading or installing PCIe cards into the FAS platform might be necessary. In the context of a FlexPod system, this can be nondisruptively performed.

Scenario

A FAS HA pair running clustered Data ONTAP 8.1.2 to 8.2 will have new PCI adapters added or will have existing PCI adapters replaced with cards that have the same ports.

Dependencies

Current nodes have already been upgraded from 8.1.2 to 8.2. Follow the section “Upgrade Data ONTAP” to nondisruptively upgrade an 8.1.2 clustered HA pair to clustered Data ONTAP 8.2.

Table 5) PCIe card upgrade—dependencies.

Dependencies	Description
Verify that the existing nodes are in an HA pair and that takeover is possible.	Use the `storage failover show` command to confirm that nodes are in an HA pair and that takeover is possible.
1. Verify that neither node has CPU utilization greater than 50%. 2. Verify that no network interface on either node has utilization greater than 50%. 3. Verify that the combined number of aggregates between the two nodes is less than 100. 4. Verify that the combined number of volumes between the two nodes is less than the maximum number of volumes supported by the Data ONTAP version and FAS model for a single node.	1. Refer to pages 16 and 24 of the Clustered Data ONTAP 8.2 Upgrade and Revert/Downgrade Guide for instructions on determining CPU and network interface utilization. 2. Run the `aggr show -nodes <<var_node01>>| <<var_node02>>` command to determine the number of aggregates owned by the two nodes. The total number of entries displayed is the total number of aggregates owned by both nodes. 3. To verify the maximum number of support volumes, refer to KB8010101.
Make sure that all cables for the nodes being replaced are clearly and correctly labeled or identifiable.	Making sure that cables are correctly labeled will reduce the risk of wrongly cabling during node replacement.

Dependencies	Description
Make sure that network failover has been properly configured.	1. Run the network `interface show -fields home-node,home-port,curr-node,curr-port,failover-policy,failover-group,data-protocol` command to show all LIFs in the cluster. Each NAS LIF with either `curr-node` or `home-node` on one of the nodes being upgraded should have a failover group assigned and should have a failover policy of either nextavail or priority. 2. Run the `network interface failover-groups show` command to view the list of failover groups and its associated ports. Each failover group being used by one of the LIFs identified above should have ports from two or more nodes assigned.
Make sure that network has been properly configured to support network failover.	Review the configuration of the data switches to make sure that switch ports, port channels, and virtual port channels are trunking the correct VLANs. When a failover occurs, a LIF will fail to a port that supports the VLAN on the FAS port, ifgrp, and VLAN interface and the corresponding switch port, port channel, or virtual port channel.
Make sure that each Vserver using SAN LUNs housed on a node being replaced has at least one LIF per node in the HA pair.	1. For a mapped LUN that belongs to an aggregate owned by either node in the HA pair, confirm that the igroup mapped to that LUN has at least one LIF from each node in the HA pair. Use the following command to confirm that an igroup has LIFs from the appropriate nodes: `lun show -mapped mapped -fields volume,vserver`. 2. For each volume listed, run `volume show -volume <<var_volume_name>>`. 3. For each aggregate listed, run `storage aggregate show -aggregate <<var_aggregate_name>> -fields nodes`. 4. For each LUN mapped to an aggregate owned by one of the nodes being replaced, make sure that the Vserver has at least one LIF per node in the HA pair using `network interface show -data-protocol fcp|iscsi -Vserver <<var_vserver_name>>`.

Dependencies	Description
Make sure that all igroups bound to portsets have LIFs associated with both nodes in the HA pair.	1. For a mapped LUN that belongs to an aggregate owned by either node in the HA pair, confirm that the igroup mapped to that LUN has at least one LIF from each node in the HA pair. Use the following command to confirm that an igroup has LIFs from the appropriate nodes: `lun show -mapped mapped -fields volume` 2. For each volume listed, run `volume show -volume <<var_volume_name>>` 3. For each aggregate listed, run `storage aggregate show -aggregate <<var_aggregate_name>> -fields nodes` 4. For each LUN mapped to an aggregate owned by one of the nodes being replaced, run `lun show -lun <<var_lun_name>> -m` 5. For each igroup shown, run `igroup show -fields portset`. 6. For each LIF in the portset, run `network interface show -lif <<var_lif_name>> -fields curr-node` 7. Each portset should have at least one LIF from each node in the HA pair to prevent outage during the head swap operation.
Verify that the controller supports the card being installed.	Refer to the Hardware Universe to verify the adapter model is supported on the FAS model and Data ONTAP version.

PCIe Card Upgrade

Each controller must be upgraded separately. After the dependencies have been satisfied, begin with step 1 for the first controller. Do not attempt to perform the PCIe card upgrade process on both nodes simultaneously.

1. Migrate the NAS, cluster management, and Vserver management LIFs away from the node being upgraded.

Note: If failover groups have been properly configured, then `network interface migrate-all` will cause all LIFs to fail to available alternate interfaces. Otherwise, each LIF might need to be migrated independently.

During Node01 Upgrade	During Node02 Upgrade

• If you have failover groups configured, run `network interface migrate-all -node <<var_node01>>`
• If you do not have failover groups configured, use the following command for each NAS and management LIF on `<<var_node01>>`, inserting the appropriate destination node and port as best suits your environment.
`network interface show -curr-node <<var_node01>>`
`network interface migrate -vserver <<var_vserver_name>> -lif <<var_lif_name>> -dest-node <<var_dest_node_name>> dest-port <<var_dest_port_name>>`
	• If you have failover groups configured, run `network interface migrate-all -node <<var_node02>>`
	• If you do not have failover groups configured, use the following command for each NAS and management LIF on `<<var_node02>>`, inserting the appropriate destination node and port as best suits your environment.
	`network interface show -curr-node <<var_node02>>`
	`network interface migrate -vserver <<var_vserver_name>> -lif <<var_lif_name>> -dest-node <<var_dest_node_name>> dest-port <<var_dest_port_name>>`

Verify that the only LIFs that remain on the node being replaced are cluster interconnect LIFs, node management LIFs, or SAN protocol LIFs. `network interface show -curr-node <<var_node01>> -fields data-protocol`	Verify that the only LIFs that remain on the node being replaced are cluster interconnect LIFs, node management LIFs, or SAN protocol LIFs. `network interface show -curr-node <<var_node02>> -fields data-protocol`

Example output:

```
vserver     lif          role         data-protocol
----------- ------------ ------------ -------------
Node-01     clus1        cluster      none
Node-01     clus2        cluster      none
Node-01     mgmt1        node-mgmt    -
infra_vserver
            fcp_lif01a   data         fcp
infra_vserver
            fcp_lif02b   data         fcp
Five entries were displayed.
```

Note: The only interfaces that should remain are the cluster LIFs (clus1, clus2), node management LIFs (mgmt1), and SAN protocol LIFs (FCP and iSCSI).

2. Perform a storage takeover by the HA partner of the node being upgraded.

During Node 01 Upgrade	During Node 02 Upgrade
`storage failover takeover –bynode <<var_node02>>`	`storage failover takeover –bynode <<var_node01>>`

3. Using a console connection, observe that this node boots to the `LOADER>` prompt. After this has been confirmed, proceed to the next step.
4. Disconnect data cables from the controller being replaced. Shut down the controller. If the controller chassis contains no other controller, turn off the power supplies and unplug the cables from the power supplies.

Note: Make sure that cables are properly tagged prior to removal so that they can be properly recabled after the upgrade.

5. Remove any cards that need to be replaced, and install any new cards. PCIe cards with network ports should be replaced with similar models that have ports with the same port numbers.

Note: Refer to the appropriate document for the specific physical card installation procedure:

– Replacing PCIe cards in a 31xx system

- Replacing PCIe cards in a 32xx system
- Replacing PCIe cards in a 60xx system
- Replacing PCIe cads in a 62xx system

6. Recable the controller and if necessary the power supplies. Power on the controller and, if necessary, the power supplies.
7. The controller should boot to the `LOADER>` prompt. Using a console connection, boot Data ONTAP from the bootloader.

```
LOADER> boot_ontap
```

8. Verify the node that was upgraded is a member of the correct cluster and in good health.

```
Cluster show
```

Example output:

```
Node                  Health  Eligibility
--------------------- ------- ------------
node-01               true    true
node-02               true    true
entries were displayed.
```

9. Verify that storage failover is possible for both nodes.

```
                              Takeover
Node           Partner        Possible State
-------------- -------------- -------- -------------------------------------
node-01        node-02        true     Connected to node-02
node-02        node-01        true     Connected to node-01
2 entries were displayed.
```

10. Return to step 1 of this procedure to upgrade the card for controller 2. After both controllers' cards have been upgraded, proceed to step 11.
11. Determine which LIFs are not currently on their home node.

```
network interface show -is-home false
```

12. Revert NAS data LIFs, Vserver management LIFs, and cluster management LIFs to their appropriate home node.

```
network interface migrate -vserver <<var_vserver_name>> -lif <<var_lif_name>> -dest-node
<<var_new_node01>> -dest-port <<var_dest_dataport>>
```

Note: Some LIFs might need to be manually migrated to the correct controller. Use `network interface migrate` to manually move a LIF back to its proper controller.

13. Verify that the network interfaces are up.

```
network interface show -curr-node <<var_new_node01>>
```

3.3 Grow a Cluster

Table 6) List of variables.

Configuration Variable Name	Description
`<<var_node03_mgmt_ip>>`	IP address of the management interface of the third node in a cluster (first node of a new HA pair being added to an existing two-node cluster).
`<<var_node03_mgmt_mask>>`	Subnet mask of the management interface of the third node in a cluster (first node of a new HA pair being added to an existing two-node cluster).

Configuration Variable Name	Description
<<var_node03_mgmt_gateway>>	Gateway IP address of the management interface of the third node in a cluster (first node of a new HA pair being added to an existing two-node cluster).
<<var_node04_mgmt_ip>>	IP address of the management interface of the fourth node in a cluster (second node of a new HA pair being added to an existing two-node cluster).
<<var_node04_mgmt_mask>>	Subnet mask of the management interface of the fourth node in a cluster (second node of a new HA pair being added to an existing two-node cluster).
<<var_node04_mgmt_gateway>>	Gateway IP address of the management interface of the fourth node in a cluster (second node of a new HA pair being added to an existing two-node cluster).
<<var_url_boot_software>>	Path to the Data ONTAP upgrade package to be downloaded.
<<var_clustername>>	Name of the existing cluster to which the new HA pair will be added.
<<var_node01>>	Name of the first node of the cluster.
<<var_node02>>	Name of the second node of the cluster.
<<var_node03>>	Name of the third node of the cluster (first node of the new HA pair being added to the cluster).
<<var_node04>>	Name of the fourth node of the cluster (second node of the new HA pair being added to the cluster).
<<var_#_disks>>	Number of disks to be assigned to a node in the cluster. This document recommends that half of the total available disks be assigned to each node.
<<var_raidsize>>	Number of disks in a RAID group within the aggregates created by each new node.
<<var_num_raid_disks>>	Number of disks in the aggregates created for each new node.
<<var_fw_pkg_location>>	Location of the service processor firmware.
<<var_node03_sp_ip>>	IP address of the service processor interface of the third node in a cluster (first node of a new HA pair being added to an existing two-node cluster).
<<var_node03_sp_mask>>	Subnet mask of the service processor interface of the third node in a cluster (first node of a new HA pair being added to an existing two-node cluster).
<<var_node03_sp_gateway>>	Gateway IP address of the service processor interface of the third node in a cluster (first node of a new HA pair being added to an existing two-node cluster).
<<var_node04_sp_ip>>	IP address of the service processor interface of the fourth node in a cluster (second node of a new HA pair being added to an existing two-node cluster).
<<var_node04_sp_mask>>	Subnet mask of the service processor interface of the fourth node in a cluster (second node of a new HA pair being added to an existing two-node cluster).

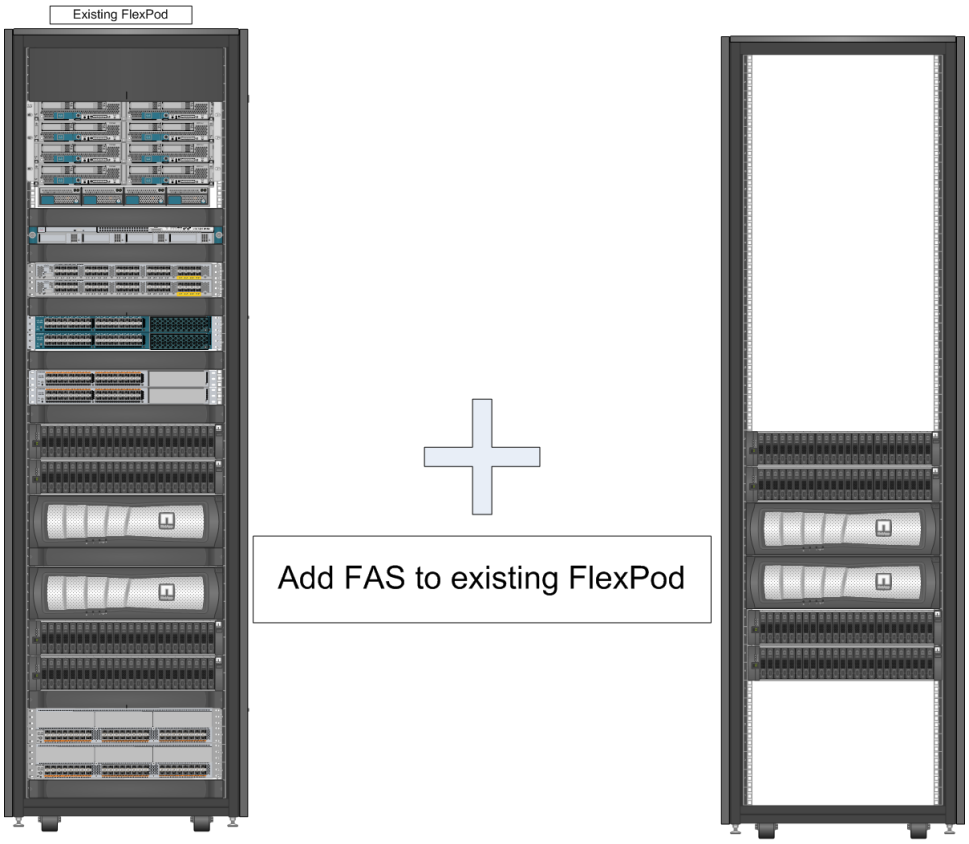
Configuration Variable Name	Description
<<var_node04_sp_gateway>>	Gateway IP address of the service processor interface of the fourth node in a cluster (second node of a new HA pair being added to an existing two-node cluster).
<<var_global_ntp_server_ip>>	IP address of your data center NTP server.
<<var_storage_admin_email>>	E-mail address of the storage administrator to be notified.
<<var_mailhost>>	IP address of the mail server to which e-mails will be sent.
<<var_nfs_vlan_id>>	VLAN number of the NFS VLAN in the FlexPod unit.
<<var_native_vlan>>	Native VLAN configured on trunked switch ports.
<<var_security_cert_node03_common_name>>	LDAP common name for node03.
<<var_country_code>>	LDAP country code corresponding to new node.
<<var_state>>	LDAP state corresponding to new node.
<<var_city>>	LDAP city corresponding to new node.
<<var_org>>	LDAP organization corresponding to new node.
<<var_unit>>	LDAP unit corresponding to new node.
<<var_vserver>>	Name of the Vserver.

The addition of a new HA pair to an existing cluster can be performed without affecting the current workloads on the cluster. Configuration of this new HA pair of controllers in the cluster to align with FlexPod technical specifications will also have no effect on the current workloads on the cluster.

Scenario

Two new controllers that will be configured as an HA pair will be joined to an existing cluster.

Figure 5) Adding an HA pair to a FlexPod system.

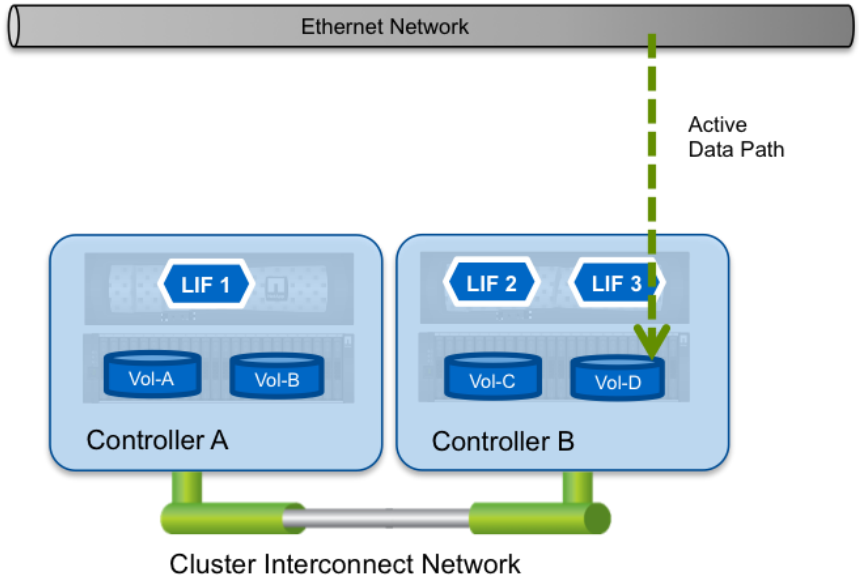


Ethernet Overview

Before

During normal operation, a clustered Data ONTAP system will have logical interfaces (LIFs) for NAS traffic configured and assigned to physical interfaces on the existing nodes in the cluster. A LIF can provide access to one or more volumes or LUNs within a Vserver. In the following example, the active data path from the Ethernet network, which could include traffic from the compute layer of the FlexPod unit or even outside resources, connects through a LIF that's assigned to an individual interface or ifgrp on controller B.

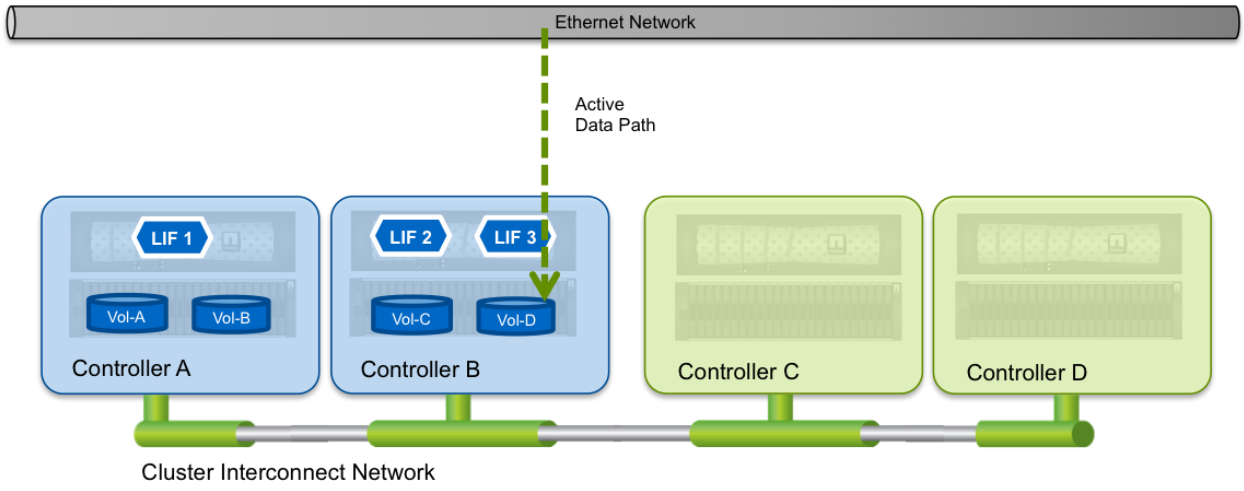
Figure 6) Ethernet before.



During

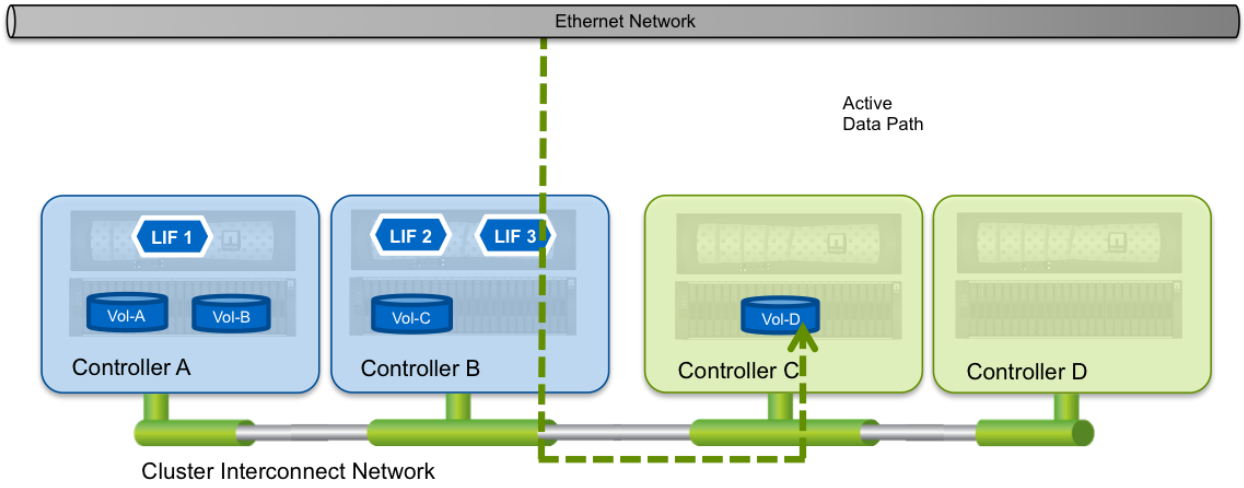
Adding an HA pair expands the cluster size, but does not automatically rebalance workloads within that cluster. In Figure 7, notice that the active data path stays the same, and the new nodes added to the cluster are not automatically utilized. Traffic and data will have to be manually moved to the new controllers to take advantage of the new resources in the cluster.

Figure 7) Ethernet during 1.



After the new controllers are active and online in the cluster, the administrator can start migrating LIFs to these new controllers. In this example, “LIF 3” was nondisruptively migrated to Controller C. This results in an unoptimized data path to the storage object (Vol-D in this case), because the traffic is entering the cluster on Controller C, but must travel over the cluster interconnect network to access data in Vol-D. Some intracluster traffic is fine, but the general best practice, especially for applications with low latency requirements, is to colocate the volumes or LUNs with the LIFs on the same controller to prevent this traffic from needing to cross the cluster interconnect.

Figure 8) Ethernet during 2.



Ethernet After

The final step in rebalancing this workload involves moving the volume to the new set of controllers. Depending on the size of the individual volume, the transfer can take minutes, hours, or more. This “vol move” migration traffic traverses the cluster interconnect network, so there is no impact on the front-end network that is accessed by the Cisco UCS servers. After the volume has finished moving nondisruptively, data access is once again local to the controller, reducing the workload on Controller B.

Figure 9) After.

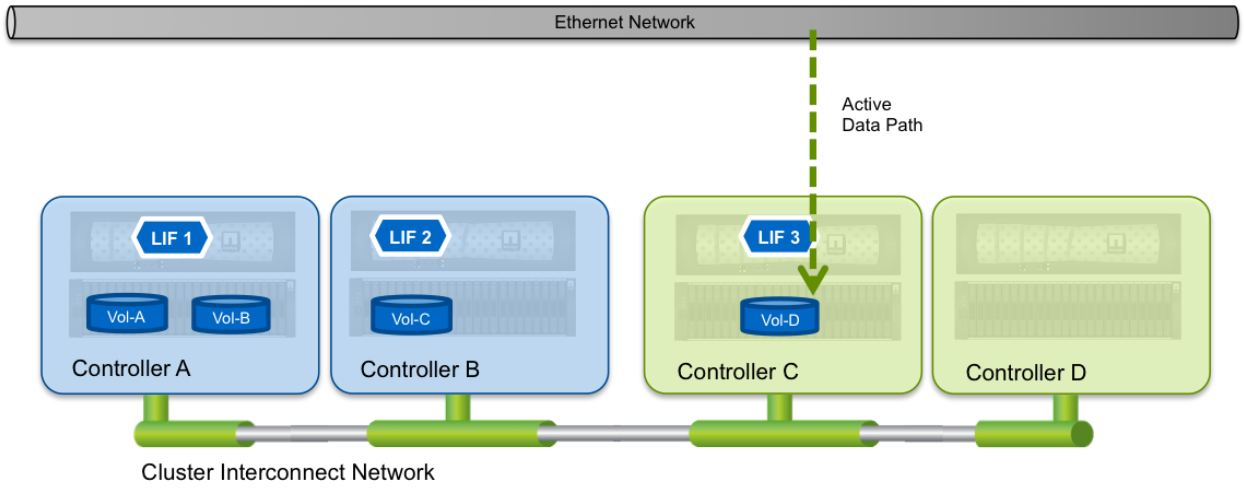


Table 7) Joining FAS to clustered Data ONTAP considerations.

Considerations	Description
Rack space/power/cooling to support new controller and/or additional shelves.	Refer to the NetApp Hardware Universe for detailed space, power, and cooling information.
Node count does not exceed maximum node count for Data ONTAP version and protocol.	Maximum node count is 24 nodes for clusters using NAS only and 8 nodes for clusters using SAN protocols. See page 97 of Clustered Data ONTAP 8.2 System Administration Guide for Cluster Administrators.

Considerations	Description
Make note of the version of Data ONTAP that the current cluster is running and that all controllers in the cluster are running the same version of Data ONTAP.	Use `version –node *` to display the current Data ONTAP version running on each node. Each node should report the same release version.

Add a New HA Pair (New from Factory) to an Existing Cluster

To join an existing clustered Data ONTAP cluster, complete the following steps:

1. Connect to the storage system console port. The console should display a `LOADER-A` prompt. However, if the storage system is in a reboot loop, press Ctrl-C to exit the autoboot loop when the following message displays:

```
Starting AUTOBOOT press Ctrl-C to abort
```

2. From the `LOADER-A` prompt:

```
printenv
```

3. If the `last-OS-booted-ver` parameter does not match the version of Data ONTAP displayed during the prerequisite steps, proceed to step 4 to load Data ONTAP 8.2 software. If Data ONTAP 8.2 is already loaded, proceed to step 16.

Note: This example will use clustered Data ONTAP 8.2P4.

4. Allow the system to boot.

```
boot_ontap
```

5. Press Ctrl-C when the `Press Ctrl-C for Boot Menu` message appears.

Note: If Data ONTAP 8.2 is not the version of software being booted, proceed with the following steps to install new software. If Data ONTAP 8.2 is the version being booted, then select option `8` and `yes` to reboot the node. Then proceed with step 15.

6. To install new software, first select option 7.

```
7
```

7. Answer yes to perform a nondisruptive upgrade.

```
y
```

8. Select e0M for the network port that should be used for the download. Port e0M should be connected to a network that has access to `<<var_url_boot_software>>`.

```
e0M
```

9. Select yes to reboot now.

```
y
```

10. Enter the IP address, netmask, and default gateway for e0M in their respective places.

During Node 3 Installation	During Node 4 Installation

`<<var_node03_mgmt_ip>>`
`<<var_node03_mgmt_mask>>`
`<<var_node03_mgmt_gateway>>`
	`<<var_node04_mgmt_ip>>`
	`<<var_node04_mgmt_mask>>`
	`<<var_node04_mgmt_gateway>>`

11. Enter the URL where the software can be found.

Note: This web server must be reachable. Test using the `ping` command.

```
<<var_url_boot_software>>
```

12. Press Enter for the username, indicating no user name.

```
Enter
```

13. Enter yes to set the newly installed software as the default to be used for subsequent reboots.

```
y
```

14. Enter yes to reboot the node.

```
y
```

Note: When installing new software, the system might perform firmware upgrades to the BIOS and adapter cards, causing reboots and possible stops at the LOADER prompt. If these actions occur, the system might deviate slightly from this procedure.

15. Press Ctrl-C to exit autoboot when this message is displayed:

```
Starting AUTOBOOT press Ctrl-C to abort…
```

16. From the `LOADER-A` prompt, enter:

```
printenv
```

Note: If `bootarg.init.boot_clustered true` is not listed, the system is not set to boot in clustered Data ONTAP.

17. If the system is not set to boot in clustered Data ONTAP, enter the following command at the LOADER prompt to make sure the system boots in clustered Data ONTAP:

```
setenv bootarg.init.boot_clustered true
setenv bootarg.bsdportname e0M
```

18. At the `LOADER-A` prompt, enter:

```
autoboot
```

19. Press Ctrl+C when the console displays `Press Ctrl-C for Boot Menu`:

```
Ctrl - C
```

20. Select option 4 to clean configuration and initialize all disks.

```
4
```

21. Answer yes to `Zero disks, reset config and install a new file system.`

```
y
```

22. Enter yes to erase all the data on the disks.

```
y
```

Note: The initialization and creation of the root volume can take 75 minutes or more to complete, depending on the number of disks attached. After initialization is complete, the storage system reboots. Continue to node 04 configuration while the disks for node 03 are zeroing. After both nodes have completed their initialization and rebooted, continue to the next steps.

23. During the node boot, the Cluster Setup wizard starts running on the console.

```
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}:
```

Note: If a login prompt displays instead of the Cluster Setup wizard, start the wizard by logging in using the factory default settings, and then enter the `cluster setup` command.

24. Enter the following command to join a cluster:

```
join
```

25. Follow the below steps to activate HA and set storage failover.

```
Non-HA mode, Reboot node to activate HA

Do you want to reboot now to set storage failover (SFO) to HA mode? {yes, no}
[yes]: Enter
```

26. After the reboot, proceed to join the cluster. Enter `Yes` to join the same cluster.

Note: During the boot of node 04, clustered Data ONTAP will detect that its storage failover partner (node 03) is part of a cluster. An example output is:

```
Existing cluster interface configuration found:

Port    MTU     IP              Netmask
e1a     9000    169.254.251.110 255.255.0.0
e2a     9000    169.254.56.206  255.255.0.0

Do you want to use this configuration? {yes, no} [yes]: Enter
```

Note: The cluster creation can take a minute or two.

27. Enter the cluster name and press Enter.

```
Enter the name of the cluster you would like to join [<<var_clustername>>]:Enter
```

Note: The node should find the cluster name and provide it as the cluster to join.

28. Set up the node.

During Node 03 Setup	During Node 04 Setup
Enter the node management interface port [e0M]: e0b Enter the node management interface IP address: <<var_node03_mgmt_ip>> Enter the node management interface netmask: Enter Enter the node management interface default gateway: Enter	Enter the node management interface port [e0M]: e0b Enter the node management interface IP address: <<var_node04_mgmt_ip>> Enter the node management interface netmask: Enter Enter the node management interface default gateway: Enter

Note: The node management interface should be in a subnet different from the cluster management interface. The node management interfaces can reside on the out-of-band management network, and the cluster management interface can be on the in-band management network.

29. Press Enter to accept the AutoSupport message.

30. Log in to the cluster management interface with the “admin” user ID and `<<var_password>>`.

31. Reboot the node.

During Node 03 Setup	During Node 04 Setup
system node reboot <<var_node03>> y	system node reboot <<var_node04>> y

32. After rebooting the node, connect to the controller’s console port.

33. When the terminal console displays `Press Ctrl-C for Boot Menu`, enter:

```
Ctrl – C
```

34. Select 5 to boot into maintenance mode.

```
5
```

35. At the question `Continue with boot?` enter:

```
y
```

36. To verify the HA status of your environment, enter:

```
ha-config show
```

Note: If either component is not in HA mode, use the `ha-config modify` command to put the components in HA mode.

37. To see disks attached to the system, enter:

```
disk show –a
```

Note: This list will also specify ownership. A brand new system should not have any disks owned.

38. Assign disks.

```
disk assign –n <<var_#_of_disks>>
```

Note: In this example we assign half the disks to each controller. Workload design could, however, dictate different percentages.

39. Reboot the controller:

```
halt
```

40. At the `LOADER-A` prompt, enter:

```
autoboot
```

41. Open an SSH connection to cluster IP or host name and log in as “admin” user with the password provided earlier.
42. Zero all spare disks in the cluster.

```
disk zerospares
```

43. Add the new node management ports to the existing failover group.

```
network interface failover-groups create -failover-group mgmt -node <<var_node03>> -port e0a
network interface failover-groups create -failover-group mgmt -node <<var_node04>> -port e0a
```

Note: In this example, the in-band failover group for the cluster management LIF is named mgmt. This will allow the cluster management LIF to fail over to any of the in-band management ports on all four nodes.

44. Verify that the cluster management LIF has been assigned to the mgmt failover group. The “failover policy” column should display “nextavail” or “priority,” the “Use Failover” Group column should display “enabled,” and the “Failover Group” column should display “mgmtchr” in this example.

```
network interface show -lif cluster_mgmt -fields failover-group,failover-policy,use-failover-
group
```

45. Create node management port failover groups for each node.

```
network interface failover-groups create -failover-group node-mgmt03 -node <<var_node03>> -port
e0b
network interface failover-groups create -failover-group node-mgmt03 -node <<var_node03>> -port
e0M
network interface failover-groups create -failover-group node-mgmt04 -node <<var_node04>> -port
e0b
network interface failover-groups create -failover-group node-mgmt04 -node <<var_node04>> -port
e0M
```

46. Assign the management port failover groups to their respective node management LIFs.

```
network interface modify -vserver <<var_node03>> -lif mgmt1 -auto-revert true -failover-group
node-mgmt03
network interface modify –vserver <<var_node04>> -lif mgmt1 –auto-revert true –failover-group
node-mgmt04
```

Flash Cache

If the new nodes contain Flash Cache cards, FlexScale™ must be enabled on the nodes to use the cards. Complete the following steps to enable Flash Cache on each node:

1. Run the following commands from the cluster management interface:

```
system node run -node <<var_node03>> options flexscale.enable on
system node run -node <<var_node03>> options flexscale.lopri_blocks off
system node run -node <<var_node03>> options flexscale.normal_data_blocks on
system node run -node <<var_node04>> options flexscale.enable on
system node run -node <<var_node04>> options flexscale.lopri_blocks off
system node run -node <<var_node04>> options flexscale.normal_data_blocks on
```

Note: Data ONTAP 8.2 and later does not require a separate license for Flash Cache.

Note: For directions on how to configure Flash Cache in metadata mode or low-priority data caching mode, refer to TR-3832: Flash Cache Best Practices Guide. Before customizing the settings, determine whether the custom settings are required or if the default settings are sufficient.

2. Execute the following command to create new aggregates:

```
aggr create -aggregate aggr03 -nodes <<var_node03>> -maxraidsize <<var_raidsize>> -diskcount
<<var_num_raid_disks>>
aggr create -aggregate aggr04 -nodes <<var_node04>> -maxraidsize <<var_raidsize>> -diskcount
<<var_num_raid_disks >>
```

Note: Retain at least one disk (select the largest disk) in the configuration as a spare. A best practice is to have at least one spare for each disk type and size.

Note: Calculate the RAID group size to allow for roughly balanced (same size) RAID groups of from 12 through 20 disks (for SAS disks) within the aggregate. For example, if 52 disks were being assigned to the aggregate, select a RAID group size of 18. A RAID group size of 18 would yield two 18-disk RAID groups and one 16-disk RAID group. Keep in mind that the default RAID group size is 16 disks, and that the larger the RAID group size, the longer the disk rebuild time in case of a failure.

Note: The aggregate cannot be created until disk zeroing completes. Use the `aggr show` command to display aggregate creation status. Do not proceed until both aggr03 and aggr04 are online.

3. Disable Snapshot copies for the two data aggregates just created.

```
node run <<var_node03>> aggr options aggr03 nosnap on
node run <<var_node04>> aggr options aggr04 nosnap on
```

4. Delete any existing Snapshot copies for the two data aggregates.

```
node run <<var_node03>> snap delete –A –a –f aggr03
node run <<var_node04>> snap delete –A –a –f aggr04
```

Service Processor

Gather information about the network and the AutoSupport settings before configuring the service processor (SP).

Configure the SP using DHCP or static addressing. If the SP uses a static IP address, verify that the following SP prerequisites have been met:

- An available static IP address
- The network netmask

- The network gateway IP
- AutoSupport information

A best practice is to configure the AutoSupport recipients and mail host before configuring the SP. Data ONTAP automatically sends AutoSupport configuration to the SP, allowing the SP to send alerts and notifications through an AutoSupport message to the system administrative recipients specified in AutoSupport. When configuring the SP, enter the name or the IP address of the AutoSupport mail host when prompted.

The service processor must be set up on each node.

1. Check the configuration and version of the service processor.

```
system node service-processor show
```

2. Using a web browser, connect to http://support.netapp.com/NOW/cgi-bin/fw. The complete instructions and release notes can be found here.
3. Navigate to the “Service Process Image for installation from the Data ONTAP prompt” page for your storage platform. This document will use the instructions for the FAS3270 model.
4. Check the latest firmware version that is available for your storage platform. If your storage system is not running the latest version, proceed to the download page for the latest release of the SP firmware for your storage platform.
5. Update the SPs on both nodes. Download the .zip file to a web server that is reachable from the cluster management interface. Issue the following command to download the firmware file to the nodes:

```
system node image get -node <<var_node03>> -package <<var_fw_pkg_location>> -replace-package true
system node image get -node <<var_node04>> -package <<var_fw_pkg_location>> -replace-package true
```

6. Verify that the nodes have been updated to the latest service processor.

```
system node service-processor image update-progress show
```

Example output:

```
                  In                              Percent
Node              Progress Start Time              Done    End Time
----------------- -------- ---------------------- ------- --------------------
Node1             no       -                       0       -
Node2             no       -                       0       -
Node3             yes      10/13/2013 20:00:34     99      -
Node4             no       -                       0       -
2 entries were displayed.
```

7. Verify that the service processors are online and running the correct firmware version.

```
system node service-processor show
```

Example output:

```
                             IP           Firmware
Node          Type Status    Configured   Version   IP Address
------------- ---- --------- ------------ --------- ----------------------
node1         SP   online    true         1.4P1     10.251.133.139
node2         SP   online    true         1.4P1     10.251.133.140
node3         SP   online    true         1.4P1     10.251.133.141
node4         SP   online    true         1.4P1     10.251.133.142
```

8. From the cluster shell, enter the following command to configure the service processor on the new node:

During Node 03 Setup	During Node 04 Setup
`system node run <<var_node03>> sp setup`	`system node run <<var_node04>> sp setup`

9. Set up the service processor.

During Node 03 Setup	During Node 04 Setup

`Would you like to configure the SP? Y`
`Would you like to enable DHCP on the SP LAN interface? no`
`Please enter the IP address of the SP[]: <<var_node03_sp_ip>>`
`Please enter the netmask of the SP[]: <<var_node03_sp_mask>>`
`Please enter the IP address for the SP gateway[]: <<var_node03_sp_gateway>>`
	`Would you like to configure the SP? Y`
	`Would you like to enable DHCP on the SP LAN interface? no`
	`Please enter the IP address of the SP[]: <<var_node04_sp_ip>>`
	`Please enter the netmask of the SP[]: <<var_node04_sp_mask>>`
	`Please enter the IP address for the SP gateway[]: <<var_node04_sp_gateway>>`

10. Enable failover on one of the two nodes.

```
storage failover modify -node <<var_node03>> -enabled true
```

11. Configure the Network Time Protocol (NTP) for each node in the cluster.

```
system services ntp server create -node <<var_node03>> -server <<var_global_ntp_server_ip>>
system services ntp server create -node <<var_node04>> -server <<var_global_ntp_server_ip>>
```

12. Configure AutoSupport.

```
system node autosupport modify -node * -state enable –mail-hosts <<var_mailhost>> -transport
https -support enable -noteto <<var_storage_admin_email>>
```

13. CDP is recommended on all nodes in a FlexPod cluster. Enable CDP using the following command.

```
node run -node <<var_node03>> options cdpd.enable on
node run -node <<var_node04>> options cdpd.enable on
```

14. Perform steps 14–17 for each Vserver in the cluster. Create a volume to be the load-sharing mirror of each Vserver in the cluster.

```
volume create –vserver <<var_vserver>> –volume root_vol_m03 –aggregate aggr03 –size 20MB –type DP
volume create –vserver <<var_vserver>> –volume root_vol_m04 –aggregate aggr04 –size 20MB –type DP
```

15. Create the mirroring relationships for each volume.

```
snapmirror create –source-path //<<var_vserver>>/rootvol –destination-path
//<<var_vserver>>/root_vol_m03 –type LS
snapmirror create –source-path //<<var_vserver>>/rootvol –destination-path
//<<var_vserver>>/root_vol_m04 –type LS
```

16. Initialize the mirroring relationship.

```
snapmirror initialize-ls-set –source-path //<<var_vserver>>/rootvol
```

17. Set an hourly (at 5 minutes past the hour) update schedule on each mirroring relationship.

```
snapmirror modify –source-path //<<var_vserver>>/rootvol –destination-path * -schedule hourly
```

18. Run the following commands as one-time commands to generate and install self-signed certificates:

```
security certificate create -vserver <<var_node03>> -common-name
<<var_security_cert_node03_common_name>> -size 2048 -country <<var_country_code>> -state
<<var_state>> -locality <<var_city>> -organization <<var_org>> -unit <<var_unit>> -email-addr
<<var_storage_admin_email>>
security certificate create -vserver <<var_node04>> -common-name
<<var_security_cert_node02_common_name>> -size 2048 -country <<var_country_code>> -state
<<var_state>> -locality <<var_city>> -organization <<var_org>> -unit <<var_unit>> -email-addr
<<var_storage_admin_email>>
```

19. To enable workload rebalancing, the new nodes should have ifgrps and VLAN interfaces that match the existing nodes. Run the following commands on the command line to create interface groups (ifgrps).

```
ifgrp create -node <<var_node03>> -ifgrp a0a -distr-func port -mode multimode_lacp
network port ifgrp add-port -node <<var_node03>> -ifgrp a0a -port e3a
network port ifgrp add-port -node <<var_node03>> -ifgrp a0a -port e4a
ifgrp create -node <<var_node04>> -ifgrp a0a -distr-func port -mode multimode_lacp
network port ifgrp add-port -node <<var_node04>> -ifgrp a0a -port e3a
network port ifgrp add-port -node <<var_node04>> -ifgrp a0a -port e4a
```

Note: All interfaces must be in the `down` status before being added to an interface group.

Note: The interface group name must follow the standard naming convention of “a<number><letter>,” where <number> is an integer in the range 0–999 without leading zeros and <letter> is a lowercase letter.

20. Create NFS VLANs.

```
network port vlan create –node <<var_node03>> -vlan-name a0a-<<var_nfs_vlan_id>>
network port vlan create –node <<var_node04>> -vlan-name a0a-<<var_nfs_vlan_id>>
```

21. Add the new NFS ports to the existing NFS failover groups.

```
network interface failover-groups create -failover-group nfs -node <<var_node03>> -port a0a-
<<var_nfs_vlan_id>>
network interface failover-groups create -failover-group nfs -node <<var_node04>> -port a0a-
<<var_nfs_vlan_id>>
```

22. To configure a clustered Data ONTAP network port to use jumbo frames (which usually have an MTU of 9,000 bytes), run the following command from the cluster shell:

```
network port modify -node <<var_node03>> -port a0a -mtu 9000

WARNING: Changing the network port settings will cause a serveral second interruption in carrier.
Do you want to continue? {y|n}: y

network port modify –node <<var_node03>> -port a0a-<<var_nfs_vlan_id>> -mtu 9000

WARNING: Changing the network port settings will cause a serveral second interruption in carrier.
Do you want to continue? {y|n}: y

network port modify –node <<var_node04>> -port a0a -mtu 9000

WARNING: Changing the network port settings will cause a serveral second interruption in carrier.
Do you want to continue? {y|n}: y

network port modify –node <<var_node04>> -port a0a-<<var_nfs_vlan_id>> -mtu 9000

WARNING: Changing the network port settings will cause a serveral second interruption in carrier.
Do you want to continue? {y|n}: y
```

Note: Because nodes are being added to the existing FlexPod unit, new VLANs might not be required. The Cisco Nexus data switch interfaces should have descriptions that correspond to the nodes and ports to which they are connected to assist with troubleshooting.

Cisco Nexus Port Configuration

1. Configure port descriptions.

On Switch 1	On Switch 2
conf t int e1/13 desc <<var_node03>>:e3a int e1/14 desc <<var_node04>>:e3a	conf t int e1/13 desc <<var_node03>>:e4a int e1/14 desc <<var_node04>>:e4a

2. The Cisco Nexus data switches should be configured to match the node interface groups and VLAN trunking configuration. Configure new port channels and virtual port channels corresponding to the new nodes using the following commands:

On Switch 1	On Switch 2
conf t int po113 desc <<var_node03>> int po114 desc <<var_node04>> int e1/13 channel-group 113 mode active no shutdown int e 1/14 channel-group 114 mode active no shutdown int po 113 switchport mode trunk switchport trunk native vlan <<var_native_vlan>> switchport trunk allowed vlan <<var_nfs_vlan_id>> spanning-tree port type edge trunk no shutdown vpc 113 int po 114 switchport mode trunk switchport trunk native vlan <<var_native_vlan>> switchport trunk allowed vlan <<var_nfs_vlan_id>> spanning-tree port type edge trunk no shutdown vpc 114 exit copy run start	conf t int po113 desc <<var_node03>> int po114 desc <<var_node04>> int e1/13 channel-group 113 mode active no shutdown int e 1/14 channel-group 114 mode active no shutdown int po 113 switchport mode trunk switchport trunk native vlan <<var_native_vlan>> switchport trunk allowed vlan <<var_nfs_vlan_id>> spanning-tree port type edge trunk no shutdown vpc 113 int po 114 switchport mode trunk switchport trunk native vlan <<var_native_vlan>> switchport trunk allowed vlan <<var_nfs_vlan_id>> spanning-tree port type edge trunk no shutdown vpc 114 exit copy run start

For additional workload rebalancing steps, see section 0, “Workload Rebalancing.”

3.4 Add Clustered Data ONTAP FAS to 7-Mode FlexPod

Table 8) Configuration variables.

Variable	Description
<<var_cluster_switch_password>>	Admin password for the cluster interconnect switches.
<<var_cluster_switch_name>>	Name of the cluster interconnect switch.
<<var_cluster_switch_mgmt0_ip>>	IP address of the cluster interconnect switch mgmt0. This should be different for each switch.
<<var_cluster_switch_mgmt0_netmask>>	Network mask for the cluster interconnect switch mgmt0.
<<var_cluser_switch_mgmt0_gw>>	Default gateway of the cluster interconnect switch mgmt0.
<<var_global_ntp_server_ip>>	IP address of the NTP server used in your data center.
<<var_transfer_protocol>>	Transfer protocol used for downloading a config file to a cluster interconnect switch. Common protocols are TFTP/FTP.
<<var_config_file_name>>	Name of the NetApp cluster interconnect config file for Cisco Nexus 5596 cluster interconnects.

Variable	Description
<<var_network_admin_email>>	E-mail address of the administrator responsible for the FlexPod network.
<<var_network_admin_phone>>	Phone number of the administrator responsible for the FlexPod network.
<<var_network_admin_address>>	Address of the administrator responsible for the FlexPod network.
<<var_mailhost>>	The SMTP server to which support e-mails should be sent.
<<var_vrf_name_mail>>	The VRF context to be used for outgoing e-mail messages from the cluster interconnect switch.
<<var_snmp_server_ip>>	IP address of the SNMP trap host in your data center.
<<var_pkg_location>>	Path to the Data ONTAP upgrade package to be downloaded. This can be an HTTP, FTP, or TFTP URL or the path of a mounted file system.
<<var_new_node01>>	The first controller in an HA pair of controllers in a cluster.
<<var_new_node02>>	The second controller in an HA pair of controllers in a cluster.
<<var_new_node01_mgmt_ip>>	IP address of the management interface of the first node of a new HA pair in a new cluster.
<<var_new_node01_mgmt_mask>>	Subnet mask of the management interface of the first node of a new HA pair in a new cluster.
<<var_new_node01_mgmt_gateway>>	Gateway IP address of the management interface of the first node of a new HA pair in a new cluster.
<<var_new_node02_mgmt_ip>>	IP address of the management interface of the second node of a new HA pair in a new cluster.
<<var_new_node02_mgmt_mask>>	Subnet mask of the management interface of the second node of a new HA pair in a new cluster.
<<var_new_node02_mgmt_gateway>>	Gateway IP address of the management interface of the second node of a new HA pair in a new cluster.
<<var_url_boot_software>>	URL of the Data ONTAP 8.2 P4 software package, reachable by the management interfaces of the new nodes.
<<var_clustername>>	Name of the new cluster.
<<var_cluster_base_license_key>>	Base cluster license for the new cluster.
<<var_password>>	Cluster administrator password.
<<var_clustermgmt_ip>>	Cluster management interface IP address.
<<var_clustermgmt_mask>>	Subnet mask of the cluster management interface.
<<var_clusermgmt_gateway>>	IP gateway of the cluster management interface.
<<var_dns_domain_name>>	Domain name of the DNS server that the cluster will use to resolve domain names.

Variable	Description
<<var_nameserver_ip>>	IP address of the DNS server that the cluster will use to resolve domain names.
<<var_node_location>>	Physical location of the new cluster.
<<var_#_disks>>	Number of disks to be assigned to a node in the cluster. This document recommends that half of the total available disks be assigned to each node.
<<var_num_disks>>	Number of disks to be assigned to aggregates in the node. This document recommends that one disk be available as a spare.
<<var_raidsize>>	Number of disks in a RAID group within the aggregates created by each new node.
<<var_num_raid_disks>>	Number of disks in the aggregates created for each new node.
<<var_sp_fw_pkg_location>>	Location of the service processor firmware.
<<var_ new_node01_sp_ip>>	IP address of the service processor interface of the third node in a cluster (first node of a new HA pair being added to an existing two-node cluster).
<<var_ new_node01_sp_mask>>	Subnet mask of the service processor interface of the third node in a cluster (first node of a new HA pair being added to an existing two-node cluster).
<<var_ new_node01_sp_gateway>>	Gateway IP address of the service processor interface of the third node in a cluster (first node of a new HA pair being added to an existing two-node cluster).
<<var_ new_node02_sp_ip>>	IP address of the service processor interface of the fourth node in a cluster (second node of a new HA pair being added to an existing two-node cluster).
<<var_ new_node02_sp_mask>>	Subnet mask of the service processor interface of the fourth node in a cluster (second node of a new HA pair being added to an existing two-node cluster).
<<var_new_node02_sp_gateway>>	Gateway IP address of the service processor interface of the fourth node in a cluster (second node of a new HA pair being added to an existing two-node cluster).
<<var_timezone>>	Time zone for the cluster.
<<var_global_ntp_server_ip>>	IP address of the data center NTP server.
<<var_snmp_server_contact>>	SNMP contact.
<<var_snmp_server_location>>	SNMP location.
<<var_snmp_community>>	SNMP v1 community name.
<<var_storage_admin_email>>	E-mail address of the storage administrator to be notified.
<<var_mailhost>>	IP address of the mail server to which e-mails will be sent.
<<var_nfs_vlan_id>>	VLAN number of the NFS VLAN in your FlexPod unit.

Variable	Description
<<var_native_vlan>>	The native VLAN configured on trunked switch ports.
<<var_security_cert_node03_common_name>>	The LDAP common name for node03.
<<var_country_code>>	LDAP country code corresponding to new node.
<<var_state>>	LDAP state corresponding to new node.
<<var_city>>	LDAP city corresponding to new node.
<<var_org>>	LDAP organization corresponding to new node.
<<var_unit>>	LDAP unit corresponding to new node.
<<var_security_cert_vserver_common_name>>	Common name for Infra_Vserver for HTTPS access.
<<var_security_certificate_vserver_authority>>	Security certificate for Infra_Vserver for HTTPS access.
<<var_security_certificate_vserver_serial_no>>	Serial number for Infra_Vserver for HTTPS access.
<<var_security_cert_cluster_common_name>>	Common name for cluster for HTTPS access.
<<var_security_certificate_cluster_authority>>	Security certificate for cluster for HTTPS access.
<<var_security_certificate_cluster_serial_no>>	Serial number for cluster for HTTPS access.
<<var_security_cert_new_node01_common_name>>	Common name for first node for HTTPS access.
<<var_security_certificate_new_node01_authority>>	Security certificate for first node for HTTPS access.
<<var_security_certificate_new_node01_serial_no>>	Serial number for first node for HTTPS access.
<<var_security_cert_new_node02_common_name>>	Common name for second node for HTTPS access.
<<var_security_certificate_new_node02_authority>>	Security certificate for second node for HTTPS access.
<<var_security_certificate_new_node02_serial_no>>	Serial number for second node for HTTPS access.
<<var_vserver_mgmt_ip>>	Infra_Vserver management IP address.
<<var_vserver_mgmt_mask>>	Infra_Vserver management interface subnet mask.
<<var_vsadmin_password>>	Password for the administrator for Infra_Vserver.
<<var_vserver>>	Name of the Vserver.

Scenario

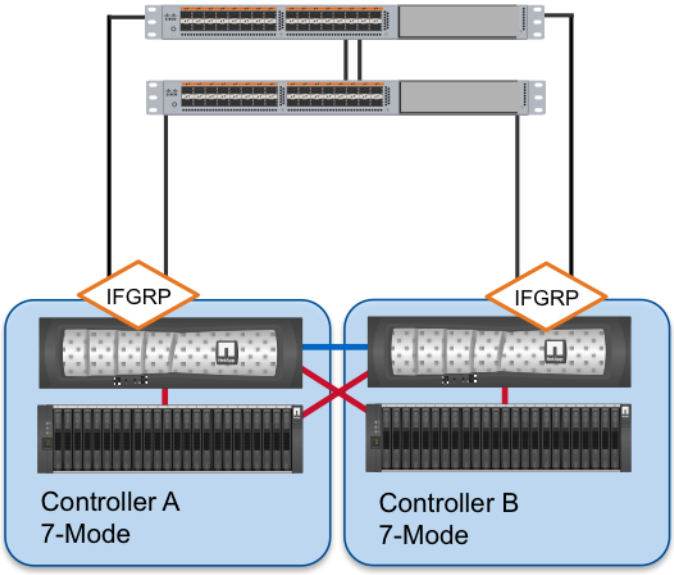
In an existing FlexPod unit that contains a FAS HA pair running any Data ONTAP 7-Mode version, a new FAS HA pair running clustered Data ONTAP 8.2 can be nondisruptively added to the existing stack without affecting the existing workloads. The two pairs within the FlexPod unit will operate independently of each other.

Overview

Before

A single HA pair running Data ONTAP 7-Mode is shown connected in a typical FlexPod configuration. Interface groups (ifgrps) are used to provide active-active connections to redundant switches in the architecture, and HA is set up between the controllers to allow them to take over and serve data from the other's disks if a controller should fail.

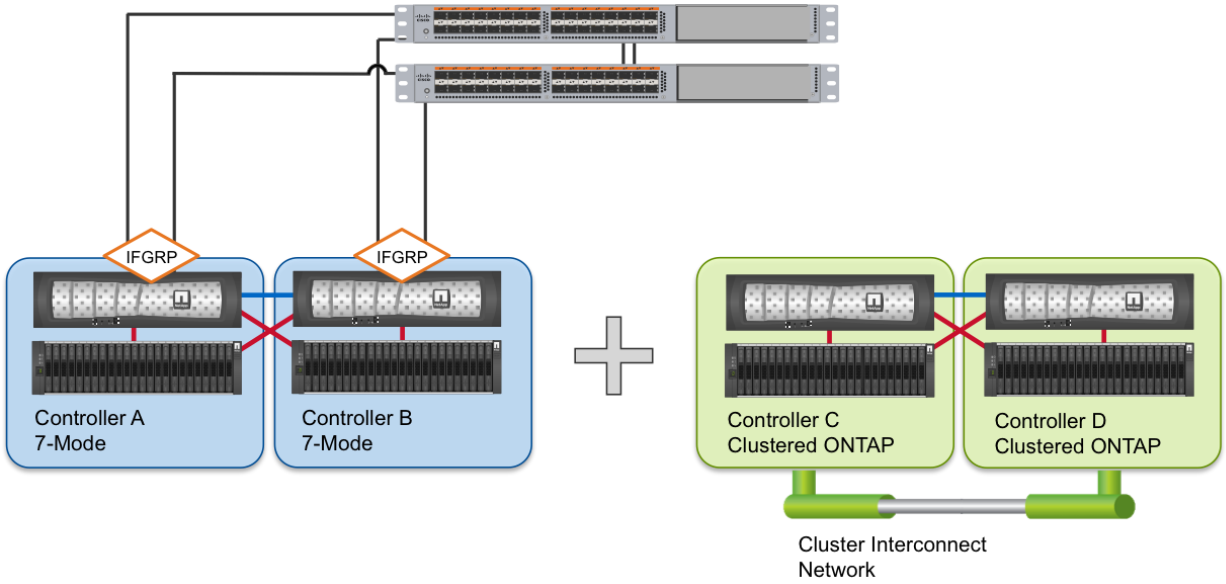
Figure 10) Single HA pair in 7-Mode.



During

Adding a clustered Data ONTAP system into an existing FlexPod system with controllers running 7-Mode will not affect the active workloads and is a fully supported operation. This is useful for deployments that might be migrating to clustered Data ONTAP, but still have data or requirements in a 7-Mode environment.

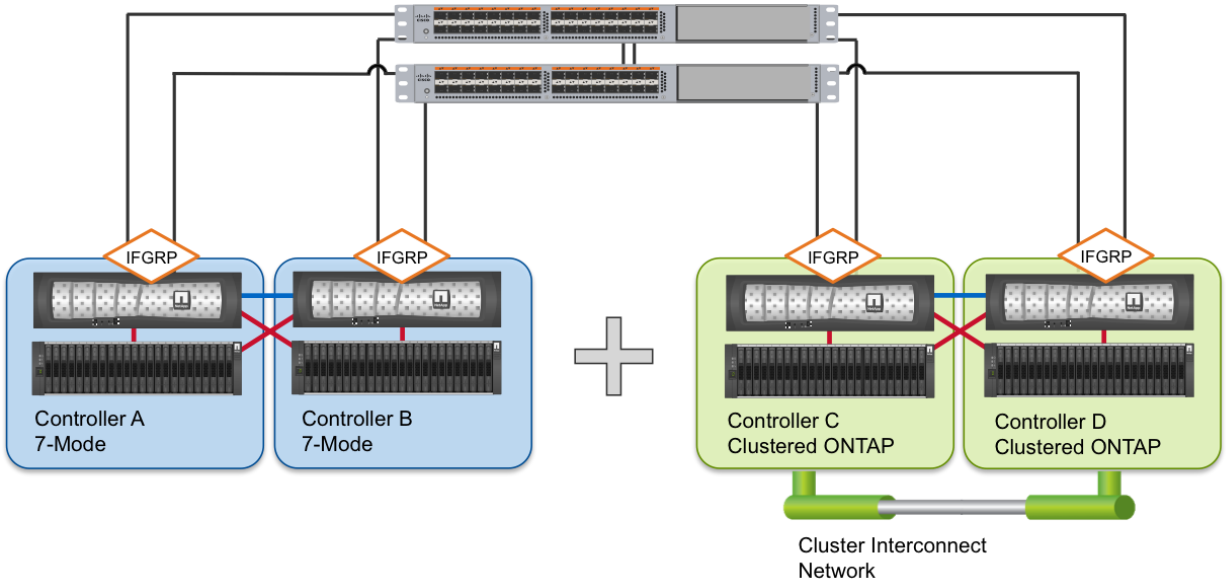
Figure 11) Data ONTAP 7-Mode HA pair and clustered Data ONTAP HA pair.



After

The administrator connects the clustered Data ONTAP system into the FlexPod system by cabling the appropriate Ethernet and SAN ports into the network. In this configuration, the network and server components can be shared between both the 7-Mode and the clustered Data ONTAP systems. Both storage systems will operate independently of each other.

Figure 12) Data ONTAP 7-Mode and clustered Data ONTAP HA pair online.



Technical Upgrade Procedure

Table 9) Upgrade procedure.

Dependencies	Description
Verify that existing Cisco Nexus data switches in the FlexPod unit have enough ports to accommodate the addition of a new FAS controller pair.	Each switch must have two available 10GbE converged network ports or two 10GbE ports and two Fibre Channel ports. This example uses two 10GbE ports.
Verify that the environment has enough available rack space and power and cooling accommodations to support new hardware such as cluster interconnect switches, controllers, and shelves.	See the Site Requirements Guide for additional detail related to your particular controller and disk shelf models.
Verify that the new disk shelves are properly connected to the new controllers.	If using SAS disk shelves, to determine the correct cabling method, refer to the SAS Disk Shelves Universal SAS and ACP Cabling Guide. Other disk shelf types and installation guides can be found at the NetApp Support site.

Note: In this example, the new clustered Data ONTAP pair will utilize two Cisco Nexus 5596 switches as cluster interconnects. The NetApp C1610 cluster interconnect switch is also supported within FlexPod. See the CN1610 Switch Setup and Configuration Guide for details on utilizing this option. If using cluster interconnect switches, dedicated cluster interconnect switches must be used. The existing Cisco Nexus 5000 switch in the FlexPod unit cannot be used for cluster interconnect traffic.

Note: If using a switchless cluster configuration, follow section 0, “Creating a Switchless Cluster,” and then skip to section “Initializing Clustered Data ONTAP 8.2” of this document.

Cisco Nexus 5596 Cluster Network Switch Configuration

Table 10) Cisco Nexus 5596 cluster network switch configuration prerequisites.

Description
• Rack and connect power to the new Cisco Nexus 5596 switches • Provide a terminal session that connects to the switch's serial console port (9600, 8, n, 1) • Connect the `mgmt0` port to the management network and be prepared to provide IP address information • Obtain password for admin • Determine switch name • Identify SSH key type (dsa, rsa, or rsa1) • Set up an e-mail server for Cisco Smart Call Home and IP connectivity between the switch and the e-mail server • Provide SNMP contact information for Cisco Smart Call Home (name, phone, street address) • Identify a Cisco.com ID associated with an appropriate Cisco SMARTnet® Service contract for Cisco Smart Call Home • Enable Cisco SMARTnet Service for the device to be registered for Cisco Smart Call Home

Initial Setup of Cisco Nexus 5596 Cluster Interconnect

The first time a Cisco Nexus 5596 cluster interconnect is accessed, it runs a setup program that prompts the user to enter an IP address and other configuration information needed for the switch to communicate over the management Ethernet interface. This information is required to configure and manage the switch. If the configuration must be changed later, the setup wizard can be accessed again by running the `setup` command in EXEC mode.

To set up the Cisco Nexus 5596 cluster interconnect, complete the following steps on both cluster interconnects.

1. Connect to the console of the Cisco Nexus switch.
2. Provide applicable responses to the setup prompts displayed on the Cisco Nexus 5596 cluster interconnect.

```
Abort Power On Auto Provisioning and continue with normal setup ?(yes/no)[n]: yes
Do you want to enforce secure password standard (yes/no): yes
Enter the password for "admin": <<var_cluster_switch_password>>
Confirm the password for "admin": <<var_cluster_switch_password>>
Would you like to enter the basic configuration dialog (yes/no):yes
Create another login account (yes/no) [n]: Enter
Configure read-only SNMP community string (yes/no) [n]: Enter
Configure read-write SNMP community string (yes/no) [n]: Enter
Enter the switch name : <<var_cluster_switch_name>>
Continue with Out-of-band (mgmt0) management configuration? (yes/no) [y]: Enter
Mgmt0 IPv4 address : <<var_cluster_switch mgmt0_ip>>
Mgmt0 IPv4 netmask : <<var_cluster_switch_mgmt0_netmask>>
Configure the default gateway? (yes/no) [y]: Enter
IPv4 address of the default gateway : <<var_cluster_switch_mgmt0_gw>>
Enable the telnet service? (yes/no) [n]: Enter
Enable the ssh service? (yes/no) [y]: Enter
Type of ssh key you would like to generate (dsa/rsa) : rsa
Number of  key bits <768-2048> : 1024
Configure the ntp server? (yes/no) [n]: yes
NTP server IPv4 address : <<var_global_ntp_server_ip>>
Enter basic FC configurations (yes/no) [n]: Enter
```

3. At the end of the setup, the configuration choices are displayed. Verify the information and save the configuration at this time.

```
Would you like to edit the configuration? (yes/no) [n]: <n>
Use this configuration and save it? (yes/no) [y]: <y>
```

Download and Install NetApp Cluster Switch Software

When the Cisco Nexus 5596 is being used as a cluster network switch with Data ONTAP 8.2, it should be running NX-OS version 5.2(1)N1(1).

The `show version` command from the switch command line interface will show the switch version currently running on the switch. If the currently running version is not 5.2(1)N1(1), go to the NetApp Support site and download and install NX-OS 5.2(1)N1(1) for the Cisco Nexus 5596 switch. Make sure both cluster interconnects are running NX-OS version 5.2(1)N1(1).

Download and Merge of NetApp Cluster Switch Reference Configuration File

Cluster network and management network switches are shipped without the configuration files installed. These files must be downloaded to the switches during deployment. Configuration files must be downloaded when the cluster network and management network switches are first installed or after the Cisco switch software is updated or reinstalled.

After the initial setup is complete, the NetApp cluster network switch reference configuration must be transferred to the switch and merged with the existing configuration. Instructions for this task and the reference configuration files for the appropriate switches are available on the NetApp Support site.

The latest Cisco Nexus OS version and reference configuration file should be used for the appropriate switch model.

After the switch installation has been completed, the Config Advisor tool can be used to confirm that the switch is running the correct OS and that the correct configuration file has been applied. See the Config Advisor site for details on how to download and use this tool.

To download configuration files to a host and install them on a Cisco Nexus 5596 switch, complete the following steps on both cluster interconnects:

1. Obtain a console connection to the switch and log into the switch with administrative access.
2. Verify the existing configuration.

```
show run
```

3. Make sure that the switch is reachable on the network.
4. Copy the configuration files to the bootflash of the switch.

```
copy <<var_transfer_protocol>>: bootflash: vrf management
```

5. Verify that the configuration file is downloaded.

```
***** Transfer of file Completed Successfully *****
Copy complete, now saving to disk (please wait)...
```

6. View the saved configuration file.

```
dir bootflash:
```

7. Merge the configuration file into the existing `running-config`.

```
copy <<var_config_file_name>> running-config
```

Note: A series of warnings regarding "`portfast`" are displayed as each port is configured.

8. Verify the success of the configuration merge by running the `show run` command and comparing its output to the contents of the configuration file (`.txt`) that was downloaded.

```
show run
```

   a. The output for both installed base switches and new switches should be identical to the contents of the configuration file for the following items:
   - `banner` (should match the expected version)
   - Switch port descriptions such as `description Cluster Node x`
   - The new ISL algorithm `port-channel load-balance Ethernet source-dest-port`

   b. The output for new switches should be identical to the contents of the configuration file for the following items:
   - Port channel
   - Policy map
   - System QoS
   - Interface
   - Boot

   c. The output for installed-base switches should have the flow control receive and send values `on` for the following items:
   - Interface port channels 1 and 2
   - Ethernet interface 1/41 through Ethernet interface 1/48

9. Copy the `running-config` to the `startup-config`.

```
copy running-config startup-config
```

Cisco Smart Call Home Setup

To configure Smart Call Home on a Cisco Nexus 5596 switch, complete the following steps:

1. Enter the mandatory system contact using the `snmp-server contact` command in global configuration mode. Then run the `callhome` command to enter callhome configuration mode.

```
NX-5596#config t
NX-5596(config)#snmp-server contact <<var_network_admin_email>>
NX-5596(config)#callhome
```

2. Configure the mandatory contact information (phone number, e-mail address, and street address).

```
NX-5596(config-callhome)#email-contact <<var_network_admin_email>>
NX-5596(config-callhome)#phone-contact <<var_network_admin_phone>>
NX-5596(config-callhome)#streetaddress <<var_network_admin_address>>
```

3. Configure the mandatory e-mail server information.

```
NX-5596(config-callhome)#transport email smtp-server <<var_mailhost>> port 25 use-vrf
<<var_vrf_name_mail>>
```

Note: The server address is an IPv4 address, IPv6 address, or the domain name of an SMTP server to which Call Home will send e-mail messages. Optional port number (default=25) and VRF may be configured.

4. Set the destination profile CiscoTAC-1 e-mail address to callhome@cisco.com.

```
NX-5596(config-callhome)#destination-profile CiscoTAC-1 email-addr callhome@cisco.com
```

5. Enable periodic inventory and set the interval.

```
NX-5596(config-callhome)#periodic-inventory notification
NX-5596(config-callhome)#periodic-inventory notification interval 30
```

6. Enable callhome, exit, and save the configuration.

```
NX-5596(config-callhome)#enable
NX-5596(config-callhome)#end
NX-5596#copy running-config startup-config
```

7. Send a callhome inventory message to start the registration process.

```
NX-5596#callhome test inventory
trying to send test callhome inventory message
successfully sent test callhome inventory message
```

8. Watch for an e-mail from Cisco regarding the registration of the switch. Follow the instructions in the e-mail to complete the registration for Smart Call Home.

SNMP Monitoring Setup

1. Configure SNMP by using the following example as a guideline.

```
NX-5596#config t
NX-5596(config)# snmp-server host <<var_snmp_server_ip>> traps { version 1 } <community>
[udp_port <number>]
NX-5596(config)# snmp-server enable traps link
```

Note: This example configures a host receiver for SNMPv1 traps and enables all link up/down traps.

Initializing Clustered Data ONTAP 8.2

Complete the following steps on new node 1 and new node 2.

1. Connect to the storage system console port. The terminal should display a `LOADER-A` prompt. However, if the storage system is in a reboot loop, press Ctrl-C to exit the autoboot loop when the terminal displays this message:

```
Starting AUTOBOOT press Ctrl-C to abort
```

2. From the `LOADER-A` prompt:

```
printenv
```

3. If the `last-OS-booted-ver` parameter is not set to 8.2, proceed to step 4 to load Data ONTAP 8.2 software. If Data ONTAP 8.2 is already loaded, proceed to step 16.
4. Allow the system to boot up.

```
boot_ontap
```

5. Press Ctrl-C when the `Press Ctrl-C for Boot Menu` message appears.

Note: If Data ONTAP 8.2 is not the version of software being booted, proceed with the following steps to install new software. If Data ONTAP 8.2 is the version being booted, then select option `8` and `yes` to reboot the node. Then proceed with step 15.

6. To install new software, first select option 7.

```
7
```

7. Answer yes to perform a software upgrade.

```
y
```

8. Select e0M for the network port to use for the download.

```
e0M
```

9. Select yes to reboot now.

```
y
```

10. Enter the IP address, netmask, and default gateway for e0M in their respective places.

During Node 1 Setup	During Node 2 Setup

`<<var_new_node01_mgmt_ip>>`
`<<var_new_node01_mgmt_mask>>`
`<<var_new_node01_mgmt_gateway>>`
	`<<var_new_node02_mgmt_ip>>`
	`<<var_new_node02_mgmt_mask>>`
	`<<var_new_node02_mgmt_gateway>>`

11. Enter the URL where the software can be found.

```
<<var_url_boot_software>>
```

Note: This web server must be reachable. Test using the `ping` utility.

12. Press Enter for the username, indicating no user name.

```
Enter
```

13. Enter yes to set the newly installed software as the default to be used for subsequent reboots.

```
y
```

14. Enter yes to reboot the node.

```
y
```

Note: When installing new software, the system might perform firmware upgrades to the BIOS and adapter cards, causing reboots and possible stops at the Loader prompt. If these actions occur, the system might deviate from this procedure.

15. Press Ctrl-C to exit autoboot when this message is displayed:

```
Starting AUTOBOOT press Ctrl-C to abort…
```

16. From the `LOADER-A` prompt, enter:

```
printenv
```

Note: If `bootarg.init.boot_clustered true` is not listed, the system is not set to boot in clustered Data ONTAP.

17. If the system is not set to boot in clustered Data ONTAP, at the LOADER prompt, enter the following command to make sure the system boots in clustered Data ONTAP:

```
setenv bootarg.init.boot_clustered true
setenv bootarg.bsdportname e0M
```

18. At the `LOADER-A` prompt, enter:

```
 autoboot
```

19. Enter special boot menu by pressing Ctrl-C when you see the following:

```
 Press Ctrl-C for Boot Menu:
```

20. Select option `4` for clean configuration and initialize all disks.

```
4
```

21. Answer `yes` to `Zero disks, reset config and install a new file system.`

```
yes
```

22. Enter `yes` to erase all the data on the disks.

```
yes
```

Note: The initialization and creation of the root volume can take 75 minutes or more to complete, depending on the number of disks attached. After initialization is complete, the storage system reboots. Continue to node 02 configuration while the disks for node 01 are zeroing.

Cluster Create and Cluster Join in Clustered Data ONTAP

The following section should be completed for both nodes. The first node in the cluster performs the `cluster create` operation. All other nodes perform a `cluster join` operation. The first node in the cluster is considered node 01. The second node in the cluster is considered node 02.

1. During the first node boot, the Cluster Setup wizard starts running on the console.

```
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}:
```

Note: If a login prompt appears instead of the Cluster Setup wizard, start the wizard by logging in using the factory default settings and then enter the `cluster setup` command.

2. Create a new cluster using the following commands.

During New Node 01 Setup	During New Node 02 Setup
`create`	`join`

3. Activate HA and set storage failover.

During New Node 01 Setup	During New Node 02 Setup
Do you intend for this node to be used as a single node cluster? {yes, no} [no]: Enter Non-HA mode, Reboot node to activate HA□□Do you want to reboot now to set storage failover (SFO) to HA mode? {yes, no}□[yes]: Enter	Non-HA mode, Reboot node to activate HA Do you want to reboot now to set storage failover (SFO) to HA mode? {yes, no} [yes]: Enter

4. Press Enter. The node will reboot. After the reboot, continue with setup.

   The system defaults are displayed.

During New Node 01 Setup	During New Node 02 Setup
System Defaults:□Private cluster network ports [e1a,e2a].□Cluster port MTU values will be set to 9000.□Cluster interface IP addresses will be automatically generated.□The cluster will be connected using network switches.□□Do you want to use these defaults? {yes, no} [yes]: Enter	This node's storage failover partner is already a member of a cluster. Storage failover partners must be members of the same cluster. The cluster setup wizard will default to the cluster join dialog. Existing cluster interface configuration found: Port MTU IP Netmask e1a 9000 169.254.251.110 255.255.0.0 e2a 9000 169.254.56.206 255.255.0.0 Do you want to use this configuration? {yes, no} [yes]: Enter

5. NetApp recommends accepting the system defaults. To accept the system defaults, press Enter.

Note: Initial cluster creation can require up to two minutes.

6. The steps to create or join a cluster are displayed. Enter appropriate responses to the prompts.

During New Node 01 Setup	During New Node 02 Setup
Enter the cluster name: <<var_clustername>> Enter the cluster base license key: <<var_cluster_base_license_key>> Creating cluster <<var_clustername>> Enter an additional license key []:	Enter the name of the cluster you would like to join [<<var_clustername>>]:Enter

7. Enter the required information as prompted.

During New Node 01 Setup	During New Node 02 Setup
Enter the cluster administrators (username "admin") password: <<var_password>> Retype the password: <<var_password>> Enter the cluster management interface port [e0a]: e0a Enter the cluster management interface IP address: <<var_clustermgmt_ip>> Enter the cluster management interface netmask: <<var_clustermgmt_mask>> Enter the cluster management interface default gateway: <<var_clustermgmt_gateway>>	N/A

8. Enter the DNS domain name.

During Node 01 Setup	During Node 02 Setup
`Enter the DNS domain names:<<var_dns_domain_name>>` `Enter the name server IP addresses:<<var_nameserver_ip>>`	N/A

Note: To use multiple name server IP addresses, separate each entry with a comma.

9. Set up the node.

During New Node 01 Setup	During New Node 02 Setup

`Where is the controller located []:<<var_node_location>>`
`Enter the node management interface port [e0M]: e0b`
`Enter the node management interface IP address: <<var_new_node01_mgmt_ip>>`
`Enter the node management interface netmask:<<var_new_node01_mgmt_mask>>`
`Enter the node management interface default gateway:<<var_new_node01_mgmt_gateway>>`
	`Where is the controller located []:<<var_node_location>>`
	`Enter the node management interface port [e0M]: e0b`
	`Enter the node management interface IP address: <<var_new_node02_mgmt_ip>>`
	`Enter the node management interface netmask:<<var_new_node02_mgmt_mask>>`
	`Enter the node management interface default gateway:<<var_new_node02_mgmt_gateway>>`

Note: The node management interface should be in a different subnet than the cluster management interface. The node management interfaces can reside on the out-of-band management network, and the cluster management interface can be on the in-band management network.

10. Press Enter to accept the AutoSupport message.
11. Reboot the node.

During New Node 01 Setup	During New Node 02 Setup
`system node reboot –node <<var_new_node01>>` `Warning: Are you sure you want to reboot the node? {y|n}: y`	`system node reboot –node <<var_new_node02>>` `Warning: Are you sure you want to reboot the node? {y|n}: y`

12. When the terminal displays `Press Ctrl-C for Boot Menu`, enter:

```
Ctrl – C
```

13. Select 5 to boot into maintenance mode.

```
5
```

14. When prompted `Continue with boot?`

```
y
```

15. To verify the HA status of your environment, run the following command:

```
ha-config show
```

Note: If either component is not in HA mode, use the `ha-config modify` command to put the components in HA mode.

16. Reboot the controller.

```
halt
```

17. At the `LOADER-A` prompt, enter:

```
autoboot
```

18. Log in to the cluster.
19. Data ONTAP assigns disks to storage controllers automatically if the `disk autoassign` setting is enabled. Verify the autoassign settings.

```
storage disk option show -fields autoassign command to verify the setting.
```

20. Reboot the node.

During New Node 01 Setup	During New Node 02 Setup
`system node reboot -node <<var_new_node01>>` `Warning: Are you sure you want to reboot the node? {y|n}: y`	`system node reboot -node <<var_new_node02>>` `Warning: Are you sure you want to reboot the node? {y|n}: y`

21. When the terminal displays `Press Ctrl-C for Boot Menu`, enter:

```
Ctrl - C
```

22. Select 5 to boot into maintenance mode.

```
5
```

23. When prompted `Continue with boot?`

```
y
```

24. To see how many disks are unowned, enter:

```
disk show -a
```

Note: No disks listed should have ownership.

25. Assign disk ownership.

```
disk assign -n <<var_#_of_disks>>
```

Note: This example allocates half the disks to each controller. However, workload design could dictate different percentages.

26. Reboot the controller.

```
halt
```

27. At the `LOADER-A` prompt, enter:

```
autoboot
```

28. Open an SSH connection to cluster IP or host name and log in to the admin user with the password provided earlier.
29. Zero all spare disks in the cluster.

```
disk zerospares
```

30. Set the auto-revert parameter on the cluster management interface.

```
network interface modify -vserver <<var_clustername>> -lif cluster_mgmt -auto-revert true
```

31. Create a management port failover group for the cluster.

```
network interface failover-groups create -failover-group mgmt -node <<var_new_node01>> -port e0a
network interface failover-groups create -failover-group mgmt -node <<var_new_node02>> -port e0a
```

32. Assign the management port failover group to the cluster management LIF.

```
network interface modify -vserver <<var_clustername>> -lif cluster_mgmt -failover-group mgmt
```

33. Create a management port failover group for the individual nodes in the cluster.

```
network interface failover-groups create -failover-group node-mgmt01 -node <<var_new_node01>> -
port e0b
network interface failover-groups create -failover-group node-mgmt01 -node <<var_new_node01>> -
port e0M
network interface failover-groups create -failover-group node-mgmt02 -node <<var_new_node02>> -
port e0b
```

```
network interface failover-groups create -failover-group node-mgmt02 -node <<var_new_node02>> -
port e0M
```

34. Assign the management port failover group to the cluster management LIF.

```
network interface modify -vserver <<var_new_node01>> -lif mgmt1 -auto-revert true -failover-group
node-mgmt01
network interface modify –vserver <<var_new_node02>> -lif mgmt1 –auto-revert true –failover-group
node-mgmt02
```

35. If you have a Flash Cache card installed, complete the following to enable it. If not, skip to step 37.

```
system node run -node <<var_new_node01>> options flexscale.enable on
system node run -node <<var_new_node01>> options flexscale.lopri_blocks off
system node run -node <<var_new_node01>> options flexscale.normal_data_blocks on
system node run -node <<var_new_node02>> options flexscale.enable on
system node run -node <<var_new_node02>> options flexscale.lopri_blocks off
system node run -node <<var_new_node02>> options flexscale.normal_data_blocks on
```

Note: Data ONTAP 8.2 and later does not require a separate license for Flash Cache.

Note: For directions on how to configure Flash Cache in metadata mode or low-priority data caching mode, refer to TR-3832: Flash Cache Best Practices Guide. Before customizing the settings, determine whether the custom settings are required or if the default settings are sufficient.

36. Execute the following command to create new aggregates:

```
aggr create -aggregate aggr01 -nodes <<var_new_node01>> -diskcount <<var_num_disks>>
aggr create -aggregate aggr02 -nodes <<var_new_node02>> -diskcount <<var_num_disks>>
```

Note: A 64-bit aggregate containing the root volume is created during the Data ONTAP setup process. To create additional 64-bit aggregates, determine the aggregate name, the node on which to create it, and the number of disks it will contain.

Note: Retain at least one disk (select the largest disk) in the configuration as a spare. A best practice is to have at least one spare for each disk type and size.

Note: The aggregate cannot be created until disk zeroing completes. Use the `aggr show` command to display aggregate creation status. Do not proceed until both aggr01 and aggr02 are online.

37. Disable Snapshot copies for the two data aggregates just created.

```
node run <<var_new_node01>> aggr options aggr01 nosnap on
node run <<var_new_node02>> aggr options aggr02 nosnap on
```

38. Delete any existing Snapshot copies for the two data aggregates.

```
node run <<var_new_node01>> snap delete –A –a –f aggr01
node run <<var_new_node02>> snap delete –A –a –f aggr02
```

Service Processor

Gather information about the network and the AutoSupport settings before configuring the service processor (SP).

Configure the SP using DHCP or static addressing. If the SP uses a static IP address, verify that the following SP prerequisites have been met:

- An available static IP address
- The network netmask
- The network gateway IP
- AutoSupport information

A best practice is to configure the AutoSupport recipients and mail host before configuring the SP. Data ONTAP automatically sends AutoSupport configuration to the SP, allowing the SP to send alerts and notifications through an AutoSupport message to the system administrative recipients specified in

AutoSupport. When configuring the SP, enter the name or the IP address of the AutoSupport mail host, when prompted.

Note: The service processor must be set up on each node.

1. Check the configuration and version of the service processor.

```
system node service-processor show
```

2. Using a web browser, connect to http://support.netapp.com/NOW/cgi-bin/fw. The complete instructions and release notes can be found here.
3. Navigate to the “Service Process Image for installation from the Data ONTAP prompt” page for your storage platform. This document will use the instructions for the FAS3270 model.
4. Check the latest firmware version that is available for your storage platform. If your storage system is not running the latest version, proceed to the download page for the latest release of the SP firmware for your storage platform.
5. Update the SPs on both nodes. Download the .zip file to a web server that is reachable from the cluster management interface. Issue the following command to download the firmware file to the nodes:

```
system node image get -node <<var_node03>> -package <<var_fw_pkg_location>> -replace-package true
system node image get -node <<var_node04>> -package <<var_fw_pkg_location>> -replace-package true
```

6. Verify that the nodes have been updated to the latest service processor.

```
system node service-processor image update-progress show
```

Example output:

```
                In                              Percent
Node            Progress Start Time             Done    End Time
--------------- -------- -------------------- ------- --------------------
Node1           no       -                     0       -
Node2           no       -                     0       -
Node3           yes      10/13/2013 20:00:34   99      -
Node4           no       -                     0       -
2 entries were displayed.
```

7. Verify that the service processors are online and running the correct firmware version.

```
system node service-processor show
```

Example output:

```
                                IP           Firmware
Node          Type Status       Configured   Version   IP Address
------------- ---- ------------ ------------ --------- ------------------------
node1         SP   online       true         1.4P1     10.251.133.139
node2         SP   online       true         1.4P1     10.251.133.140
```

Configure the Service Processor

1. From the cluster shell, enter the following command:

During New Node 01 Setup	During New Node 02 Setup
`system node run <<var_new_node01>> sp setup`	`system node run <<var_new_node01>> sp setup`

2. Set up the service processor.

During New Node 01 Setup	During New Node 02 Setup
Would you like to configure the SP? Y Would you like to enable DHCP on the SP LAN interface? no Please enter the IP address of the SP[]: <<var_new_node01_sp_ip>> Please enter the netmask of the SP[]: <<var_new_node01_sp_mask>> Please enter the IP address for the SP gateway[]: <<var_new_node01_sp_gateway>>	Would you like to configure the SP? Y Would you like to enable DHCP on the SP LAN interface? no Please enter the IP address of the SP[]: <<var_new_node02_sp_ip>> Please enter the netmask of the SP[]: <<var_new_node02_sp_mask>> Please enter the IP address for the SP gateway[]: <<var_new_node02_sp_gateway>>

3. Verify the status of storage failover.

```
storage failover show
```

4. Both the nodes <<var_new_node01>> and <<var_new_node02>> must be capable of performing a takeover. The "Takeover Possible" column should display "true" for both nodes.
5. If takeover is not possible, enable failover on one of the two nodes.

```
storage failover modify -node <<var_node01>> -enabled true
```

Note: Storage takeover on the partner node is automatically enabled.

6. Verify HA status for two-node clusters:

```
cluster ha show
```

Note: This step is not applicable for clusters with more than two nodes.

7. Enable HA mode for two-node clusters only.

```
cluster ha modify -configured true
Do you want to continue? {y|n}: y
```

Note: Do not run this command for clusters with more than two nodes because it will cause problems with failover.

8. Verify that hardware assist is correctly configured and if needed modify the partner IP address.

```
storage failover hwassist show
storage failover modify –hwassist-partner-ip <<var_new_node02_mgmt_ip>> -node <<var_new_node01>>
storage failover modify –hwassist-partner-ip <<var_new_node01_mgmt_ip>> -node <<var_new_node02>>
```

9. Run the following commands on the command line to create interface groups (ifgrps).

```
ifgrp create -node <<var_new_node01>> -ifgrp a0a -distr-func port -mode multimode_lacp
network port ifgrp add-port -node <<var_new_node01>> -ifgrp a0a -port e3a
network port ifgrp add-port -node <<var_new_node01>> -ifgrp a0a -port e4a
ifgrp create -node <<var_new_node02>> -ifgrp a0a -distr-func port -mode multimode_lacp
network port ifgrp add-port -node <<var_new_node02>> -ifgrp a0a -port e3a
network port ifgrp add-port -node <<var_new_node02>> -ifgrp a0a -port e4a
```

Note: This type of interface group requires two or more Ethernet interfaces and a switch that supports LACP. Therefore, make sure that the switch is configured properly.

Note: All interfaces must be in the down status before being added to an interface group.

Note: The interface group name must follow the standard naming convention of "a<number><letter>," where <number> is an integer in the range 0–999 without leading zeros and <letter> is a lowercase letter.

10. Create NFS VLANs.

```
network port vlan create –node <<var_new_node01>> -vlan-name a0a-<<var_nfs_vlan_id>>
network port vlan create –node <<var_new_node02>> -vlan-name a0a-<<var_nfs_vlan_id>>
```

11. Disable flow control.

```
net port mod -node <node> -port <port> -flowcontrol-admin none
```

12. Repeat step 2 for each cluster node port.

Note: Flow control is configured on each physical port, even if it is a member of an interface group.

Note: Some models of Ethernet adapter such as the Unified Target Adapter (UTA) do not allow the flow control setting to be changed. In this instance, disabling flow control at the switch is recommended.

Note: Changing flow control settings disrupts the network connection for several seconds.

13. To configure a clustered Data ONTAP network port to use jumbo frames (which usually have an MTU of 9,000 bytes), run the following command from the cluster shell:

```
network port modify -node <<var_new_node01>> -port a0a -mtu 9000

WARNING: Changing the network port settings will cause a serveral second interruption in carrier.
Do you want to continue? {y|n}: y

network port modify –node <<var_new_node01>> -port a0a-<<var_nfs_vlan_id>> -mtu 9000

WARNING: Changing the network port settings will cause a serveral second interruption in carrier.
Do you want to continue? {y|n}: y

network port modify –node <<var_ new_node02>> -port a0a -mtu 9000

WARNING: Changing the network port settings will cause a serveral second interruption in carrier.
Do you want to continue? {y|n}: y

network port modify –node <<var_ new_node02>> -port a0a-<<var_nfs_vlan_id>> -mtu 9000

WARNING: Changing the network port settings will cause a serveral second interruption in carrier.
Do you want to continue? {y|n}: y
```

14. Set the time zone for the cluster.

```
timezone <<var_timezone>>
```

Note: For example, in the eastern United States, the time zone is `America/New_York`.

15. Set the date for the cluster.

```
date <ccyymmddhhmm.ss>
```

Note: The format for the date is `<[Century][Year][Month][Day][Hour][Minute].[Second]>`; for example, `201309081735.17`

16. Configure the Network Time Protocol (NTP) for each node in the cluster.

```
system services ntp server create -node <<var_ new_node01>> -server <<var_global_ntp_server_ip>>
system services ntp server create -node <<var_ new_node02>> -server <<var_global_ntp_server_ip>>
```

17. Configure SNMP basic information, such as the location and contact. When polled, this information is visible as the `sysLocation` and `sysContact` variables in SNMP.

```
snmp contact <<var_snmp_server_contact>>
snmp location "<<var_snmp_server_location>>"
snmp init 1
options snmp.enable on
```

18. Set the shared secret plain-text password, which is called a community.

```
snmp community delete all
snmp community add ro <<var_snmp_community>>
```

Note: Use the `delete all` command with caution. If community strings are used for other monitoring products, the `delete all` command will remove them.

19. Create a user called `snmpv3user`.

```
security login create -username snmpv3user -authmethod usm -application snmp
```

Note: SNMPv3 requires that a user be defined and configured for authentication.

20. Enter the authoritative entity's Engine ID and select `md5` as the authentication protocol.

Note: Use `security snmpusers` to view the Engine ID.

21. Enter an eight-character minimum-length password for the authentication protocol, when prompted.
22. Select `des` as the privacy protocol.
23. Enter an eight-character minimum-length password for the privacy protocol, when prompted.
24. Configure AutoSupport.

```
system node autosupport modify -node * -state enable –mail-hosts <<var_mailhost>> -transport
https -support enable -noteto <<var_storage_admin_email>>
```

Note: AutoSupport sends support summary information to NetApp through HTTPS.

25. Enable CDP on Data ONTAP.

```
node run -node <<var_new_node01>> options cdpd.enable on
node run -node <<var_new_node02>> options cdpd.enable on
```

Note: To be effective, CDP must also be enabled on directly connected networking equipment such as switches and routers.

26. To create a Vserver, run the Vserver setup wizard.

```
vserver setup

Welcome to the Vserver Setup Wizard, which will lead you through
the steps to create a virtual storage server that serves data to clients.

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 your answers to previous questions, and
"exit" if you want to quit the Vserver Setup Wizard. Any changes
you made before typing "exit" will be applied.

You can restart the Vserver Setup Wizard by typing "vserver setup". To accept a default
or omit a question, do not enter a value.

Vserver Setup wizard creates and configures only data Vservers.
If you want to create a Vserver with Infinite Volume use the vserver create command.

Step 1. Create a Vserver.
You can type "back", "exit", or "help" at any question.
```

27. Enter the Vserver's name.

```
<<var_vserver>>
```

28. Select the Vserver data protocols to configure.

```
Choose the Vserver data protocols to be configured {nfs, cifs, fcp, iscsi, ndmp}: nfs,fcp
```

29. Select the Vserver client services to configure.

```
Choose the Vserver client services to configure {ldap, nis, dns}:Enter
```

30. Enter the Vserver's root volume aggregate.

```
Enter the Vserver's root volume aggregate {aggr01, aggr02} [aggr01]:aggr01
```

31. Enter the Vserver language setting. English is the default (C.UTF8).

```
Enter the Vserver language setting, or "help" to see all languages [C]:Enter
```

32. Enter the Vserver's security style.

```
Enter the Vserver root volume's security style {mixed, ntfs, unix} [unix]: Enter
```

33. Answer no to `Do you want to create a data volume`?

```
Do you want to create a data volume?  {yes, no} [Yes]: no
```

34. Answer no to `Do you want to create a logical interface?`

```
Do you want to create a logical interface?  {yes, no} [Yes]: no
```

35. Answer no to `Do you want to Configure FCP? {yes, no} [yes]: no.`

```
Do you want to Configure FCP? {yes, no} [yes]: no
```

36. Add the two data aggregates to the `<<var_vserver>>` aggregate list for NetApp Virtual Console.

```
vserver modify –vserver <<var_vserver>> –aggr-list aggr01, aggr02
```

37. Create a volume to be the load sharing mirror of the infrastructure Vserver root volume on each node.

```
volume create –vserver <<var_vserver>> –volume root_vol_m01 –aggregate aggr01 –size 20MB –type DP
volume create –vserver <<var_vserver>> –volume root_vol_m02 –aggregate aggr02 –size 20MB –type DP
```

38. Create the mirroring relationships.

```
snapmirror create –source-path //<<var_vserver>>/rootvol –destination-path
//<<var_vserver>>/root_vol_m01 –type LS
snapmirror create –source-path //<<var_vserver>>/rootvol –destination-path
//<<var_vserver>>/root_vol_m02 –type LS
```

39. Initialize the mirroring relationship.

```
snapmirror initialize-ls-set –source-path //<<var_vserver>>/rootvol
```

40. Set an hourly (at 5 minutes past the hour) update schedule on each mirroring relationship.

```
snapmirror modify –source-path //<<var_vserver>>/rootvol –destination-path * -schedule hourly
```

41. Create the FC service on each Vserver. This command also starts the FC service and sets the FC alias to the name of the Vserver.

```
fcp create -vserver <<var_vserver>>
```

42. Increase the privilege level to access the certificate commands.

```
set -privilege diag
Do you want to continue? {y|n}: y
```

43. Generally, a self-signed certificate is already in place. Check it by using the following command:

```
security certificate show
```

44. Run the following commands as one-time commands to generate and install self-signed certificates:

Note: Use the `security certificate delete` command to delete expired certificates.

```
security certificate create -vserver <<var_vserver>> -common-name
<<var_security_cert_vserver_common_name>> -size 2048 -country <<var_country_code>> -state
<<var_state>> -locality <<var_city>> -organization <<var_org>> -unit <<var_unit>> -email-addr
<<var_storage_admin_email>>
security certificate create -vserver <<var_clustername>> -common-name
<<var_security_cert_cluster_common_name>> -size 2048 -country <<var_country_code>> -state
<<var_state>> -locality <<var_city>> -organization <<var_org>> -unit <<var_unit>> -email-addr
<<var_storage_admin_email>>
security certificate create -vserver <<var_node01>> -common-name
<<var_security_cert_node01_common_name>> -size 2048 -country <<var_country_code>> -state
<<var_state>> -locality <<var_city>> -organization <<var_org>> -unit <<var_unit>> -email-addr
<<var_storage_admin_email>>
security certificate create -vserver <<var_node02>> -common-name
<<var_security_cert_node02_common_name>> -size 2048 -country <<var_country_code>> -state
```

```
<<var_state>> -locality <<var_city>> -organization <<var_org>> -unit <<var_unit>> -email-addr
<<var_storage_admin_email>>
```

45. Configure and enable SSL and HTTPS access and disable telnet access.

```
system services web modify -external true -sslv3-enabled true
Do you want to continue {y|n}: y
system services firewall policy delete -policy mgmt -service http -action allow
system services firewall policy create -policy mgmt -service http -action deny -ip-list 0.0.0.0/0
system services firewall policy delete -policy mgmt -service telnet -action allow
system services firewall policy create -policy mgmt -service telnet -action deny -ip-list
0.0.0.0/0
```

46. Use the `security certificate` command to obtain the values for the parameters that would be needed in the following step.

```
security certificate show
```

Example output:

```
Vserver    Serial Number   Common Name                              Type
---------- --------------- ---------------------------------------- ------------
FAS3250-Cluster
           52173806        FAS3250-Cluster.cert                     server
    Certificate Authority: FAS3250-Cluster.cert
          Expiration Date: Sat Aug 23 15:53:03 2014
```

47. Continue configuration.

```
security ssl modify -vserver <<var_vserver>> -common-name
<<var_security_cert_vserver_common_name>> -server-enabled true -client-enabled false -ca
<<var_security_certificate_vservser_authority>> -serial
<<var_security_certificate_vserver_serial_no>>

security ssl modify -vserver <<var_clustername>> -common-name
<<var_security_cert_cluster_common_name>> -server-enabled true -client-enabled false -ca
<<var_security_certificate_cluster_authority>> -serial
<<var_security_certificate_cluster_serial_no>>

security ssl modify -vserver <<var_ new_node01>> -common-name
<<var_security_cert_new_node01_common_name>> -server-enabled true -client-enabled false -ca
<<var_security_certificate_ new_node01_authority>> -serial <<var_security_certificate_
new_node01_serial_no>>

security ssl modify -vserver <<var_ new_node02>>-common-name <<var_security_cert_
new_node02_common_name>> -server-enabled true -client-enabled false -ca
<<var_security_certificate_ new_node02_authority>> -serial <<var_security_certificate_
new_node02_serial_no>>

set –privilege admin
```

Note: It is normal for some of these commands to return an error message stating that the entry does not exist.

48. Secure the default rule for the default export policy and create the FlexPod export policy.

```
vserver export-policy rule modify -vserver <<var_vserver>> -policyname default -ruleindex 1 -
rorule never -rwrule never -superuser none
vserver export-policy create –vserver <<var_vserver>> FlexPod
```

49. Add the infrastructure Vserver administrator and Vserver administration logical interface in the out-of-band management network.

```
network interface create –vserver <<var_vserver>> –lif vsmgmt –role data –data-protocol none –
home-node <<var_new_node02>> -home-port e0a –address <<var_vserver_mgmt_ip>> -netmask
<<var_vserver_mgmt_mask>> -status-admin up –failover-policy nextavail –firewall-policy mgmt –
auto-revert true –failover-group mgmt

network routing-groups route create –vserver <<var_vserver>> –routing-group
d<<var_clustermgmt_ip>> -destination 0.0.0.0/0 –gateway <<var_clustermgmt_gateway>>
```

```
security login password –username vsadmin –vserver <<var_vserver>>
Enter a new password:  <<var_vsadmin_password>>
Enter it again:  <<var_vsadmin_password>>

security login unlock –username vsadmin –vserver <<var_vserver>>
```

Update Cisco Nexus Data Switch Configurations

1. Log in to the Cisco Nexus switch with administrative access.
2. Configure port descriptions.

On Switch 1	On Switch 2
conf t int e1/13 desc <<var_new_node01>>:e3a int e1/14 desc <<var_ new_node02>>:e3a	conf t int e1/13 desc <<var_ new_node01>>:e4a int e1/14 desc <<var_ new_node02>>:e4a

3. The Cisco Nexus data switches should be configured to match the node interface groups and VLAN trunking configuration. Configure new port channels and virtual port channels corresponding to the new nodes.

On Switch 1	On Switch 2
conf t int po113 desc <<var_new_node01>> int po114 desc <<var_new_node02>> int e1/13 channel-group 113 mode active no shutdown int e 1/14 channel-group 114 mode active no shutdown int po 113 switchport mode trunk switchport trunk native vlan <<var_native_vlan>> switchport trunk allowed vlan <<var_nfs_vlan_id>> spanning-tree port type edge trunk no shutdown vpc 113 int po 114 switchport mode trunk switchport trunk native vlan <<var_native_vlan>> switchport trunk allowed vlan <<var_nfs_vlan_id>> spanning-tree port type edge trunk no shutdown vpc 114 exit copy run start	conf t int po113 desc <<var_new_node01>> int po114 desc <<var_new_node02>> int e1/13 channel-group 113 mode active no shutdown int e 1/14 channel-group 114 mode active no shutdown int po 113 switchport mode trunk switchport trunk native vlan <<var_native_vlan>> switchport trunk allowed vlan <<var_nfs_vlan_id>> spanning-tree port type edge trunk no shutdown vpc 113 int po 114 switchport mode trunk switchport trunk native vlan <<var_native_vlan>> switchport trunk allowed vlan <<var_nfs_vlan_id>> spanning-tree port type edge trunk no shutdown vpc 114 exit copy run start

Note: Additional configuration might be required to utilize specific protocol functionality on the new cluster.

3.5 Add Shelf/Shelves to Existing Storage Controller

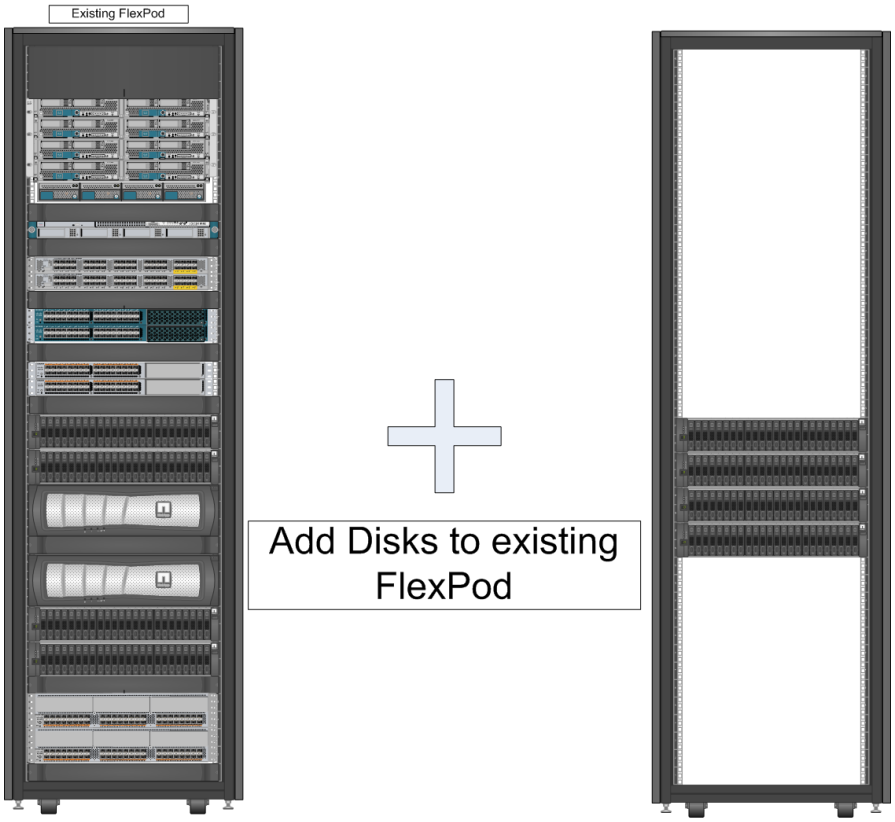
Table 11) List of variables.

Configuration Variable Name	Description
<<var_num_disks>>	Number of disks that are to be assigned to a node. This document recommends that half of the available disks be assigned to each node connected to the shelf.
<<var_vserver_name>>	Name of a Vserver that contains a volume that should be moved to another aggregate.
<<var_volume_name>>	Name of a volume that should be moved to another aggregate.
<<var_dest_aggregate_name>>	Name of a destination aggregate to which a target volume should be moved. This aggregate must be accessible by the volume owning the Vserver.
<<var_aggregate_name>>	Name of the aggregate to which the additional disks will be added.
<<var_disk1>>, <<var_disk2>>	Name of the disks to be added to the aggregate.

Scenario

Given an existing clustered HA pair running Data ONTAP 8.1.2–8.2, a new shelf can be nondisruptively added. After the shelves have been added, the disks can be added to existing aggregates, or new aggregates can be added, both nondisruptively. If new aggregates are added, volumes/LUNs can be nondisruptively moved to the new aggregates.

Figure 13) Add shelf/shelves to existing storage controller.



Before

Figure 14) MPHA SAS cabling.

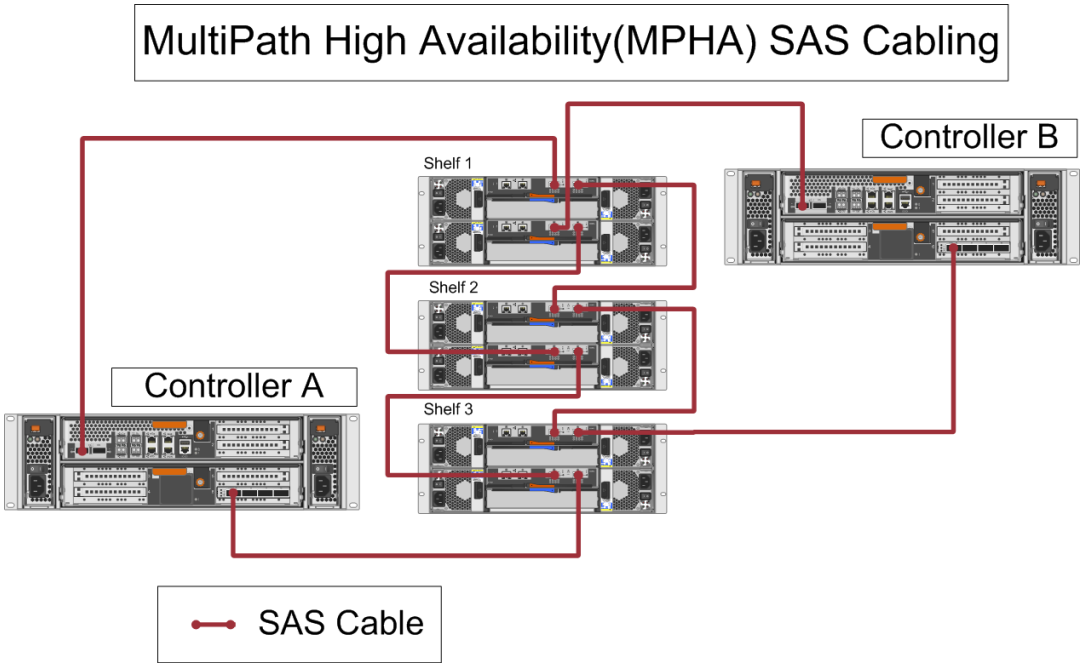
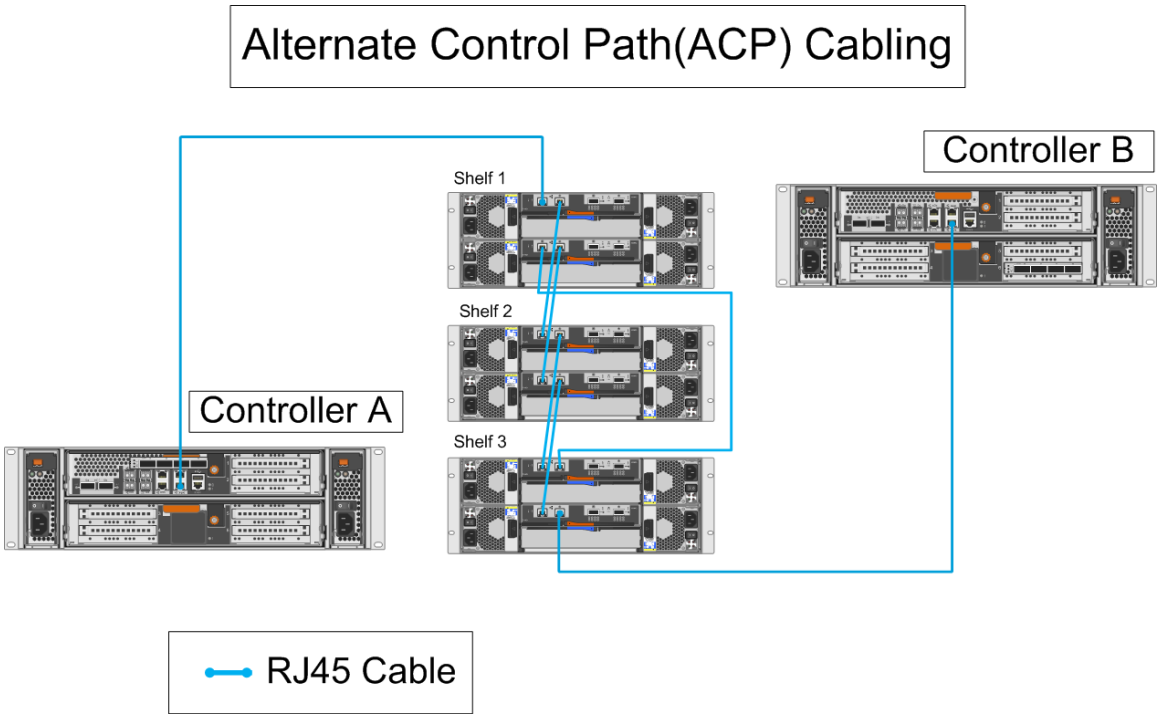


Figure 15) ACP cabling.



During

Figure 16) MPHA SAS cabling.

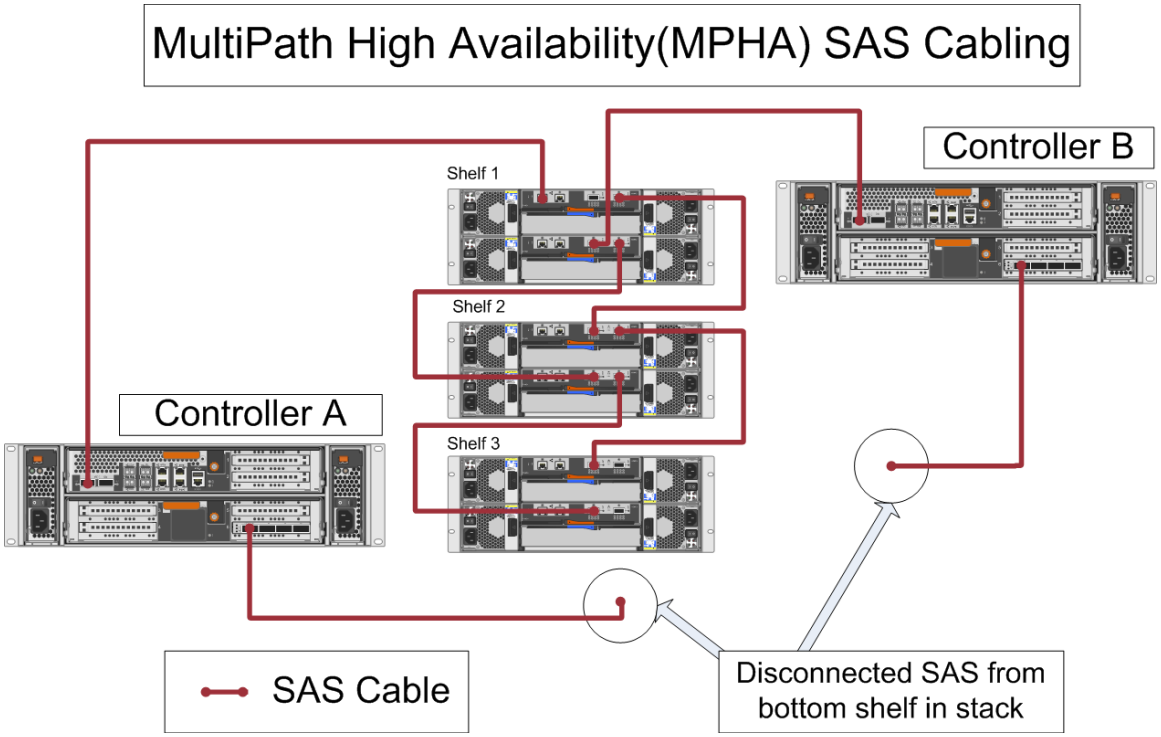
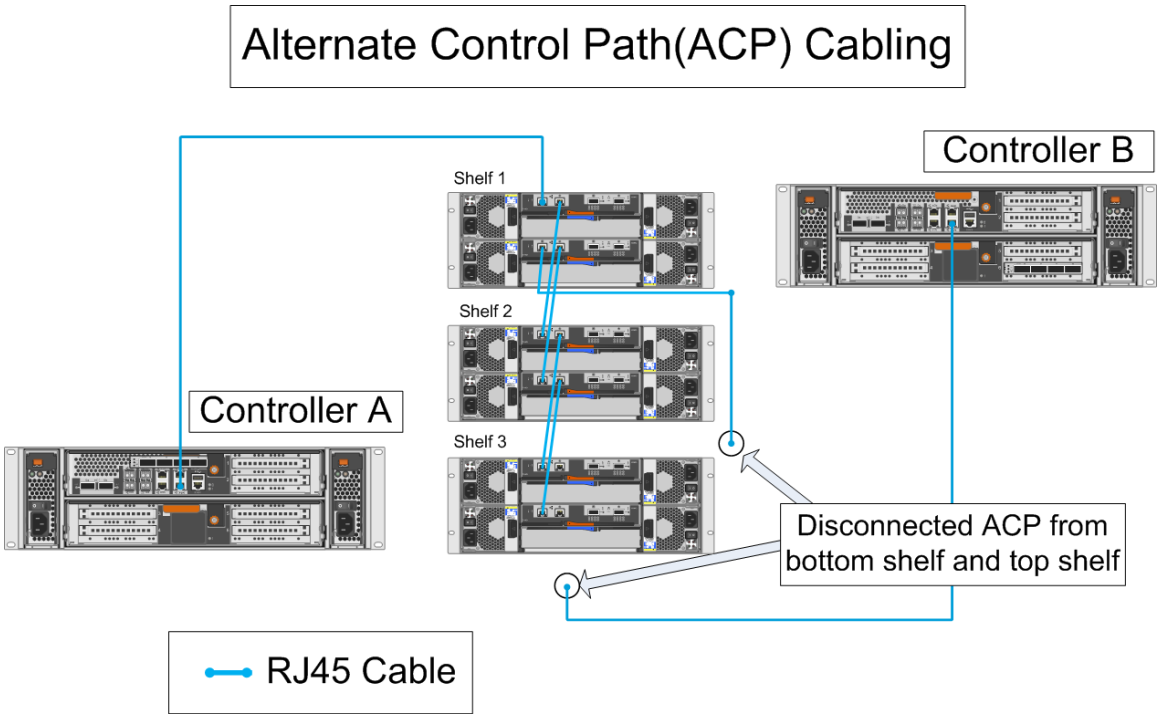


Figure 17) ACP cabling.



After

Figure 18) MPHA SAS cabling.

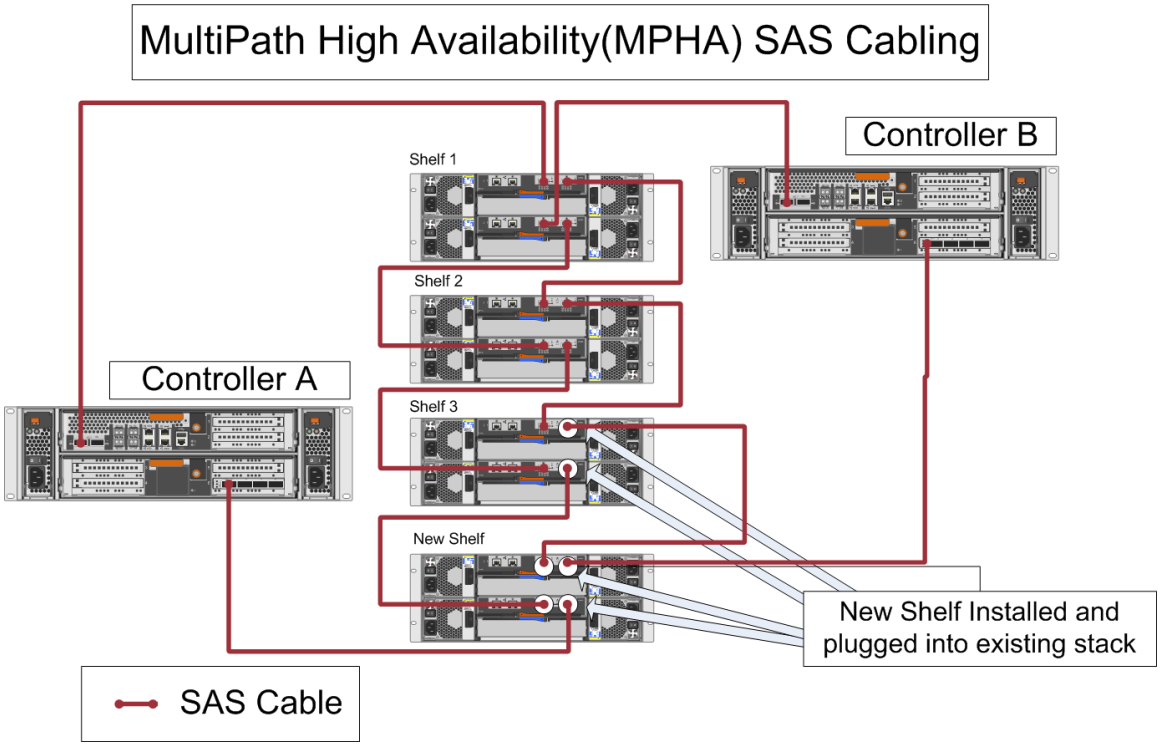
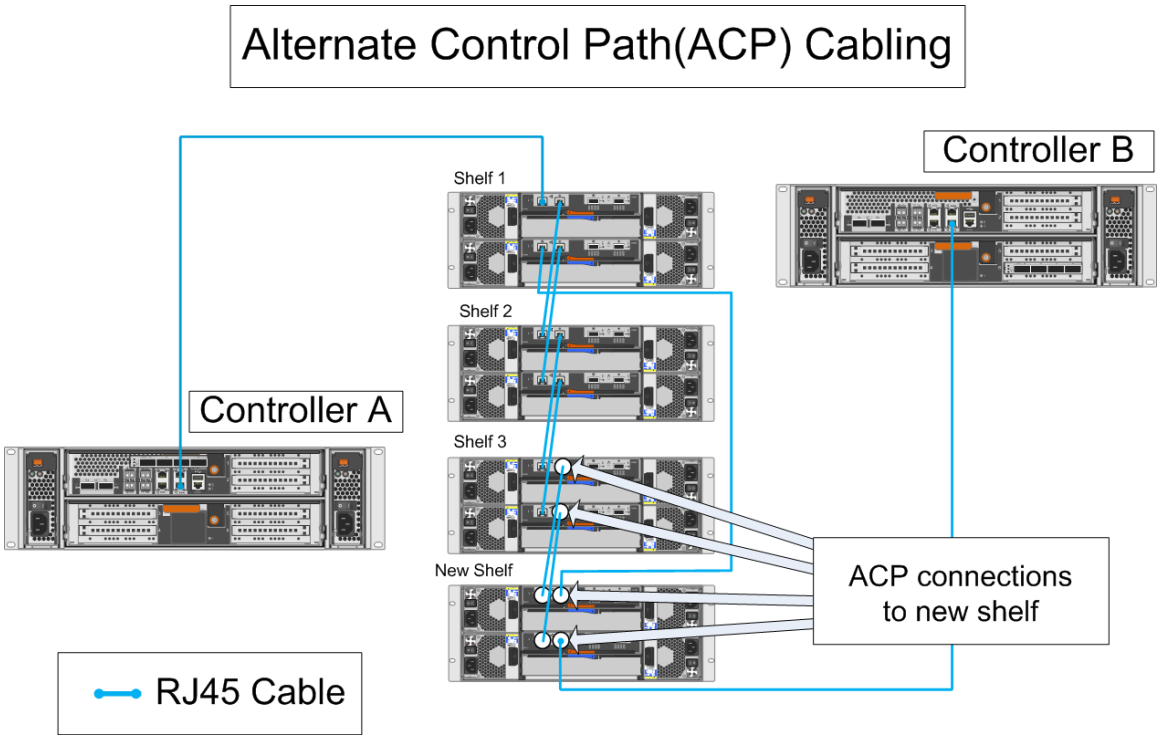


Figure 19) ACP cabling.



Procedure

This section will demonstrate two methods that can be used to install new disk shelves in an existing HA pair of controllers running clustered Data ONTAP 8.1.2 or higher within the context of a FlexPod unit. The first method assumes that two spare SAS ports exist in each controller in the HA pair. The second method assumes that the new shelf or shelves will connect to an existing stack of shelves.

Prerequisites	Description
Verify that the environment has enough available rack space and power and cooling accommodations to support the new disk shelves.	See NetApp Hardware Universe for the space, power, and cooling requirements for the appropriate disk shelf.
If adding shelves to an existing stack of shelves, make sure that desired quantity of shelves does not exceed the maximum number of shelves in a stack for the given Data ONTAP version and FAS model.	The maximum number of shelves allowed in a stack depends on the FAS model and Data ONTAP version. See NetApp Hardware Universe for the appropriate FAS model and Data ONTAP version.
Make sure that the desired quantity of shelves does not exceed the maximum number of shelves for the given Data ONTAP version and FAS model.	The maximum number of shelves allowed depends on the FAS model and Data ONTAP version. See NetApp Hardware Universe for the appropriate FAS model and Data ONTAP version.

Prerequisites	Description
Make sure that the desired disk shelves are compatible with the given version of Data ONTAP and model of shelf. **Note:** SAS shelves can only be connected to SAS stacks. FC-connected shelves can only be connected to FC-connected stacks.	See NetApp Hardware Universe for the appropriate FAS model and Data ONTAP version to determine the compatible shelves for the appropriate FAS model and Data ONTAP version.
If adding shelves and creating a new stack of shelves, make sure that the FAS controllers have enough SAS or storage FC ports available to accommodate the new stack.	Each controller in the HA pair should have two available storage ports of the appropriate type.
Make sure that all existing shelves and cables are installed and working properly.	Config Advisor can be used to confirm that all shelves and cables have been installed and are working properly. See the Config Advisor site for details on how to download and use this tool.

Adding Shelves to Existing Stack (Hot Add)

Note: It is assumed that SAS and ACP have been cabled by following the Universal SAS and ACP Cabling Guide.

To hot add disk shelves, complete the following steps:

1. Rack the shelves.
2. Connect shelves to existing stack using provided cables and in method prescribed in the appropriate shelf cabling document (https://library.netapp.com/ecm/ecm_get_file/ECMM1280392).
3. Power on the shelf and change shelf ID to a valid ID that is unique to the storage system.
4. Power cycle the shelf/shelves.
5. Unplug the bottom SAS cables from the existing stack that go to the storage controllers.

Note: Be sure to unplug one cable from IOXM-A loop and one from IOXM-B loop to avoid an outage.

Note: Connectivity to the storage system is maintained from the SAS cables at the top of the existing stack (MPHA).

6. Daisy chain the SAS ports of the new shelf into the existing stack that was just disconnected from the storage controllers.
7. Plug the bottom of the stack back into the storage controllers where it was disconnected previously.

Note: If all of the disks should not be assigned to the first node that detects them, disable the `disk.autoassign` option.

8. Connect an Ethernet cable from dedicated port on node 01 to the new disk shelf's IOM A ACP square port (ACP).
9. Connect an Ethernet cable from dedicated port on node 02 to the new disk shelf's IOM B ACP circle port (ACP).
10. Connect an Ethernet cable from IOM A ACP circle port to IOM B ACP square port (ACP).

Note: If it is only a single controller, complete steps 8 and 10 only (skip step 9).

11. Verify all cables are securely attached and LED indicators illuminate next to the disk shelf's SAS ports.
12. Log in to nodeshell of the node connected to the new shelf of disks.

```
node run -node <<var_nodename>>
```

13. Verify disks are seen and owned by the proper node.

```
disk show -a
```

14. Verify disk auto assign setting.

```
options disk.auto_assign
```

Note: If the `options disk.auto_assign` is set to `off`, manually assign your disks using the `disk assign` command. If `options disk_autoassign` is `on`, the first system that sees the disk will take ownership of it.

15. Verify that the disk shelf firmware is the most current version.

```
sysconfig -v
……
                Shelf   3: IOM3  Firmware rev. IOM3 A: 0160 IOM3 B: 0160
                Shelf   4: IOM3  Firmware rev. IOM3 A: 0160 IOM3 B: 0160
                Shelf  13: IOM3  Firmware rev. IOM3 A: 0160 IOM3 B: 0160
……
```

Note: In the above example, 0160 is the disk shelf firmware version for shelf 3.

Note: Go to the disk shelf firmware information on the NetApp Support site to determine the most current disk shelf firmware version.

Note: If Downrev states “yes,” update your firmware with the `disk_fw_update` command.

16. Exit nodeshell by pressing Ctrl-D.
17. New aggregates may not be created with the disks, or the disks can be added to existing aggregates.

Adding New Stack to Existing HA Pair

1. Rack the shelves.
2. Cable shelves according to the Universal SAS and ACP Cabling Guide using appropriate SAS cables.
3. Cable ACP to the existing ACP loop in the other shelf stack according to the Universal SAS and ACP Cabling Guide.
4. Power on the shelf and change shelf ID to a valid ID that is unique to the storage system.
5. Power cycle the shelf/shelves.
6. Connect the shelves to available SAS ports on controller as recommended in to the Universal SAS and ACP Cabling Guide.
7. Verify shelf is seen by the cluster pair and is unowned.

```
Run local
Disk show -a
```

Example output:

```
……
6a.6000300137.10 icef4-stcl1-04(1575111434)    Pool0  EC47PC5024HH          icef4-stcl1-
04(1575111434)
0a.6000299754.15 icef4-stcl1-04(1575111434)    Pool0  EC47PC101PSD          icef4-stcl1-
04(1575111434)
6a.6000300137.14 icef4-stcl1-04(1575111434)    Pool0  EC47PC5022T2          icef4-stcl1-
04(1575111434)
6a.6000300137.13 icef4-stcl1-04(1575111434)    Pool0  EC47PC5022K4          icef4-stcl1-
04(1575111434)
0a.6000299754.2 icef4-stcl1-04(1575111434)    Pool0  EC47PC1014UJ          icef4-stcl1-
04(1575111434)
0a.6000299754.7 icef4-stcl1-04(1575111434)    Pool0  EC47PC1014UH          icef4-stcl1-
04(1575111434)
6a.6000300137.5 icef4-stcl1-04(1575111434)    Pool0  EC47PC50236C          icef4-stcl1-
04(1575111434)
```

```
NOTE: Currently 24 disks are unowned. Use 'disk show -n' for additional information.
```

8. Assign disks to the appropriate controller.

```
disk assign -n <<var_num_disks>>
```

Note: `Disk assign all` can be used to assign all unowned disks seen by the storage system.

Note: After the ownership has been assigned, boot into maintenance mode to remove or change it.

Rebalancing Load Across Aggregates

1. Create new aggregates based on FlexPod best practices.
2. To relocate volumes/LUNs to new aggregates, run the following command. See Moving a volume for more information.

```
vol move start -vserver <<var_vserver_name>> -volume <<var_volume_name>> -destination-aggregate
<<var_dest_aggregate_name>>
```

Note: Remember to move corresponding LIFs or to create new SAN LIFs when moving a volume to an aggregate owned by a node different from the one it is currently on in order to avoid an indirect data path.

Growing Aggregates

1. Add disks to an aggregate using one of the following commands:

```
storage aggregate add-disks -aggregate <<var_aggregate_name>> -diskcount <<var_num_disks>>
storage aggregate add-disks -aggregate <<var_aggregate_name>> -disklist
<<var_disk1>>,<<var_disk2>>,…
```

Note: This should only be done in cases where the disk types are the same. New aggregates should be created and volumes should be moved if using different disks.

Note: If creating a hybrid aggregate using Flash Pool, see Increasing the size of an aggregate for more information.

Note: See Reallocate Best Practices guide for additional recommendations and scheduling details.

3.6 Permanently Removing a Controller from a Clustered Data ONTAP Cluster

Table 12) List of variables.

Configuration Variable Name	Description
<<var_retiring_node01>>	First node of an HA pair that is being retired and removed from a cluster
<<var_retiring_node02>>	Second node of an HA pair that is being retired and removed from a cluster
<<var_controller_name>>	Name of the controller being retired
<<var_other_node>>	Name of a controller in the cluster that is not being retired
<<var_retiring_node01_vpc>>	The switch virtual port channel number corresponding to the first node being retired
<<var_retiring_node02_vpc>>	The switch virtual port channel number corresponding to the second node being retired
<<var_retiring_node01_vfc>>	The switch virtual Fibre Channel number corresponding to the first node being retired
<<var_retiring_node02_vfc>>	The switch virtual Fibre Channel number corresponding to the second node being retired
<<var_retiring_node01_port-channel_number>>	The switch port channel number corresponding to the first node being retired
<<var_retiring_node02_port-channel_number>>	The switch port channel number corresponding to the second node being retired
<<var_retiring_node_01_port>>	Switch port corresponding to the first node being retired
<<var_retiring_node_02_port>>	Switch port corresponding to the second node being retired
<<var_vserver>>	Vserver name
<<var_volume_name>>	Volume name
<<var_dest_aggr_name>>	Destination aggregate name
<<var_cutover_time_in_sec>>	Cutover time in seconds

Scenario

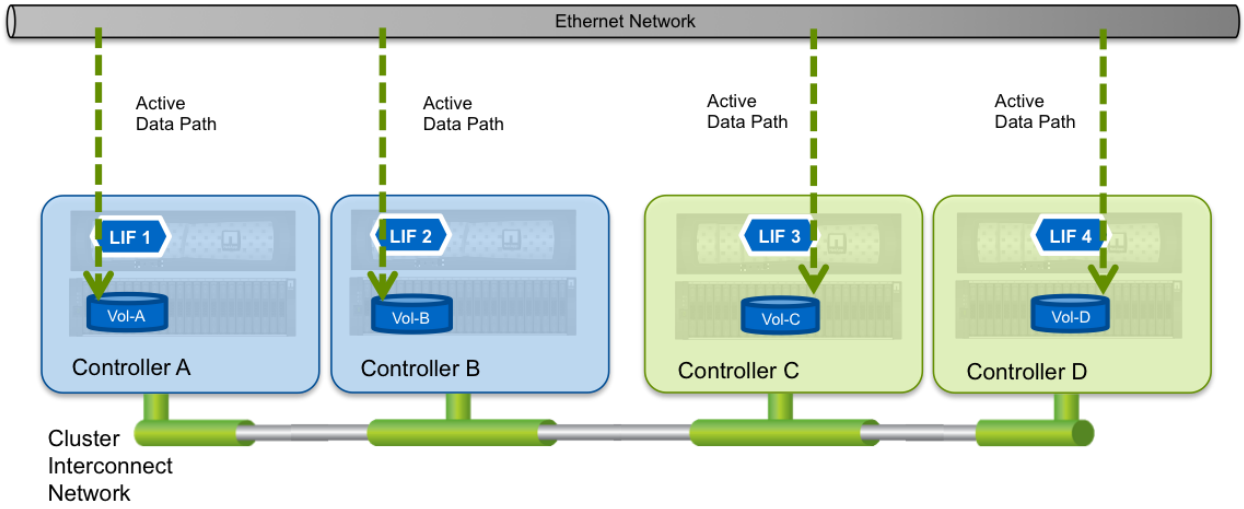
As part of the lifecycle of a FlexPod unit, controllers might need to be permanently retired from a cluster. Retiring an HA pair of controllers from a cluster is a nondisruptive activity.

Overview

Before

In a typical clustered Data ONTAP cluster, multiple nodes can contain multiple LIFs and volumes, as shown in the following NAS-only example.

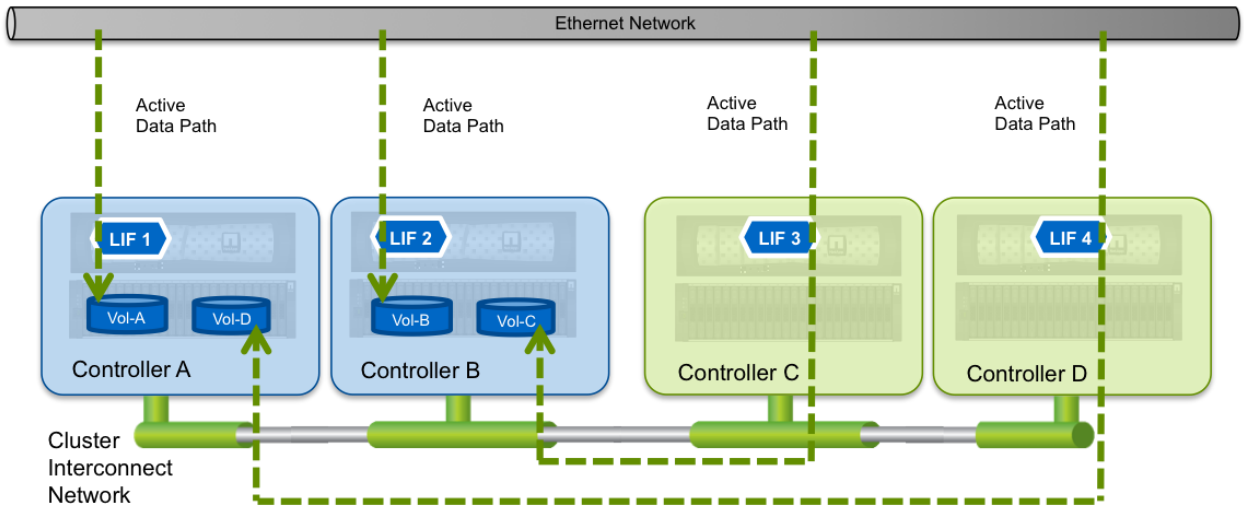
Figure 20) Before.



During

If, for example, an HA pair of controllers is coming off lease, the administrator must take a few steps to migrate both the data and the data paths from those controllers in order to remove them from the cluster. The first step involves migrating the volumes using the “volume move” procedure to any surviving nodes in the cluster that can handle both the space and I/O requirements.

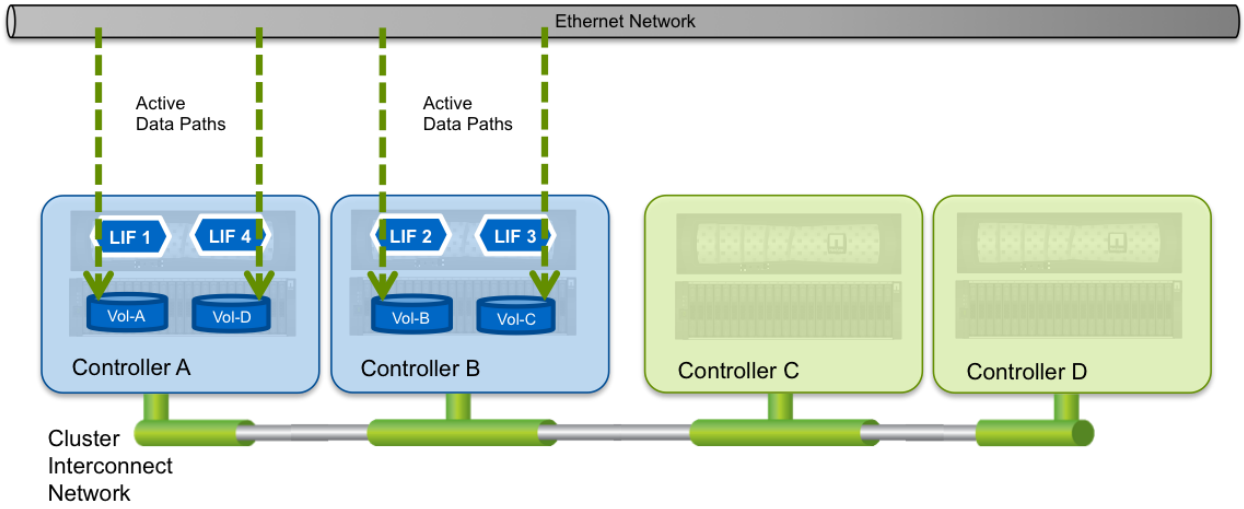
Figure 21) During 1.



During 2

After the volumes containing the data have been migrated, the LIFs must be migrated as well. When that completes, all active data and paths to that data will have been migrated off of the HA pair of controllers that will be removed from the cluster.

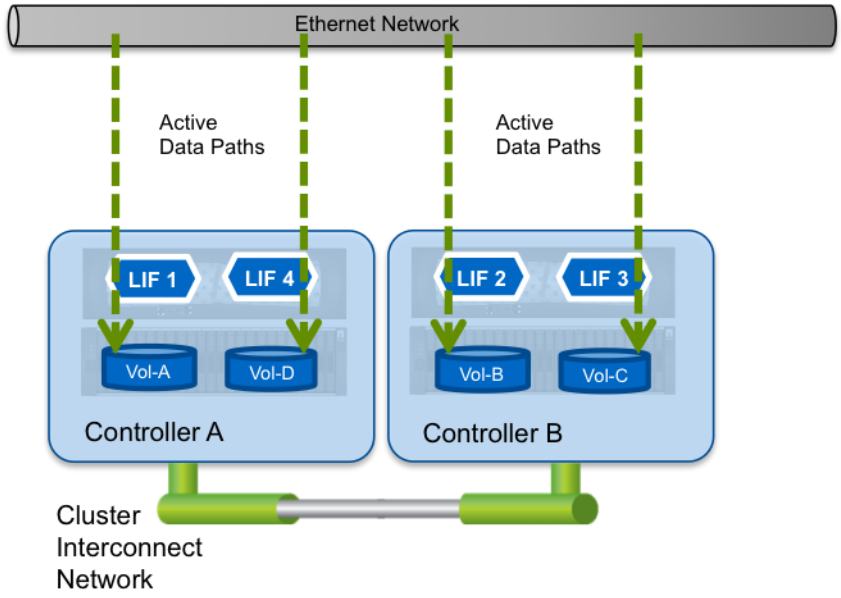
Figure 22) During 2.



After

Finally, the HA pair may now be removed, and the surviving nodes continue to serve data from all of the volumes and LIFs that existed at the beginning of this procedure.

Figure 23) After.



Dependencies

Dependencies	Description
Verify that the other nodes in the cluster can accommodate the LIFs currently owned by the nodes that will be retired.	The other nodes in the cluster must have the same port types. The total number of LIFs in the cluster must not exceed the total number of LIFs allowed for the remaining nodes in the cluster.

Dependencies	Description
Verify that the other nodes in the cluster can accommodate the volumes currently owned by the nodes that will be retired. **Note:** This step verifies that enough space exists but does not verify that the required IOPS headroom is present. See section "Measuring Utilization" for additional guidance for verifying IOPS availability.	Use `aggr show -nodes <<var_retiring_node01>>|<<var_retiring_node02>>` to display the aggregates owned by the nodes being retired. For each aggregate shown above, use the `volume show -aggregate <<var_aggregate_name>> -fields size`, used to determine the volumes located on aggregates that will be retired. The sum of the Used column for all of the volumes shown must be available in aggregates owned by other nodes. The sum of the Size column for all of the volumes might be necessary depending on the space guarantee of the volumes.
All volumes must be evacuated from the nodes being retired.	See section "Workload Rebalancing" for details on nondisruptively moving volumes away from aggregates owned by the nodes being retired.
All LIFs must be evacuated from the nodes being retired.	See section "Workload Rebalancing" for details on nondisruptively moving volumes away from aggregates owned by the nodes being retired.

Retiring a Controller

After the dependencies have been met, the following procedure should be used to retire both nodes of the HA pair that needs to be retired.

1. Use the following command to confirm that volumes and LIFs have been removed from the controller:

```
network interface show -curr-node <<var_controller_name>>
```

Example output:

```
            Logical    Status     Network            Current     Current Is
Vserver     Interface  Admin/Oper Address/Mask       Node        Port    Home
----------- ---------- ---------- ------------------ ----------- ------- --
node-01
            clus1        up/up    169.254.168.97/16  node-01     e1a     true
            clus2        up/up    169.254.89.245/16  node-01     e2a     true
            mgmt1        up/up    10.251.112.47/26   node-01     e0M     true
```

Note: The only remaining LIFs should be the cluster LIFs and the node management LIF as shown above.

2. Make sure that no LIFs have the target controller set as the home node using the following command:

```
network interface show -home-node <<var_controller_name>>
```

Example output:

```
            Logical    Status     Network            Current     Current Is
Vserver     Interface  Admin/Oper Address/Mask       Node        Port    Home
----------- ---------- ---------- ------------------ ----------- ------- --
node-01
            clus1        up/up    169.254.168.97/16  node-01     e1a     true
            clus2        up/up    169.254.89.245/16  node-01     e2a     true
            mgmt1        up/up    10.251.112.47/26   node-01     e0M     true
```

Note: The only LIFs that appear should be the cluster LIFs and node management LIF as shown above. If other LIFs appear in the list, use `network interface modify -vserver`

`<<var_vserver_name>> -lif <<var_lif_name>> -home-node <<var_new_node_name>> -home-port <<var_home_port_name>>` to change the home node and port of a LIF.

3. Make sure that the target node does not have epsilon. Use the following commands to determine which controller has epsilon and to move it if necessary.
   a. First, change to advanced privilege mode.

```
set -privilege adv

Warning: These advanced commands are potentially dangerous; use them only when directed to do so
by NetApp personnel.
Do you want to continue? {y|n}: y
```

   b. Use the following command to display which controller owns epsilon.

```
cluster show
```

   Example output:

```
Node            Health  Eligibility   Epsilon
--------------- ------- ------------  ------------
node-01         true    true          true
node-02         true    true          false
node-03         true    true          false
node-04         true    true          false
4 entries were displayed.
```

   c. Use the following commands to remove and reassign epsilon:

```
cluster modify -node <<var_controller_name>> -epsilon false
cluster modify -node <<var_other_node>> -epsilon true
```

Note: See Understanding quorum and epsilon for more details.

4. Remove the node from the cluster using the following command. Type y and press Enter to continue when prompted.

```
cluster unjoin -node <<var_controller_name>>
```

   Example output:

```
Warning: This command will unjoin node "node-01" from the cluster. You must unjoin the failover
         partner as well. After the node is successfully unjoined, erase its configuration and
         initialize all disks by using the "Clean configuration and initialize all disks (4)"
option
         from the boot menu.
Do you want to continue? {y|n}: y
[Job 7111] Job is queued: Cluster unjoin of Node:node-01 with UUID:a8bae730-a295-11dc-b8a8-
c9e14c3[Job 7111] Cleaning cluster database[Job 7111] Job succeeded: Cluster unjoin succeeded
If applicable, also unjoin the node's HA partner, and then clean its configuration and initialize
all disks via the boot menu. Run "debug vreport" to address any remaining aggregate or volume
issues.
```

5. From the serial connection to the node, the following will be displayed after the node reboots. At this point the controller can be safely turned off and decabled.

```
******************************
*                            *
* Press Ctrl-C for Boot Menu. *
*                            *
******************************

This node was removed from a cluster. Before booting, use
option (4) to initialize all disks and setup a new system.
Normal Boot is prohibited.

Please choose one of the following:
```

```
(1) Normal Boot.
(2) Boot without /etc/rc.
(3) Change password.
(4) Clean configuration and initialize all disks.
(5) Maintenance mode boot.
(6) Update flash from backup config.
(7) Install new software first.
(8) Reboot node.
Selection (1-8)?
```

Note: If the remaining cluster has only two nodes, then cluster HA should be turned on by using the command `cluster ha modify -configured true`.

6. Repeat steps 1 through 4 above for the HA partner of the node that was retired.

Configuring the Switch

1. Remove port channels, vPCs, and vFC and shut down inactive interfaces on the Cisco Nexus data switches using the following commands on each switch.

```
conf t
no vpc <<retiring_node01_vpc>>
no vpc <<retiring_node02_vpc>>
no interface vfc <<retiring_node01_vfc>>
no interface vfc <<retiring_node02_vfc>>
no interface port-channel <<retiring_node_01_port-channel_number>>
no interface port-channel <<retiring_node_02_port-channel_number>>
interface ethernet <<retiring_node_01_port>>
no description
shutdown
exit
interface ethernet <<retiring_node_02_port>>
no description
shutdown
exit
end
copy run start
```

Appendix

Aggregate Relocation-Based Controller Upgrade

In a FlexPod unit running clustered Data ONTAP 8.2 or higher, it is possible to nondisruptively upgrade controller hardware between certain FAS models. The method might be faster and less complex than expanding, migrating workloads, and contracting a cluster, especially in environments with large volumes or large quantities of volumes and LIFs.

Aggregate relocation (ARL)-based controller upgrades might in some cases better suit an environment than temporarily expanding, migrating workloads, and contracting a cluster. For example, in situations where a data center might be space, power, or cooling limited, or in cases where the controllers must be replaced but the disks and shelves do not require replacement, an ARL-based controller upgrade might better meet business needs.

Consult with your NetApp sales or support representative for more details on this procedure.

Workload Rebalancing

If FlexPod has been modified to address performance concerns for existing workloads, additional procedures might be necessary to begin effectively using the newly acquired resources.

The following rebalancing procedures may generally be performed nondisruptively and can be categorized as follows:

Table 13) Workload rebalancing procedures.

Procedure	Alleviates
Relocating logical storage objects (LUNs, volumes, and so on)	Overutilization of disk I/O
Relocating logical storage objects (LUNs, volumes, and so on)	Overutilization of controller CPU I/O
Relocating data path objects (LIFs)	Overutilization of network path I/O

Either the affected workload or neighboring workloads using the same network/controller/disk resources can be moved to resolve performance issues. Many considerations should be made to determine which and whether a workload should be moved. This includes the size and data access patterns of the workload, the workload's sensitivity to minor data traffic disruptions, and other Data ONTAP features.

Measuring Utilization

Because workload rebalancing can be nondisruptively performed, rebalancing should be considered both during lifecycle events and during day-to-day operations. Overutilization of resources might cause performance issues and can be avoided by intelligently rebalancing workloads.

Prior to performing any rebalancing operation, it is important to understand how the target node, network port, or aggregate will be affected by the change. NetApp offers a variety of tools and resources for measuring and estimating the impact of workloads.

Go to http://support.netapp.com or contact Technical Support for additional tools and resources for measuring utilization.

Rebalancing Workload for Volumes and LUNs

Volumes can be moved from one aggregate to another. Relocating a volume to another aggregate owned by the same controller might alleviate issues due to disk overutilization. Relocating a volume to another aggregate owned by a different controller might alleviate both disk overutilization and controller CPU utilization issues.

LUNs in a volume will be automatically moved when the containing volume is moved. It is not possible to move a LUN independently of the volume.

Moving a Volume

This procedure assumes that the reader has:

- Determined the workload that should be relocated and its corresponding volumes.
- Made sure that the host timeout setting for the workload is greater than the cutover time used during the move operation.
- Determined an appropriate destination aggregate.
- Cluster administrator privileges with CLI access to the cluster.

Note: This procedure is not recommended if the workload host has a timeout setting less than 30 seconds.

1. Determine the aggregate currently housing the volume that will be moved.

```
volume show -volume <<var_volume_name>>
```

2. Determine whether the chosen target aggregate is valid for the volume move.

```
volume move target-aggr show -vserver <<var_vserver>> -volume <<var_volume_name>>
```

Example output:

```
Aggregate Name   Available Size   Storage Type
```

```
--------------   --------------   ------------
aggr01           4.35TB           SAS
aggr02           6.69TB           SAS
aggr03           7.11TB           SAS
3 entries were displayed.
```

Note: Aggregates must have enough space available, belong to a node in the same cluster as the current aggregate, and belong to the Vserver that owns the volume.

3. View the aggregates that a Vserver may use.

```
vserver show -fields aggr-list
```

Example output:

```
vserver       aggr-list
------------- --------------------------
Infra_Vserver aggr01,aggr02,aggr03,aggr04
```

4. Add the new aggregate without removing the existing aggregates from the list of aggregates available to the Vserver.

```
vserver modify  -aggr-list <list-from-previous-output>,<new-aggregate-name>
```

5. Move a target volume:

```
vol move start -vserver <<var_vserver>> -volume <<var_volume_name>> -destination-aggregate
<<var_dest_aggr_name>> -cutover-window <<var_cutover_time_in_seconds>>
```

6. Monitor the status of the volume move.

```
volume move show
```

Example output:

```
Vserver   Volume     State    Move Phase Percent-Complete Time-To-Complete
--------- ---------- -------- ---------- ---------------- ----------------
Infra_Vserver
          Nfsvol     healthy  cutover     -                -
```

7. After the volume move is completed successfully, running the `volume move show` command will display:

```
This table is currently empty.
```

8. This procedure is finished after the cutover is complete. If the volume now resides on an aggregate owned by a node different from the LIF used to access the volume, it is highly recommended to move the LIF as well.

Rebalancing NAS LIFs

NAS (NFS/CIFS) network interfaces may be nondisruptively moved to different ports, ifgrps, or VLAN interfaces on the same controller or on a different controller. The FlexPod architecture recommends using ifgrps and VLAN interfaces that correspond to virtual port channels within the network layer.

If a FlexPod unit has been consistently configured according to best practices, the network layer VLANs should be consistent across controllers. The following steps can be taken to make sure that the correct configuration is in place. This should be performed prior to moving a LIF to protect against a disruptive data path failure.

Verifying That VLANs Are Consistently Mapped

This procedure assumes that the administrator has CLI access to both switches within the FlexPod unit. It assumes that FlexPod best practices have been followed with regard to the creation of port channels, virtual port channels, and ifgrps.

1. Log in to the Cisco Nexus switch with administrative access.

2. If interfaces and port channels have been named according to best practices, use the following command on each switch to determine which ports correspond to the target node.

```
Show interface status
```

Example output:

```
--------------------------------------------------------------------------------
Port          Name            Status    Vlan      Duplex  Speed   Type
--------------------------------------------------------------------------------
Eth1/1        node-01:e3a     connected trunk     full    10G     SFP-H10GB-C
Eth1/2        node-02:e4a     connected trunk     full    10G     SFP-H10GB-C
[Output Omitted]
Po10          vPC Peer-link   connected trunk     full    10G     --
Po11          node-01         connected trunk     full    10G     --
Po12          node-01         connected trunk     full    10G     --
[Output Omitted]
```

Note: In this example, each controller has a single ifgrp for all SAN/NAS interfaces. Each ifgrp corresponds to a vPC configured between the two data network switches in the FlexPod unit.

Note: If interfaces and port channels have not been named according to best practices, use `show cdp neighbors` to determine which ports correspond to which controllers and then run `show running-config interface <<var_interface_name>>` to determine the port channel membership of each port.

3. Use the following command on either switch to make sure that the virtual port channels between the two switches are operational and configured to trunk the same VLANs:

```
show vpc brief
```

Example output:

```
Legend:
                (*) - local vPC is down, forwarding via vPC peer-link

vPC domain id                     : 23
Peer status                       : peer adjacency formed ok
vPC keep-alive status             : peer is alive
Configuration consistency status: success
Per-vlan consistency status       : success
Type-2 consistency status         : success
vPC role                          : primary
Number of vPCs configured         : 9
Peer Gateway                      : Disabled
Dual-active excluded VLANs        : -
Graceful Consistency Check        : Enabled

vPC Peer-link status
---------------------------------------------------------------------
id   Port   Status Active vlans
--   ----   ------ --------------------------------------------------
1    Po10   up     1905,3191-3194

vPC status
----------------------------------------------------------------------------
id     Port        Status Consistency Reason                     Active vlans
------ ----------- ------ ----------- -------------------------- -----------
11     Po11        up     success     success                    1905,3191-3
                                                                 194
12     Po12        up     success     success                    1905,3191-3
                                                                 194
[Output omitted]
```

Note: The Active VLANs column should display the same set of VLANs for the vPCs that correspond to the port channels identified in the previous step.

4. If any VLANs need to be added, add them to the port channel on each switch using the following commands:

```
conf t
interface port-channel <<var_port_channel_number>>
switchpot trunk allowed vlan add <<var_vlan_number>>
end
copy run start
```

Note: Because the VLAN is already in use and trunked to another controller, it should already exist on both switches and be trunked by the vPC peer links of both switches. Use `show vlan` to review the existence of the VLAN on both switches and `show running-config interface port-channel <port-channel-number>` for the vPC peer link port channel to confirm.

Modify and Migrate a LIF

The FlexPod architecture generally uses ifgrps and virtual port channels on the FAS controllers and Cisco Nexus switches, respectively. This procedure assumes that the reader has determined the workload that should be relocated and its corresponding LIFs.

1. Verify the current node and port and the home node and port for the LIF being relocated.

```
network interface show -lif <<var_lif_name>> -fields curr-node,curr-port,home-node,home-port
```

2. Migrate the LIF to the target node and ifgrp or VLAN.

```
network interface migrate -vserver <<var_vserver_name>> -lif <<var_lif_name>> -dest-node
<<var_dest_node_name>> -dest-port <<var_dest_port_name>>
```

3. Modify the home node and port of the LIF.

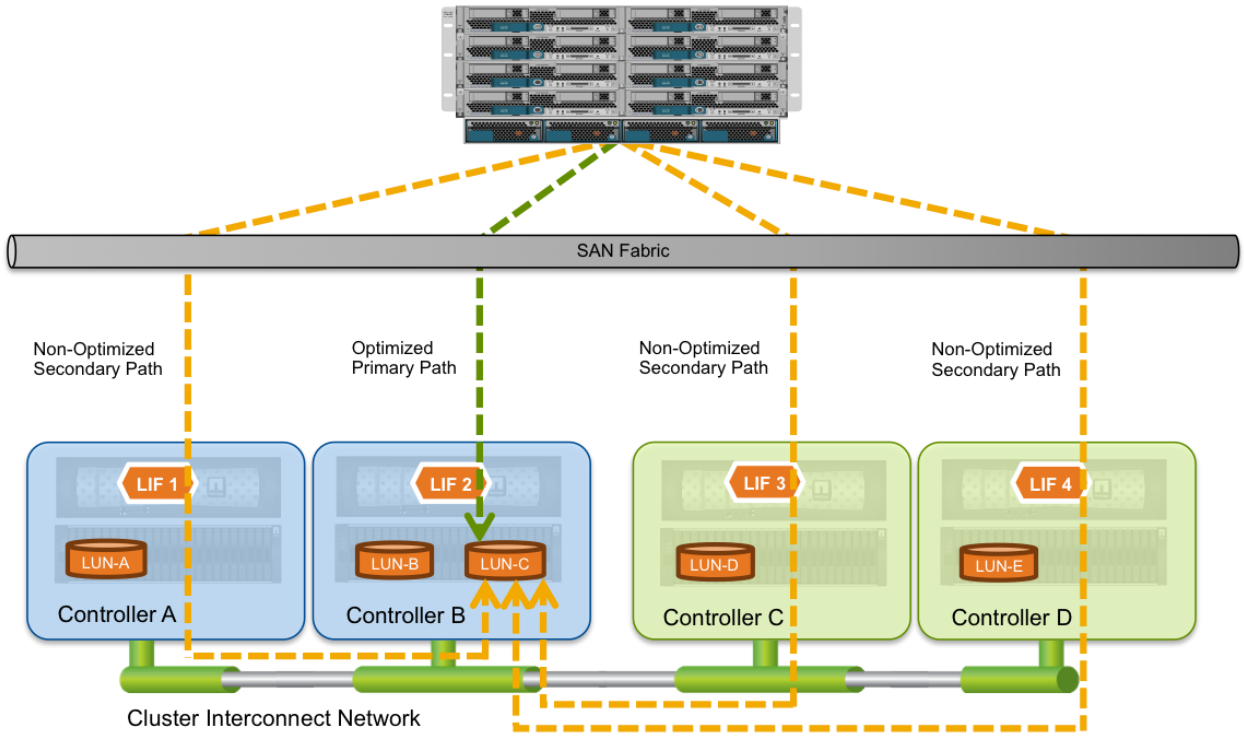
```
network interface modify -vserver <<var_vserver_name>> -lif <<var_lif_name>> -home-node
<<var_home_node_name>> -home-port <var_home_port_name>>
```

4. Use `network interface show -vserver <<var_vserver_name>> -lif <<var_lif_name>>` to confirm that the LIF has been moved and is operational.

Rebalancing SAN LIFs

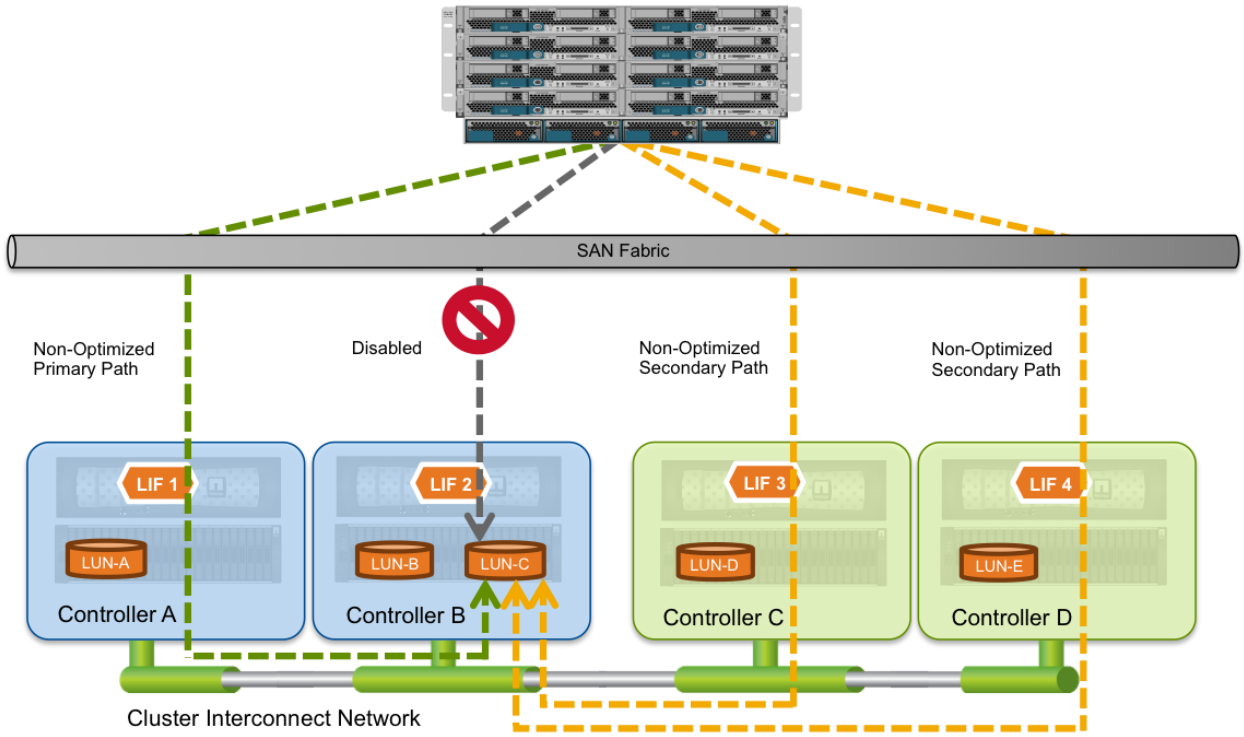
SAN LIFs can be relocated to other ports on the same controller or to other ports on another controller within the same cluster, but they must be inactive to be relocated. If the workload supports ALUA, then within a FlexPod unit configured according to best practices, this is a nondisruptive operation.

Figure 24) Before.



Temporarily making a SAN LIF inactive to move it within a cluster is a nondisruptive operation within a FlexPod unit because FlexPod prescribes that a LUN be accessible by at least two LIFs on at least two controllers within a cluster. Because the FlexPod unit and workload rely on ALUA to find the most optimized path, link failures will trigger path selection automatically. See Figure 25 for an example.

Figure 25) After.



No LIF movements are required if a LUN is being moved to a controller that already has SAN LIFs in the correct zone or VLAN. Path optimization will occur automatically and without disruption to the workload.

When moving a LUN to a controller that does not yet have SAN LIFs in the correct zone or VLAN, it is recommended that the LIFs of the nonoptimized path be relocated to the target controller prior to moving the volume. After the volume has been moved, the LIFs that were on the original controller should be relocated to the HA pair of the target controller.

Note: Alternatively, new LIFs could be created on the destination controller. This could require substantial configuration changes to the network and compute layers and will not be covered in this document.

Verifying That VLANs Are Correctly and Consistently Mapped (iSCSI)

This procedure assumes that the administrator has CLI access to both switches within the FlexPod unit. This procedure assumes that FlexPod best practices have been followed with regard to the creation of port channels, virtual port channels, and ifgrps.

1. Log in to the Cisco Nexus switch with administrative access.
2. If interfaces and port channels have been named according to best practices, use the following command on each switch to determine which ports correspond to the target node.

```
Show interface status
```

Example output:

```
--------------------------------------------------------------------------------
Port          Name            Status    Vlan      Duplex  Speed   Type
--------------------------------------------------------------------------------
Eth1/1        node-01:e3a     connected trunk     full    10G     SFP-H10GB-C
Eth1/3        node-03:e3a     connected trunk     full    10G     SFP-H10GB-C
[Output Omitted]
Po10          vPC Peer-link   connected trunk     full    10G     --
```

```
Po11          node-01         connected trunk     full    10G     --
Po13          node-03         connected trunk     full    10G     --
[Output Omitted]
```

Note: In this example, each controller has a single ifgrp for all SAN/NAS interfaces. Each ifgrp corresponds to a vPC configured between the two switches.

Note: If interfaces and port channels have not been named according to best practices, use `show cdp neighbors` to determine which ports correspond to which controllers and then run `show running-config interface <<var_interface_name>>` to determine the port channel membership of each port.

3. Use the following command on either switch to make sure that the virtual port channels between the two switches are operational and configured to trunk the same VLANs.

```
show vpc brief
```

Example output:

```
Legend:
                (*) - local vPC is down, forwarding via vPC peer-link

vPC domain id                     : 23
Peer status                       : peer adjacency formed ok
vPC keep-alive status             : peer is alive
Configuration consistency status: success
Per-vlan consistency status     : success
Type-2 consistency status       : success
vPC role                          : primary
Number of vPCs configured       : 9
Peer Gateway                      : Disabled
Dual-active excluded VLANs      : -
Graceful Consistency Check      : Enabled

vPC Peer-link status
---------------------------------------------------------------------
id   Port   Status Active vlans
--   ----   ------ -----------------------------------------------------
1    Po10   up     1905,3191-3194

vPC status
----------------------------------------------------------------------------
id     Port        Status Consistency Reason                     Active vlans
------ ----------- ------ ----------- -------------------------- -----------
11     Po11        up     success     success                    911-912,
                                                                  1905,3191-3
                                                                  194
13     Po13        up     success     success                    911-912,
                                                                  1905,3191-3
                                                                  194
[Output omitted]
```

Note: The active VLANs column should display the same set of VLANs for the vPCs that corresponds to the port channels identified in the previous step. The VLANs associated with iSCSI fabric A and iSCSI fabric B should appear in both sets of lists.

4. If any VLANs need to be added, add them to the port channel on each switch.

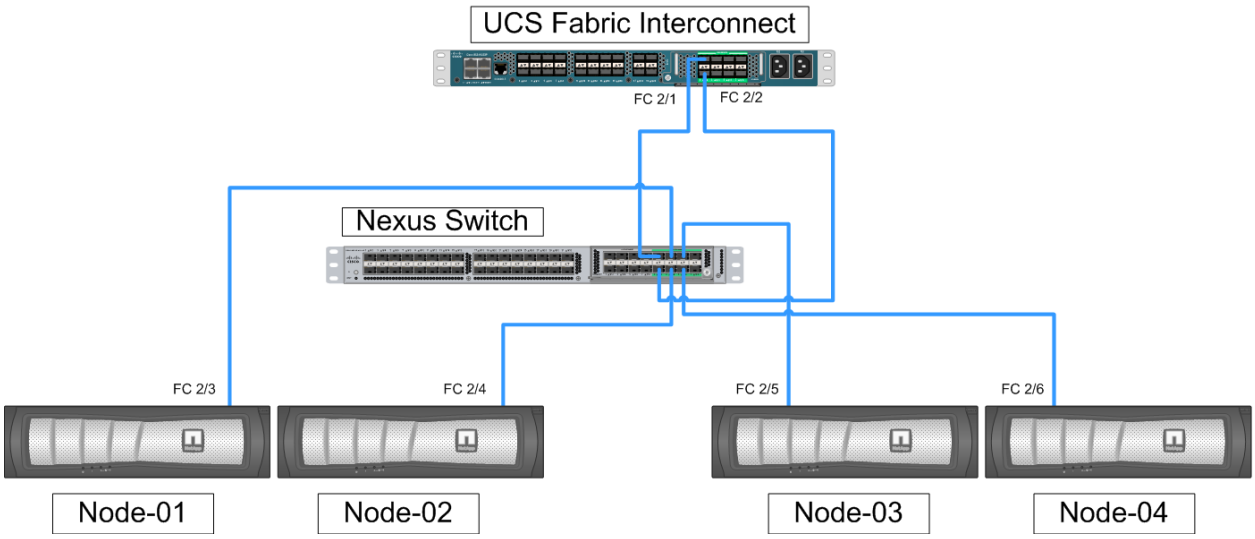
```
conf t
interface port-channel <<var_port-channel-number>>
switchpot trunk allowed vlan add <<var_vlan-number>>
end
copy run start
```

Note: Because the VLAN is already in use and trunked to another controller, it should already exist on both switches and be trunked by the vPC peer links of both switches. Use `show vlan` to review the existence of the VLAN on both switches and `show running-config interface port-channel <port-channel-number>` for the vPC peer link port channel to confirm.

Verifying That VSANs Are Correctly Configured (FC)

This procedure assumes that the administrator has CLI access to both Cisco Nexus switches in the environment. This procedure also assumes that each controller in the cluster has an FC port connected to each Cisco Nexus switch. See Figure 26 for an example.

Figure 26) Verify FC port configuration.



1. On each switch, review the VSANs configured and confirm that the FC port associated with the source controller is in the same VSAN as the port associated with the target controller. In this case node-01 is connected to port FC 2/3 and node-03 is connected to port FC 2/5. Use the following command to check the settings:

```
show vsan membership
```

Example output:

```
vsan 110 interfaces:
    fc2/1            fc2/2            fc2/3            fc2/4
    fc2/5            fc2/6            san-port-channel 1
vsan 4079(evfp_isolated_vsan) interfaces:

vsan 4094(isolated_vsan) interfaces:
```

2. If necessary use the following commands to add the FC port to the correct VSAN:

For Switch 1	For Switch 2
conf t vsan database vsan <<var_vsan_A>> interface fc <<var_interface_slot>>/<<var_interface_number>> end copy run start	conf t vsan database vsan <<var_vsan-b>> interface fc <<var_interface_slot>>/<<var_interface_number>> end copy run start

3. Additionally, the device alias database on both switches should be updated to reflect the new LIF location for each LIF that is relocated. Update the device alias for a given LIF using the following commands:

```
conf t
device-alias database
```

4. For each LIF that needs to be renamed, run the following command:

```
no device-alias name <<var_old_LIF_name>>
device-alias name <<var_controller_name_port_name_lif_name>> pwwn <var_LIF_pwwn>>
```

5. Commit the changes and save the configuration.

```
device-alias commit
end
copy run start
```

Making Sure That FCoE VSANs Are Correctly Mapped (FCoE)

This procedure assumes that the administrator has CLI access to both Cisco Nexus switches in the environment. This procedure also assumes that each controller in the cluster has a UTA port connected to each Cisco Nexus switch.

1. On each switch, review the VSANs configured on the switch and confirm that the virtual Fibre Channel (vFC) port associated with the source controller is in the same VSAN as the vFC port associated with the target controller. In this example node-01 is connected to vfc 11 and node-03 does not yet have a vFC created. Use the following command to verify the relationship between a vFC and controller.

```
show interface brief
```

   Example output:

```
[Output ommitted]
--------------------------------------------------------------------------------
Interface  Vsan   Admin  Admin   Status      Bind                Oper   Oper
                  Mode   Trunk               Info                Mode   Speed
                         Mode                                           (Gbps)
--------------------------------------------------------------------------------
vfc11      110    F     on     trunking    port-channel11           TF   auto
vfc12      110    F     on     trunking    port-channel12           TF   auto
[Output ommitted]
```

2. Assuming that port channels have been named correctly, use the following command to determine to which controller each port channel connects.

```
show interface status
```

   Example output:

```
--------------------------------------------------------------------------------
Port          Name               Status    Vlan      Duplex  Speed   Type
--------------------------------------------------------------------------------
[Output omitted]
Po10          vPC peer-link      connected trunk     full    10G     --
Po11          node-01            connected trunk     full    10G     --
Po12          node-02            connected trunk     full    10G     --
Po13          node-03            connected trunk     full    10G     --
Po14          node-04            connected trunk     full    10G     --
[Output omitted]
```

3. Use the following to determine the vFC membership in the appropriate VSAN.

```
show vsan membership
```

   Example output:

```
vsan 110 interfaces:
    fc2/1            fc2/2            san-port-channel-1
    vfc11            vfc12
vsan 4079(evfp_isolated_vsan) interfaces:

vsan 4094(isolated_vsan) interfaces:
```

4. If necessary, use the following commands to add and bind a virtual Fibre Channel port to the FCoE port connected to the controller and add the virtual FC port to the correct VSAN.

For Switch 1	For Switch 2
conf t interface vfc <<var_vfc-number>> desc <<var_controller-name>>:FCOE bind interface port-channel <<var_port-channel-number>> switchport trunk allowed vsan <<var_vsan-number>> no shutdown vsan database vsan <<var_vsan-A>> interface vfc <<var_vfc-number>> end copy run start	conf t interface vfc <<var_vfc-number>> desc <<var_controller-name>>:FCOE bind interface port-channel <<var_port-channel-number>> switchport trunk allowed vsan <<var_vsan-number>> no shutdown vsan database vsan <<var_vsan-b>> interface fc <<var_interface-slot>>/<<var_interface-number>> end copy run start

5. Additionally, the device alias database on both switches should be updated to reflect the new LIF location for each LIF that is relocated. For each switch, update the device alias for each moved LIF using the following commands:

```
conf t
device-alias database
```

6. For each LIF that needs to be renamed:

```
no device-alias name <<var_old-LIF-name>>
device-alias name <<var_controller-name-port-name-lif-name>> pwwn <<var_LIF_pwwn>>
```

7. Finally, commit the changes and save the configuration.

```
device-alias commit
end
copy run start
```

Disabling, Moving, and Reenabling a SAN LIF

The following procedure should be executed for each LIF that needs to be relocated. SAN LIFs should not be moved at the same time.

1. To move a SAN LIF, it must first be disabled. Disable a SAN LIF using the following command:

```
network interface modify -vserver <<var_vserver_name>> -lif <<var_lif_name>> -status-admin down
```

2. After the SAN LIF has been disabled, it can be relocated to the correct controller on the correct FC/FCoE port or VLAN interface. Relocate a SAN LIF using the following command:

```
network interface modify -vserver <<var_vserver_name>> -lif <<var_lif_name>> -home-node <<target-node-name>> -home-port <<var_port_name>>
```

3. After the LIF is moved, it can be safely returned to an operational state using the following command:

```
network interface modify -vserver <<var_vserver_name>> -lif <<var_lif_name>> -status-admin up
```

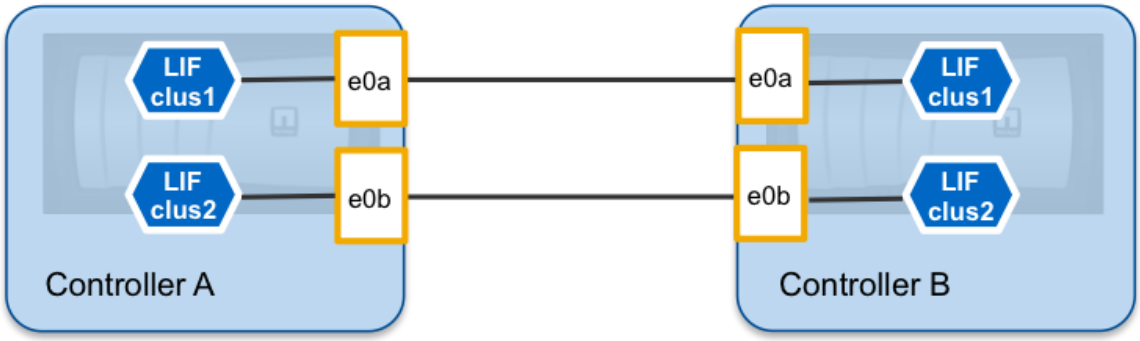
4. Confirm that the LIF is operational and on the correct controller and port using the following command:

```
network interface show -vsrever <<var_vserver_name>> -lif <<var_lif_name>>
```

Creating a Switchless Cluster

Clustered Data ONTAP supports creating a switchless two-node cluster. Rather than requiring separate cluster interconnect switches, each of the redundant cluster interconnect network ports on one controller will connect directly to a corresponding cluster interconnect network port on the other controller in the HA pair, providing a back-to-back connection.

Figure 27) Switchless cluster configuration.



The following steps should be followed on the first node in the cluster.

Note: The nodes must be running Data ONTAP 8.2 or later.

Note: The switchless cluster feature cannot be used with more than two nodes.

Note: Each storage system must have two dedicated cluster ports providing redundant cluster-network connections.

1. During the first node boot, the Cluster Setup wizard starts running on the console.

```
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}:
```

Note: If a login prompt appears instead of the Cluster Setup wizard, start the wizard by logging in using the factory default settings and then enter the `cluster setup` command.

2. Enter the following command to create a new cluster:

```
create
```

3. Follow the below steps to activate HA and set storage failover.

```
Do you intend for this node to be used as a single node cluster? {yes, no} [no]: Enter

Non-HA mode, Reboot node to activate HA

Do you want to reboot now to set storage failover (SFO) to HA mode? {yes, no}
[yes]: Enter
```

4. After the reboot, continue with cluster create.
5. The system defaults are displayed.

```
System Defaults:
Private cluster network ports [e1a,e1b].
Cluster port MTU values will be set to 9000.
Cluster interface IP addresses will be automatically generated.
The cluster will be connected using network switches.

Do you want to use these defaults? {yes, no} [yes]:
```

5. NetApp recommends accepting the system defaults. To accept the system defaults, press Enter.

Note: Cluster is created; this can take a minute or two.

6. The steps to create a cluster are displayed.

```
Enter the cluster name: <<var_clustername>>
Enter the cluster base license key: <<var_cluster_base_license_key>>
Creating cluster <<var_clustername>>
Enter an additional license key []:
```

Note: For this validated architecture we recommend you install license keys for SnapRestore®, NFS, FCP, FlexClone®, and SnapManager® Suite. After you finish entering the license keys, press Enter.

```
Enter the cluster administrators (username "admin") password: <<var_password>>
Retype the password: <<var_password>>
Enter the cluster management interface port [e0a]: e0a
Enter the cluster management interface IP address: <<var_clustermgmt_ip>>
Enter the cluster management interface netmask: <<var_clustermgmt_mask>>
Enter the cluster management interface default gateway: <<var_clustermgmt_gateway>>
```

7. Enter the DNS domain name.

```
Enter the DNS domain names:<<var_dns_domain_name>>
Enter the name server IP addresses:<<var_nameserver_ip>>
```

Note: If you have more than one name server IP address, separate them with a comma.

8. Set up the node.

```
Where is the controller located []:<<var_node_location>>
Enter the node management interface port [e0M]: Enter
Enter the node management interface IP address: <<var_node01_mgmt_ip>>
Enter the node management interface netmask:<<var_node01_mgmt_mask>>
Enter the node management interface default gateway:<<var_node01_mgmt_gateway>>
```

Note: The node management interface should be in a different subnet than the cluster management interface. The node management interfaces can reside on the out-of-band management network, and the cluster management interface can be on the in-band management network.

9. Press Enter to accept the AutoSupport message.
10. Log in to the node and set the privilege mode to advanced.

```
set -privilege advanced
```

11. Enable the switchless cluster.

```
network options switchless-cluster modify true
```

12. Make sure the switchless cluster is enabled.

```
FAS2240-Cluster::*> network options switchless-cluster show

Enable Switchless Cluster: true
```

13. Set the privilege mode back to admin.

```
set -privilege admin
```

14. Reboot node 01.

```
system node reboot –node <<var_node01>>
Warning: Are you sure you want to reboot the node? {y|n}: y
```

15. When you see `Press Ctrl-C for Boot Menu`, enter:

```
Ctrl – C
```

16. Select 5 to boot into maintenance mode.

```
5
```

17. When prompted `Continue with boot?`, enter `y`.
18. To verify the HA status of your environment, run the following command:

```
ha-config show
```

Note: If either component is not in HA mode, use the `ha-config modify` command to put the components in HA mode.

19. Reboot the controller.

```
halt
```

20. At the `LOADER-A` prompt, enter:

```
autoboot
```

21. Log in to the cluster.
22. Data ONTAP would assign disks to Storage Controllers automatically if the disk autoassign setting was turned on. Use the `storage disk option show -fields autoassign` command to verify the setting.
23. If disk autoassign was turned on, skip to section 3.3 ,“Grow a Cluster.” Else continue with step 24.
24. Reboot node 01.

```
system node reboot –node <<var_node01>>
Warning: Are you sure you want to reboot the node? {y|n}: y
```

25. When you see `Press Ctrl-C for Boot Menu`, enter:

```
Ctrl – C
```

26. Select 5 to boot into maintenance mode.

```
5
```

27. When prompted `Continue with boot?`, enter `y`.
28. To see how many disks are unowned, enter:

```
disk show -a
```

Note: No disks should be owned in this list.

29. Assign disks.

Note: This reference architecture allocates half the disks to each controller. However, workload design could dictate different percentages.

```
disk assign –n <<var_#_of_disks>>
```

30. Reboot the controller.

```
halt
```

31. At the `LOADER-A` prompt, enter:

```
autoboot
```

Refer to the Interoperability Matrix Tool (IMT) on the NetApp Support site to validate that the exact product and feature versions described in this document are supported for your specific environment. The NetApp IMT defines the product components and versions that can be used to construct configurations that are supported by NetApp. Specific results depend on each customer's installation in accordance with published specifications.

NetApp provides no representations or warranties regarding the accuracy, reliability, or serviceability of any information or recommendations provided in this publication, or with respect to any results that may be obtained by the use of the information or observance of any recommendations provided herein. The information in this document is distributed AS IS, and the use of this information or the implementation of any recommendations or techniques herein is a customer's responsibility and depends on the customer's ability to evaluate and integrate them into the customer's operational environment. This document and the information contained herein may be used solely in connection with the NetApp products discussed in this document.

www.netapp.com