Technical Report

Best Practices for Scalable SAN in Clustered Data ONTAP 8.2

Benjamin Krieger, NetApp

June 2013 | TR-4080

Version 2.0

TABLE OF CONTENTS

1 Introduction....4

1.1 Audience....4

1.2 Caveats....4

2 Overview of Clustered Data ONTAP....5

2.1 Scalable SAN in Clustered Data ONTAP....5

2.2 Volume Configuration....5

2.3 Host Connectivity....6

2.4 Path Selection....7

2.5 Path Selection Changes....8

2.6 Scalable SAN Configuration Differences from 7-Mode....11

2.7 iSCSI and the Partner Interface....12

2.8 Fibre Channel and NPIV....12

2.9 Portsets....13

2.10 Management Interfaces....14

3 Scalable SAN Key Value Propositions....17

3.1 SVM as Unified Target and Unit of Management....17

3.2 Cluster-Wide Consistency Groups....18

3.3 Migrating Resources Between Cluster Nodes....18

4 Host Integration Examples....18

4.1 NetApp Host Utilities Kit....18

4.2 Microsoft Windows and Clustered Data ONTAP....19

4.3 Data ONTAP DSM....20

4.3.1 NetApp SnapDrive....22

4.4 UNIX or Linux and Clustered Data ONTAP....22

4.5 IBM AIX and Clustered Data ONTAP....24

Version History....25

References....25

LIST OF TABLES

Table 1) Configuration differences between clustered Data ONTAP and Data ONTAP operating in 7-Mode. 11

LIST OF FIGURES

Figure 4) MPIO and ALUA path changes during a port or switch failure. 9

Figure 5) MPIO and ALUA path changes during a volume move. 10

Figure 6) MPIO and ALUA path changes after a volume move.. 11

1 Introduction

NetApp® Clustered Data ONTAP® 8.2 is the second clustered Data ONTAP release to support SAN protocols after their introduction in 8.1. This paper presents an overview of the differences between 7-Mode and clustered SAN implementations from the point of view of SAN-attached hosts. It also describes the best practices for leveraging the high-availability and data mobility features of clustered Data ONTAP.

1.1 Audience

This paper is intended for system and storage architects who design iSCSI, Fibre Channel (FC), and Fibre Channel over Ethernet (FCoE) solutions with NetApp storage appliances running clustered Data ONTAP 8.2. It assumes that the reader:

- Has a general knowledge of NetApp hardware and software solutions
- Is somewhat familiar with block-access protocols such as Fibre Channel and iSCSI

1.2 Caveats

This document is not meant to be a general introduction to clustered Data ONTAP administration. This area is covered by the Clustered Data ONTAP System Administration Guide for Cluster Administrators and by the Clustered Data ONTAP SAN Administration Guide.

SAN-related limits for Data ONTAP clusters that use SAN protocols can be found in the Clustered Data ONTAP SAN Configuration Guide.

For the regularly updated and complete matrix of tested and supported SAN configurations, refer to the NetApp Support Matrix at http://support.netapp.com/matrix/.

2 Overview of Clustered Data ONTAP

Storage controllers running clustered Data ONTAP are referred to as *nodes.* These nodes are aggregated into a *clustered system.* The nodes in the cluster communicate with each other continuously, coordinate cluster activities, and move data transparently from node to node by using redundant paths to a dedicated cluster network that consists of two 10 Gigabit Ethernet switches.

Although the basic unit of a cluster is the node, nodes are added to the cluster as part of a high-availability (HA) pair. As with Data ONTAP operating in 7-Mode, HA pairs enable high availability by communicating with each other over an HA interconnect (separate from the dedicated cluster network) and by maintaining redundant connections to the HA pair's disks. Also, like Data ONTAP operating in 7-Mode, disks are not shared between HA pairs, although shelves may contain disks that belong to either member of an HA pair.

Clusters are administered on a whole-cluster rather than a per-node basis, and data is served from one or more storage virtual machines (SVMs). Each SVM is configured to own storage, in the form of volumes provisioned from a physical aggregate, and logical interfaces (LIFs), assigned either to a physical Ethernet network or to Fibre Channel target ports. Logical disks (LUNs) are created inside an SVM's volumes and mapped to hosts to provide them with storage space. SVMs are node independent and cluster based; they can make use of physical resources such as volumes or network ports anywhere in the cluster.

2.1 Scalable SAN in Clustered Data ONTAP

When an SVM is first created and a block protocol (FC or iSCSI) is enabled, the SVM gets either a Fibre Channel worldwide name (WWN) or an iSCSI qualified name (IQN), respectively. This identifier is used irrespective of which physical node is being addressed by a host, with Data ONTAP making sure that SCSI target ports on all of the cluster nodes work together to present a virtual, distributed SCSI target to hosts that are accessing block storage.

In practice, this means that no matter which physical node a host is communicating with, it is communicating with the same SCSI target. This method of access presents new opportunities for data resiliency and mobility, and it also has implications for best practices when accessing data by using block protocols on a cluster.

Best Practice

When creating iSCSI or Fibre Channel LIFs for the first time for an existing SVM, make sure that the Fibre Channel and/or iSCSI service for that SVM has been created and is turned on by using the `fcp show` or `iscsi show` command, or by navigating to the Cluster → Vserver → Configuration → Protocols pane in OnCommand® System Manager.

Note: This step is not necessary if the SVM was originally set up to serve these protocols by using either the `vserver setup` command or the System Manager Vserver Setup Wizard.

2.2 Volume Configuration

When provisioning volumes in a cluster or in Data ONTAP operating in 7-Mode, many considerations regarding deduplication, space reservations, and storage efficiency are the same. One major difference is that volumes in clustered Data ONTAP are oriented to SVM containers instead of to individual nodes, and a side effect is that they can be mapped into an SVM-wide global namespace for the purpose of exporting file systems by using NFS or CIFS protocols. However, the presence or absence of a given volume in the global namespace has no effect on data that is served by using Fibre Channel or iSCSI.

Best Practice

Volumes that contain LUNs do not need to be junctioned to the global namespace to serve data by using block protocols; they only require an igroup-to-LUN mapping.

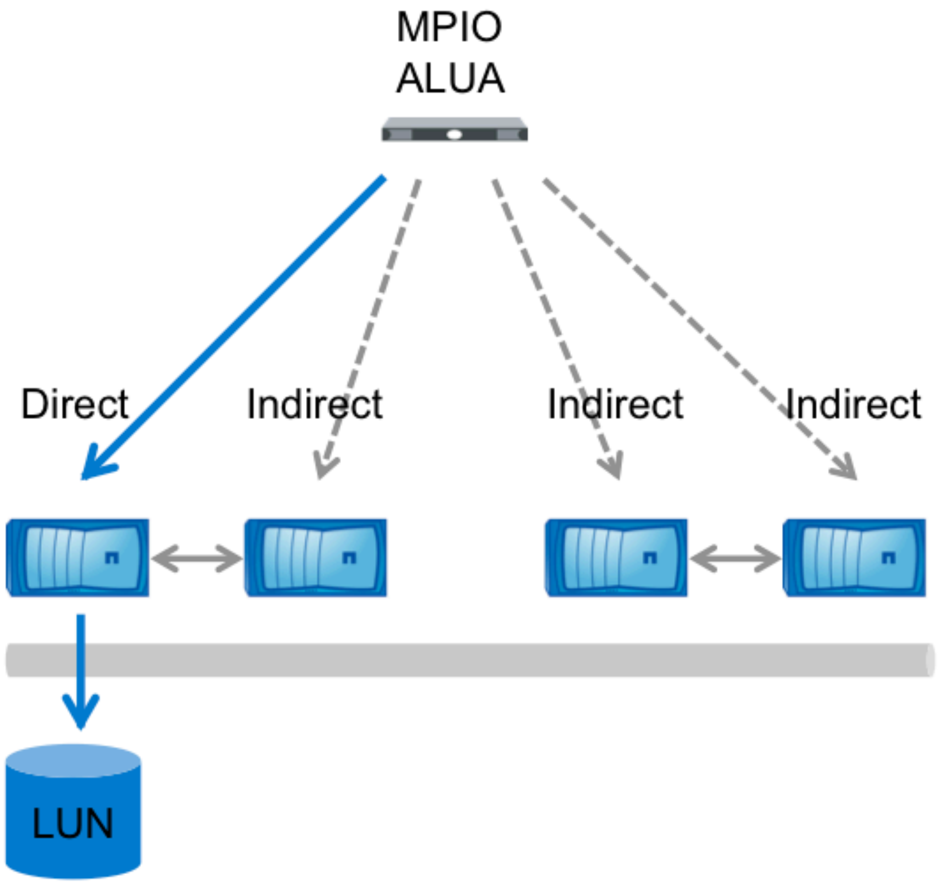
2.3 Host Connectivity

Hosts that access data served by clustered Data ONTAP using a block protocol are expected to make use of the asymmetrical logical unit access (ALUA) extension to the SCSI protocol to determine which paths are direct and which are indirect. The ALUA standard refers to direct paths as active/optimized and to indirect paths as active/nonoptimized. All ALUA information is requested and delivered in-band, using the same iSCSI or Fibre Channel connection that used for data.

The status of a given path is discoverable by a host that sends a path status inquiry down each of the paths it has discovered. This path status inquiry can be triggered when the storage system sends extra data along with the result of a SCSI request to inform a host that path statuses have been updated and that their priorities should be rediscovered.

ALUA is a well-known and widely deployed standard and is a requirement for access to data served by clustered Data ONTAP. Any operating systems tested and qualified to work with clustered Data ONTAP block access protocols will support ALUA.

Figure 1) ALUA and MPIO with direct and indirect paths.



2.4 Path Selection

Even though every LIF owned by an SVM accepts writes and read requests for its LUNs, only one of the cluster nodes actually owns the disks backing that LUN at any given moment. This effectively divides available paths to a LUN into two types: direct and indirect paths.

A *direct path* for a LUN is a path where an SVM's LIFs and the LUN being accessed reside on the same node. To go from a physical target port to disk, it is not necessary to traverse the cluster network.

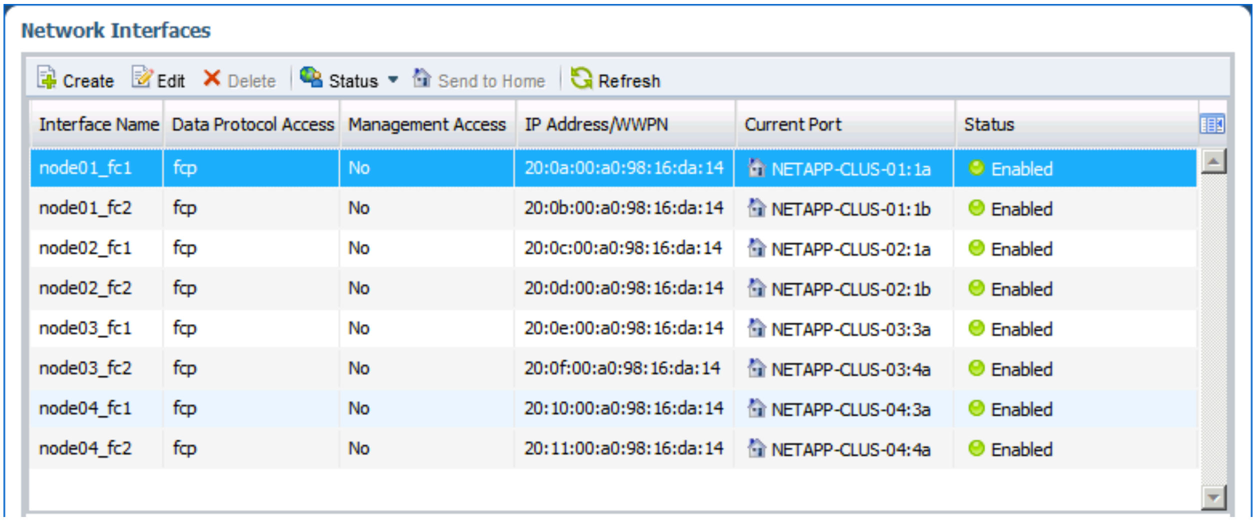
Indirect paths are data paths where an SVM's LIFs and the LUN being accessed reside on different nodes. Data must traverse the cluster network in order to go from a physical target port to disk. Because the cluster network is fast and highly available, this does not add a great deal of time to the round trip, but it is not the maximally efficient data path.

Because every host communicates only with SVMs that use physical resources anywhere in the cluster, in practice this means that all connections to a cluster are managed by multipath I/O (MPIO) software running on the host that is accessing LUNs, with the result that only direct paths are used during normal operation.

Best Practice

All SVMs should be assigned LIFs on each cluster node and each Fibre Channel fabric or Ethernet network. For instance, if a four-node cluster is connected to two independent Fibre Channel fabrics, A and B, using its 3a and 4a Fibre Channel target ports, an SVM that serves data by using Fibre Channel should have eight LIFs, on node1:3a, node1:4a, node2:3a, node2:4a, node3:3a, node3:4a, node4:3a, and node4:4a. Clusters with more than four nodes should limit the number of paths used to access any given LUN for ease of manageability and to respect operating system path count limitations (see section 2.9, "Portsets").

Figure 2) Block protocol LIFs in OnCommand System Manager 3.0.



Interface Name	Data Protocol Access	Management Access	IP Address/WWPN	Current Port	Status
node01_fc1	fcp	No	20:0a:00:a0:98:16:da:14	NETAPP-CLUS-01:1a	Enabled
node01_fc2	fcp	No	20:0b:00:a0:98:16:da:14	NETAPP-CLUS-01:1b	Enabled
node02_fc1	fcp	No	20:0c:00:a0:98:16:da:14	NETAPP-CLUS-02:1a	Enabled
node02_fc2	fcp	No	20:0d:00:a0:98:16:da:14	NETAPP-CLUS-02:1b	Enabled
node03_fc1	fcp	No	20:0e:00:a0:98:16:da:14	NETAPP-CLUS-03:3a	Enabled
node03_fc2	fcp	No	20:0f:00:a0:98:16:da:14	NETAPP-CLUS-03:4a	Enabled
node04_fc1	fcp	No	20:10:00:a0:98:16:da:14	NETAPP-CLUS-04:3a	Enabled
node04_fc2	fcp	No	20:11:00:a0:98:16:da:14	NETAPP-CLUS-04:4a	Enabled

For administrators who are used to using clustered Data ONTAP with NAS protocols such as NFS and CIFS, there is a distinction to be made between LIFs that serve these protocols and LIFs that serve block iSCSI or Fibre Channel. NAS LIFs can be freely moved from node to node, or they can belong to a failover group that specifies to which node and port they will move during an HA or port failover. SAN LIFs, by comparison, represent the endpoint of a number of paths, all established simultaneously between SCSI initiator and SCSI target, and the host's MPIO software manages which paths actually receive I/O.

Because of this difference in behavior, Ethernet LIFs that serve data by using the iSCSI protocol cannot also serve data by using a NAS protocol.

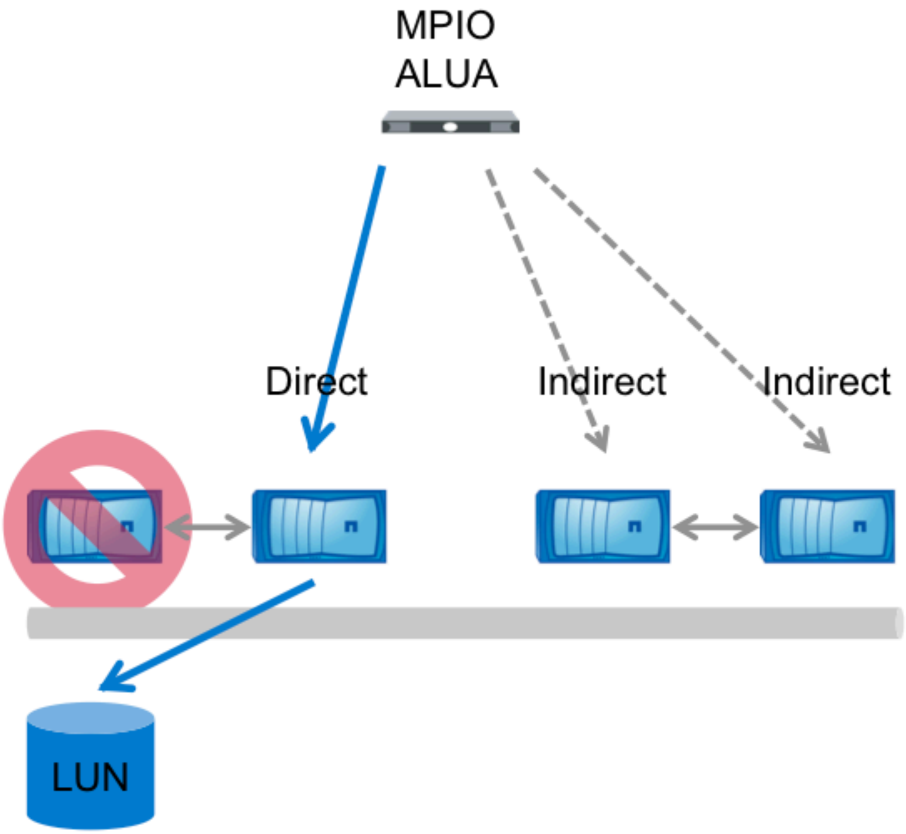
2.5 Path Selection Changes

There are three events that could change the path selected by a host to access data on a cluster.

HA Failover

In an HA failover event, LIFs on the down node are seen as offline, and LIFs on the HA partner that has taken over for the down node are now direct paths. This state changes automatically due to ALUA path inquiry, and no administrative changes are necessary.

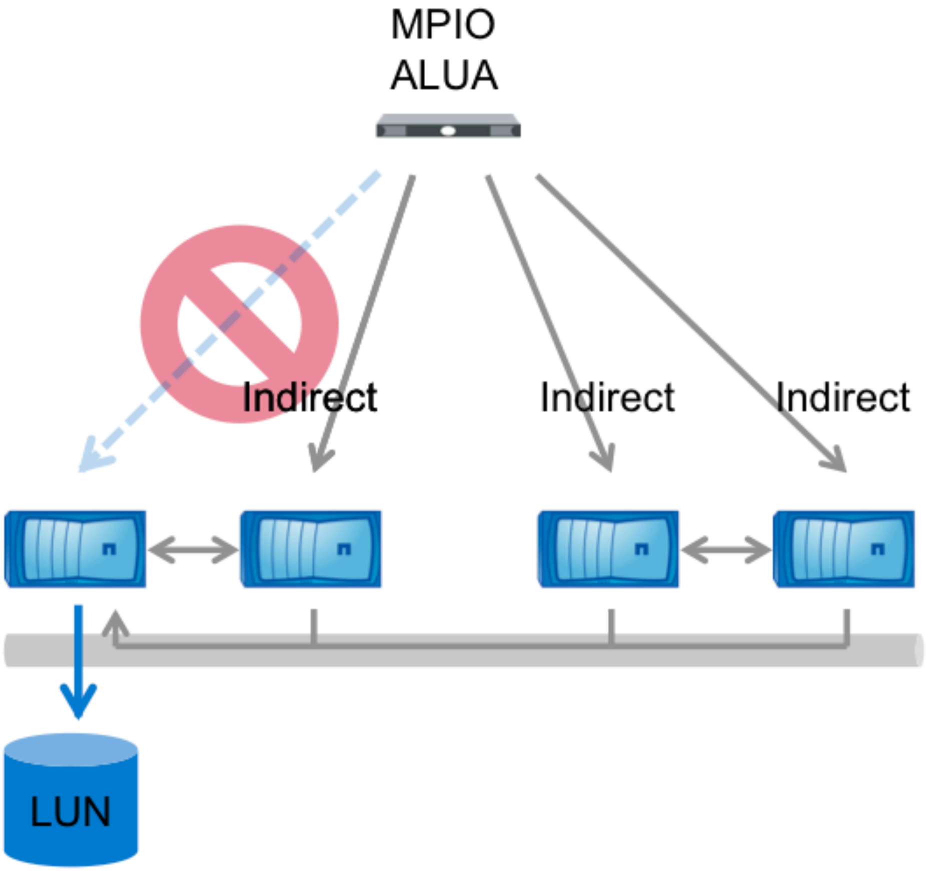
Figure 3) MPIO and ALUA path changes during an HA failover.



Port or Switch Failure

In a port or switch failure, no more direct paths are available. Path priority remains the same, and MPIO software running on the host selects alternate indirect paths until a direct path becomes available again. The ALUA path states do not change.

Figure 4) MPIO and ALUA path changes during a port or switch failure.



Volume Movement

A volume is moved transparently between nodes by using **`volume move`** functionality. When the cutover occurs and the volume's new node begins to handle read and write requests, the path status is updated so that the new node has direct paths and the old node has indirect paths. All paths remain available at all times.

Figure 5) MPIO and ALUA path changes during a volume move.

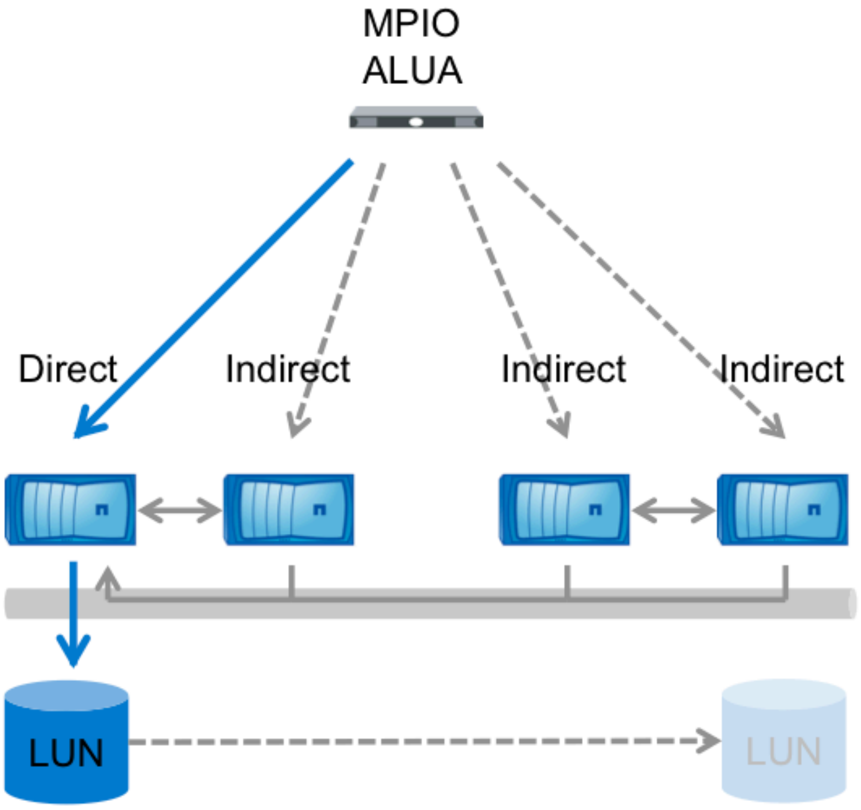
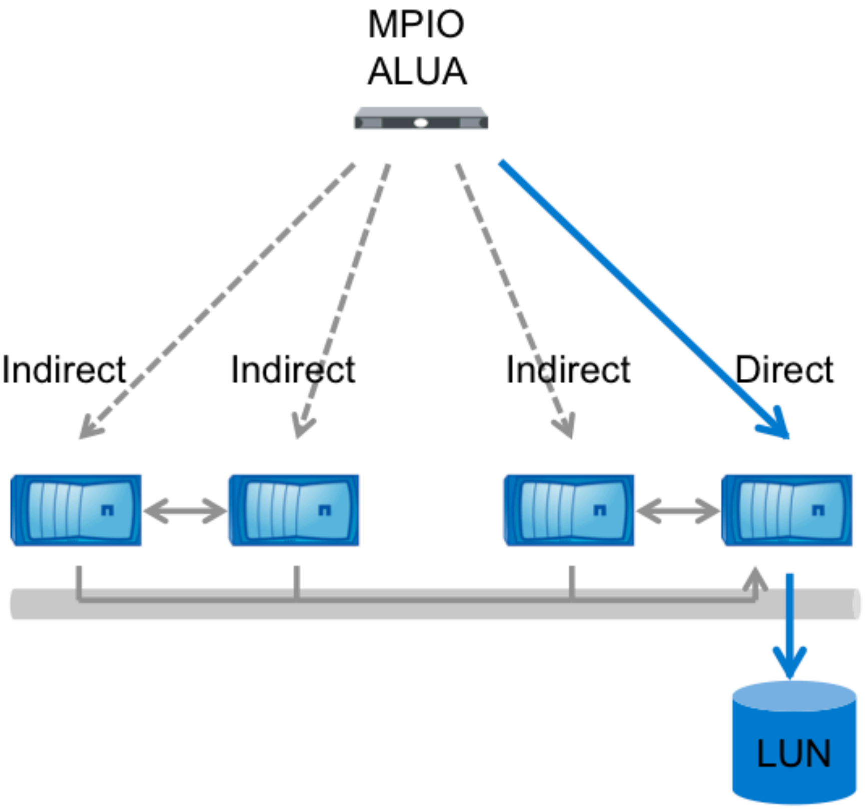


Figure 6) MPIO and ALUA path changes after a volume move.



2.6 Scalable SAN Configuration Differences from 7-Mode

Table 1 shows some key differences between clustered Data ONTAP and Data ONTAP operating in 7-Mode with respect to SAN configuration and behavior. The differences are covered in detail in the following sections.

Table 1) Configuration differences between clustered Data ONTAP and Data ONTAP operating in 7-Mode.

Configuration detail	Data ONTAP operating in 7-Mode	Clustered Data ONTAP
iSCSI secondary paths	Failover to partner interface	ALUA based
Fibre Channel indirect paths	Over NVRAM interconnect	Over cluster network
Fibre Channel ports	Physical ports	Virtual ports (NPIVs)
Portsets	Optional	Recommended
Interface with SnapDrive	Any active Ethernet port	SVM management LIF
FC and iSCSI service scope	Per-node/per-HA pair	Per SVM

2.7 iSCSI and the Partner Interface

Data ONTAP operating in 7-Mode uses a partner interface to provide redundancy for iSCSI connections. During an HA failover event, the HA partner of the affected storage controller has a predesignated partner interface that is brought online with the IP address of the taken-over HA partner’s iSCSI target interface, and I/O resumes.

In clustered Data ONTAP, iSCSI partner interfaces are not assigned when configuring iSCSI. Instead, iSCSI connections are made to the node where the data resides, to its HA partner for redundancy, and optionally to additional nodes in the same cluster. Instead of the path reappearing after an HA takeover on the HA partner, MPIO on the host determines that the primary path is no longer available and resumes I/O over the already-established connection to the taken over node’s HA partner. This means that, unlike in 7-Mode, iSCSI connections to a cluster have both direct and indirect paths to data managed by host MPIO software, operating in the same way as Fibre Channel paths.

2.8 Fibre Channel and NPIV

Clustered Data ONTAP uses N_Port ID Virtualization (NPIV) to permit every logical interface to log into an FC fabric with its own World Wide Port Name (WWPN), rather than using a single WWNN and associated WWPNs based on the address of the HA pair’s physical FC target adapters, as when operating in 7-Mode. This permits a host connected to the same FC fabric to communicate with the same SCSI target regardless of which physical node owns which LIF: The virtual port presents the SCSI target service and sends and receives data.

Best Practice

NPIV is required for Fibre Channel LIFs to operate correctly. Before creating FC LIFs, make sure that any fabrics attached to a clustered Data ONTAP system have NPIV enabled.

When using Cisco® NX-OS, the status of NPIV can be checked with the **`show npiv status`** command.

```
N5K-A# show npiv status
NPIV is enabled
```

When using Brocade FabOS, the **`portcfgshow`** command shows NPIV capability and status.

```
BRCD_8K:admin> portcfgshow
Ports of Slot 0    0   1   2   3     4   5   6   7     8   9  10  11    12  13  14  15
----------------+---+---+---+---+-----+---+---+---+-----+---+---+---+-----+---+---+---+---
Speed             AN  AN  AN  AN    AN  AN  AN  AN    10  10  10  10    10  10  10  10
Fill Word          0   0   0   0     0   0   0   0     -   -   -   -     -   -   -   -
AL_PA Offset 13   ..  ..  ..  ..    ..  ..  ..  ..    ..  ..  ..  ..    ..  ..  ..  ..
Trunk Port        ON  ON  ON  ON    ON  ON  ON  ON     -   -   -   -     -   -   -   -
Long Distance     ..  ..  ..  ..    ..  ..  ..  ..    ..  ..  ..  ..    ..  ..  ..  ..
VC Link Init      ..  ..  ..  ..    ..  ..  ..  ..    ..  ..  ..  ..    ..  ..  ..  ..
Locked L_Port     ..  ..  ..  ..    ..  ..  ..  ..    ..  ..  ..  ..    ..  ..  ..  ..
Locked G_Port     ..  ..  ..  ..    ..  ..  ..  ..    ..  ..  ..  ..    ..  ..  ..  ..
Disabled E_Port   ..  ..  ..  ..    ..  ..  ..  ..    ..  ..  ..  ..    ..  ..  ..  ..
Locked E_Port     ..  ..  ..  ..    ..  ..  ..  ..    ..  ..  ..  ..    ..  ..  ..  ..
ISL R_RDY Mode    ..  ..  ..  ..    ..  ..  ..  ..    ..  ..  ..  ..    ..  ..  ..  ..
RSCN Suppressed   ..  ..  ..  ..    ..  ..  ..  ..    ..  ..  ..  ..    ..  ..  ..  ..
Persistent Disable..  ..  ..  ..    ..  ..  ..  ..    ..  ..  ..  ..    ..  ..  ..  ..
LOS TOV enable    ..  ..  ..  ..    ..  ..  ..  ..    ..  ..  ..  ..    ..  ..  ..  ..
```

```
NPIV capability   ON  ON  ON  ON    ON  ON  ON  ON    ON  ON  ON  ON    ON  ON  ON  ON
NPIV PP Limit    126 126 126 126   126 126 126 126   126 126 126 126   126 126 126 126
QOS E_Port        AE  AE  AE  AE    AE  AE  AE  AE    ..  ..  ..  ..    ..  ..  ..  ..
EX Port           ..  ..  ..  ..    ..  ..  ..  ..    ..  ..  ..  ..    ..  ..  ..  ..
Mirror Port       ..  ..  ..  ..    ..  ..  ..  ..    ..  ..  ..  ..    ..  ..  ..  ..
Rate Limit        ..  ..  ..  ..    ..  ..  ..  ..    ..  ..  ..  ..    ..  ..  ..  ..
Fport Buffers     ..  ..  ..  ..    ..  ..  ..  ..    ..  ..  ..  ..    ..  ..  ..  ..
Port Auto Disable ..  ..  ..  ..    ..  ..  ..  ..    ..  ..  ..  ..    ..  ..  ..  ..
CSCTL mode        ..  ..  ..  ..    ..  ..  ..  ..    ..  ..  ..  ..    ..  ..  ..  ..
```

From the storage administration console, it is not possible to inquire about NPIV status on an attached switch directly, but examining the local FC topology can show whether fabric switch ports have NPIV enabled. In the following example, NPIV must be enabled, because port 2/1 has multiple attached WWPNs, some of which are virtual ports.

```
cluster::> node run -node node01 fcp topology show
   Switch Name: N5K-A
 Switch Vendor: Cisco Systems, Inc.
Switch Release: 5.0(2)N2(1a)
 Switch Domain: 200
    Switch WWN: 20:66:00:0d:ec:b4:94:01

Port     Port WWPN                State    Type     Attached WWPN           Port ID
-----------------------------------------------------------------------------------
2/1      20:01:00:0d:ec:b4:94:3f  Online   F-Port   50:0a:09:83:8d:4d:bf:f1 0xc80033
                                                    20:1c:00:a0:98:16:54:8c 0xc80052*
                                                    20:0e:00:a0:98:16:54:8c 0xc80034*
                                                    20:10:00:a0:98:16:54:8c 0xc8003f
2/3      20:03:00:0d:ec:b4:94:3f  Online   F-Port   50:0a:09:83:8d:3d:c0:1c 0xc8002c
```

Best Practice

Physical WWPNs (beginning with 50:0a:09:8x) do not present a SCSI target service and should not be included in any zone configurations on the FC fabric. Use only virtual WWPNs (visible in the output of the **`network interface show`** command and in the System Manager Cluster → Vserver → Configuration → Network Interface pane, as shown in Figure 2.

2.9 Portsets

Clusters with more than two nodes are likely to have more paths than has commonly been the case in the past. Clusters attached to more than one fabric, or with nodes attached more than once per fabric, can quickly multiply the number of potential paths available.

This presents the following potential problems to the storage administrator:

- Having a large number of target ports can be good for redundancy, but it can become operationally burdensome. In an FC environment, it requires larger, more complex zones and zonesets; a larger table of WWPNs belonging to cluster SVMs to keep track of; or, in the case of an iSCSI environment, a large number of sessions to be established for every host that requires a LUN.
- Many operating systems have an upper limit to the number of paths it is feasible for them to access. Especially for hosts that have many paths and many LUNs, this can lead to LUN enumeration or path status problems.
- Some demanding, high-throughput workloads can benefit from having their traffic segregated from less critical traffic to reduce contention, but ALUA path statuses provide no mechanism to prioritize one direct path over another.

Portsets permit administrators to mask an interface group (igroup) so that the LUNs that are mapped to it are available on a subset of the total number of available target ports. This functionality is available in both clustered Data ONTAP and 7-Mode. Portsets are much more likely to be useful in clustered Data ONTAP, because higher potential path counts are supported.

To ensure both a direct path to data and availability/redundancy in the case of an HA failover or nondisruptive operation event, the only paths required are to the node that contains the volume with the data being accessed and its HA partner.

Note: A LIF that is currently a member of a portset cannot be modified until it is removed from the portset. It can be added to the portset after modification, but care should be taken to leave enough LIFs in the portset to satisfy host requirements for a path to data.

Best Practice

Portsets should be used in clustered Data ONTAP to prune the total number of available paths to a number sufficient for redundancy and direct data access without presenting unnecessary paths. Present no more than eight paths per LUN for a given igroup. It is not necessary to have a cluster of more than four nodes in order to use portsets. It is a best practice to use them on clusters with a smaller total node count so that as nodes are added, they are already in place.

There are two non–mutually-exclusive methods to consider:

- For clusters with more than one interface per Ethernet network or fabric, use portsets to limit hosts to a subset of the total number of ports per node. For example, if FC ports 0a and 0c on each node are attached to fabric A, and ports 0b and 0d are attached to fabric B, use a portset to limit every host to either a portset that contains ports 0a and 0b or a portset that contains ports 0c and 0d.
- Use portsets to limit a host's access to a subset of the total number of available nodes. Care should be taken to make sure that hosts have access to both members of an HA pair so that in case of a takeover, the remaining node presents a direct path to its data. Additionally, make sure that volumes are not moved to a node that will not present a direct path after the move is complete. (This would not disrupt access to data, but would force data over an indirect path.)

2.10 Management Interfaces

Because LIFs belonging to SVMs that serve data by using block protocols cannot also be used for administration purposes, and because the logical unit of management in clustered Data ONTAP is the SVM, every SVM must have a management interface in addition to interfaces that are serving data using block protocols.

This interface should have the following properties:

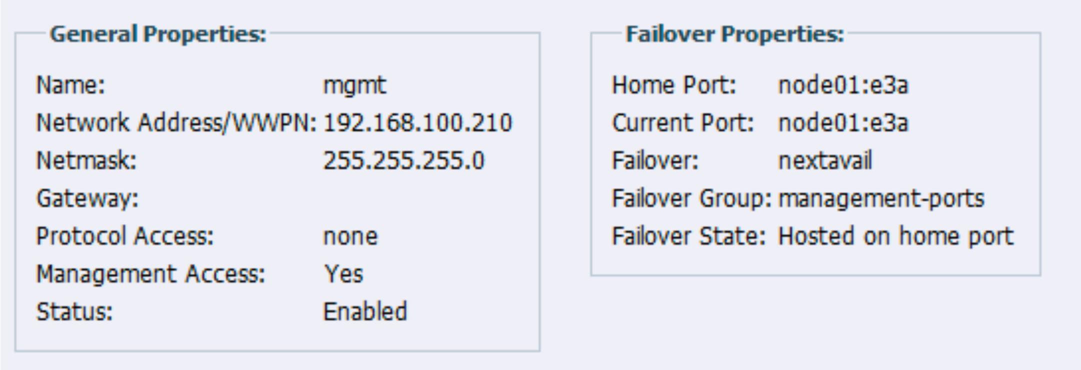
- A LIF type of `data`
- No data protocols assigned (`-data-protocols none`)
- A firewall policy that permits management access (`-firewall-policy mgmt`)

Example:

```
cluster::> network interface show -lif mgmt -instance

                    Vserver Name: vs
               Logical Interface: mgmt
                            Role: data
                   Data Protocol: none
                       Home Node: node01
                       Home Port: e3a
                    Current Node: node02
                    Current Port: e3a
              Operational Status: up
                 Extended Status: -
                         Is Home: false
                 Network Address: 192.168.100.210
                         Netmask: 255.255.255.0
                 IPv4 Link Local: -
             Bits in the Netmask: 24
              Routing Group Name: d192.168.100.0/24
           Administrative Status: up
                 Failover Policy: nextavail
                 Firewall Policy: mgmt
                     Auto Revert: false
              Use Failover Group: management-ports
                        DNS Zone: none
             Failover Group Name:
                        FCP WWPN: -
```

Figure 7) Detail pane of a management LIF in OnCommand System Manager.



General Properties:

Name: mgmt
Network Address/WWPN: 192.168.100.210
Netmask: 255.255.255.0
Gateway:
Protocol Access: none
Management Access: Yes
Status: Enabled

Failover Properties:

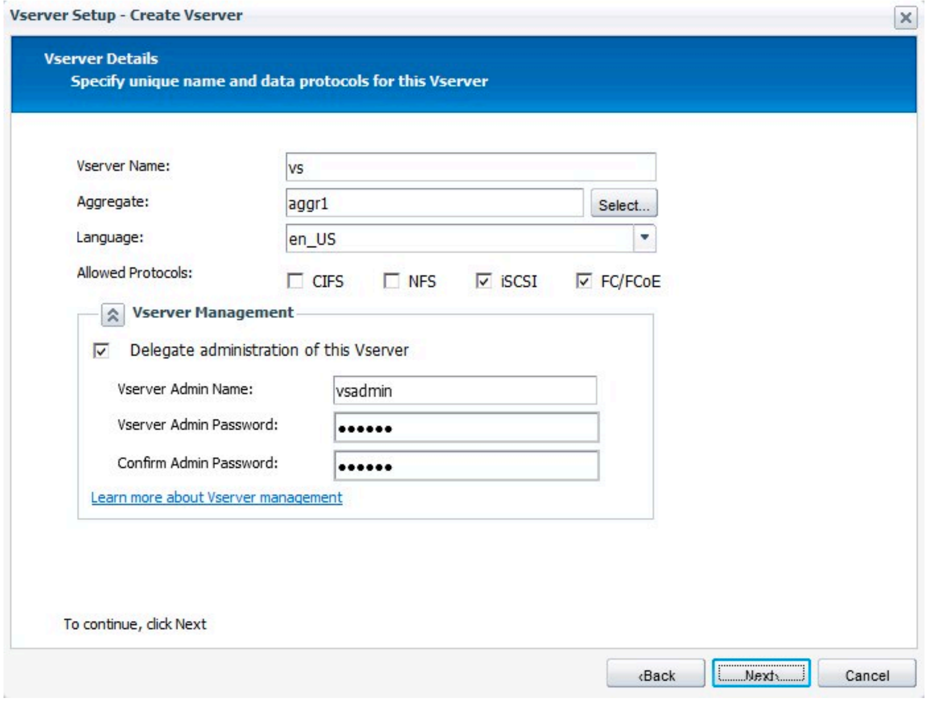
Home Port: node01:e3a
Current Port: node01:e3a
Failover: nextavail
Failover Group: management-ports
Failover State: Hosted on home port

Best Practice

The management LIF should be assigned to a failover group that makes it accessible to hosts that might need to contact it for data management purposes, such as creating or managing Snapshot™ copies by using NetApp SnapDrive®. For more information on failover groups, see "Configuring Failover Groups for LIFs" in the Clustered Data ONTAP Network Management Guide.

Additionally, an SVM-level administrative account should be available. The default account created during SVM creation is the `vsadmin` account, but it must first be assigned a password with the **`security login password`** command and then unlocked by using the **`security login unlock`** command. For more details, see "Delegating Administration to a Vserver Administrator" in the Clustered Data ONTAP System Administration Guide.

Note: When using System Manager 2 to configure SVMs, a management interface and a management user can be created as part of the Vserver Setup Wizard process by selecting the “Delegate administration of this Vserver” checkbox.



Vserver Setup - Create Vserver

Vserver Details

Specify unique name and data protocols for this Vserver

Vserver Name: vs

Aggregate: aggr1

Language: en_US

Allowed Protocols: □ CIFS □ NFS ☑ iSCSI ☑ FC/FCoE

Vserver Management

☑ Delegate administration of this Vserver

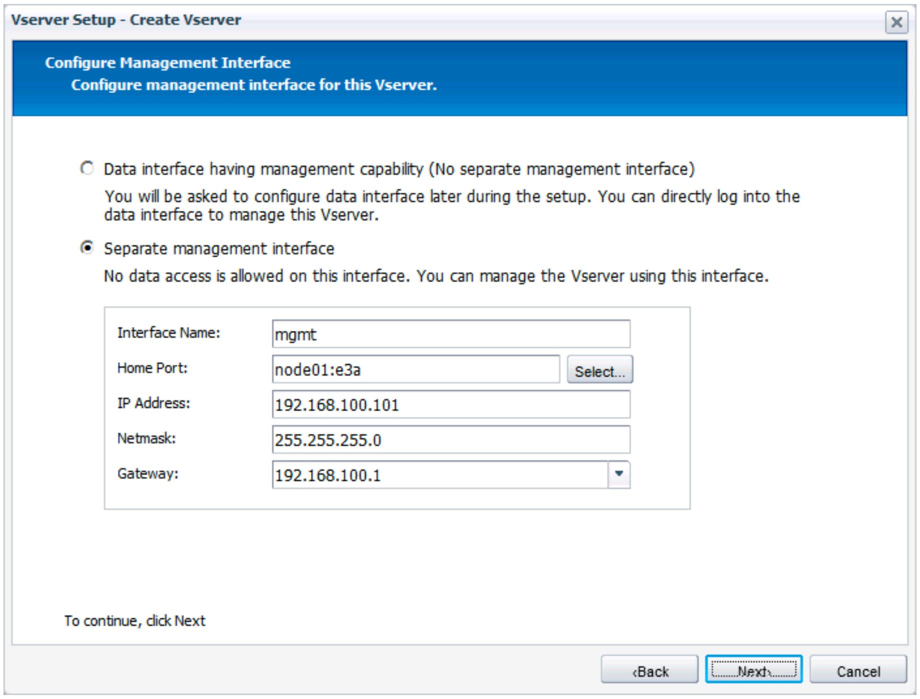
Vserver Admin Name: vsadmin

Vserver Admin Password: ••••••

Confirm Admin Password: ••••••

Learn more about Vserver management

To continue, click Next



Vserver Setup - Create Vserver

Configure Management Interface

Configure management interface for this Vserver.

○ Data interface having management capability (No separate management interface)

You will be asked to configure data interface later during the setup. You can directly log into the data interface to manage this Vserver.

● Separate management interface

No data access is allowed on this interface. You can manage the Vserver using this interface.

Interface Name: mgmt

Home Port: node01:e3a

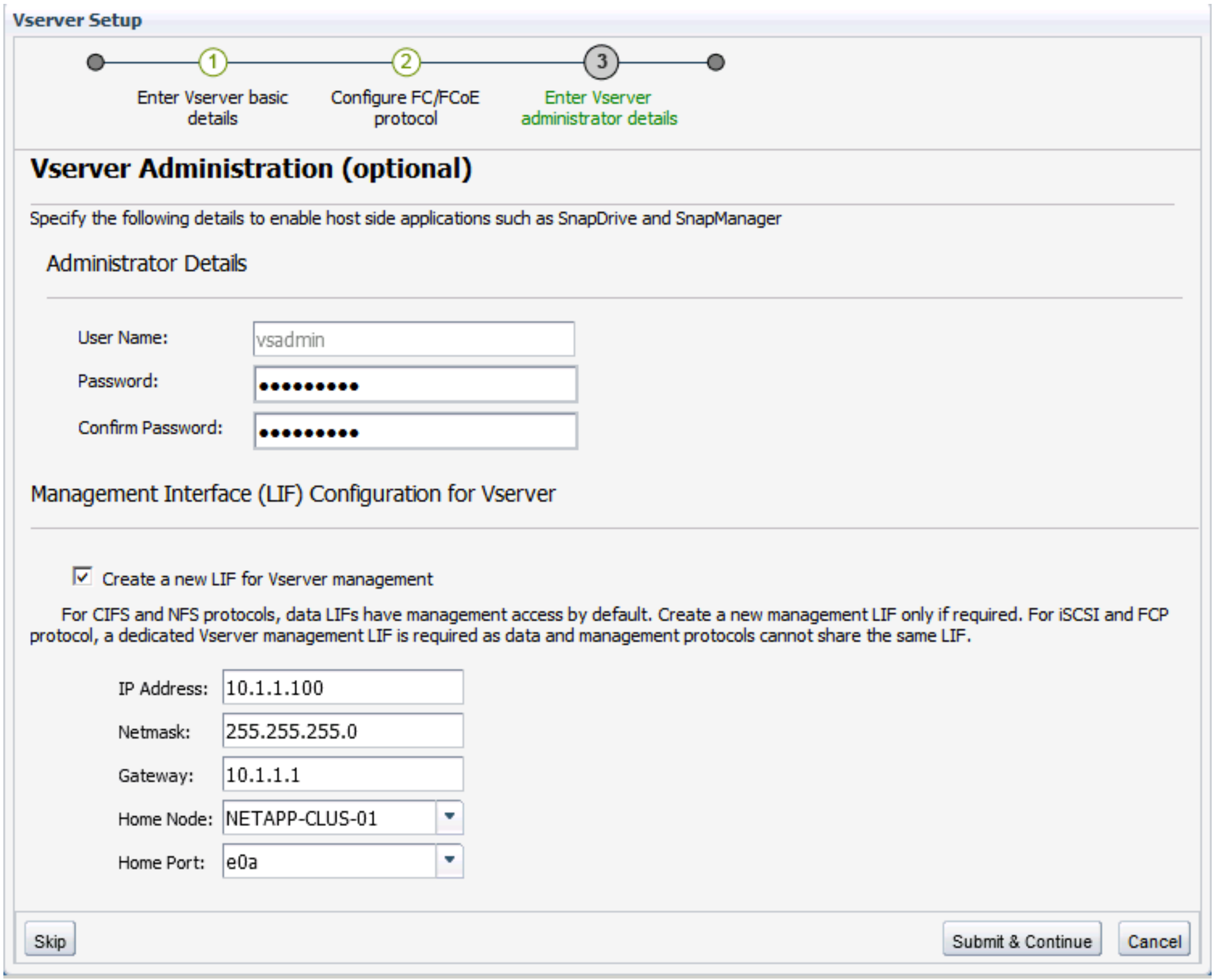
IP Address: 192.168.100.101

Netmask: 255.255.255.0

Gateway: 192.168.100.1

To continue, click Next

In System Manager 3, setting up an SVM administrator and creating a management interface is automatically part of the Vserver Setup Wizard workflow.


Vserver Setup
1 Enter Vserver basic details
2 Configure FC/FCoE protocol
3 Enter Vserver administrator details

Vserver Administration (optional)

Specify the following details to enable host side applications such as SnapDrive and SnapManager

Administrator Details

User Name: vsadmin

Password: ●●●●●●●●●●

Confirm Password: ●●●●●●●●●●

Management Interface (LIF) Configuration for Vserver

☑ Create a new LIF for Vserver management

For CIFS and NFS protocols, data LIFs have management access by default. Create a new management LIF only if required. For iSCSI and FCP protocol, a dedicated Vserver management LIF is required as data and management protocols cannot share the same LIF.

IP Address: 10.1.1.100

Netmask: 255.255.255.0

Gateway: 10.1.1.1

Home Node: NETAPP-CLUS-01

Home Port: e0a

3 Scalable SAN Key Value Propositions

This section highlights some of NetApp’s principal design goals for the clustered Data ONTAP architecture. These goals included providing a unified architecture at scale that enables nondisruptive operations for data mobility, performance optimization, capacity planning, and even system-level hardware replacement. Although this is not an exhaustive list of key features now available, it does show how scalable SAN and clustered Data ONTAP are set apart from the rest of the storage market.

3.1 SVM as Unified Target and Unit of Management

Data ONTAP operating in 7-Mode, when running as a member of an HA configuration, already presents a single WWNN to an attached Fibre Channel fabric. Clustered Data ONTAP extends this single WWNN on an SVM basis to every member of a cluster, so that every node presents the same target, and also permits multiple targets to coexist on the same physical hardware.

The same concept also applies to storage management. Since all data is served from volumes associated with an SVM and from an iSCSI or FC target configured as part of an SVM, a cluster is administered on a per-SVM basis, rather than the time-consuming process of administering storage a single node at a time.

This focus on management at the SVM level means that it is possible to implement a secure multi-tenancy model of storage management

3.2 Cluster-Wide Consistency Groups

Beginning with clustered Data ONTAP 8.2, Snapshot consistency groups are supported. When running in 7-Mode, consistency groups are a way for Snapshot copies on multiple storage controllers to be taken simultaneously, allowing a host with LUNs served from volumes on multiple storage controllers to ensure consistency across all of them.

Clustered Data ONTAP translates consistency groups to a cluster-oriented format. Rather than directing a Snapshot copy to be taken on multiple storage controllers at once, a host can take a copy across multiple cluster nodes and volumes simultaneously with a single command. Consistency groups in clustered Data ONTAP work on a per-SVM basis, so any volumes owned by the SVM that is receiving the command are candidates for a Snapshot copy.

3.3 Migrating Resources Between Cluster Nodes

During normal operations there is no need for LIFs or volumes to move from one cluster node to another, but in some circumstances nondisruptive migration of either volumes or LIFs from one node to another might be desirable; for example, replacing one HA pair of a cluster with an upgraded HA pair in a way that is transparent to hosts accessing LUN data.

Migrating volumes from one node to another requires only that the destination node be able to provide a direct path for the host (see “Volume Movement" in section 2.5).

Migrating a LIF from one node and port to another can be made less administratively burdensome by modifying rather than deleting and recreating it; its IP address or WWPN remains the same, so no fabric zoning or host changes are needed. SAN LIFs can be modified only when the LIF (but not the port) in question is offline. SAN LIFs can be set administratively offline by using the **`network interface modify -status-admin down`** command.

Best Practice
Do not exceed the cluster size limit when making changes to cluster membership. For information on the cluster size limit when using block protocols, refer to the Clustered Data ONTAP SAN Configuration Guide.

4 Host Integration Examples

4.1 NetApp Host Utilities Kit

Installation of the Host Utilities Kit sets timeout and other operating-system-specific values to their recommended defaults, and includes utilities for examining LUNs provided by NetApp storage, whether clustered or operating in 7-Mode. The Host Utilities Kit became clustered Data ONTAP aware for the operating system with the version specified in Figure 8. See the NetApp Interoperability Matrix Tool for complete details for a given NetApp tested and supported configuration.

Figure 8) Scalable SAN support by operating system and NetApp Host Utilities Kit version.

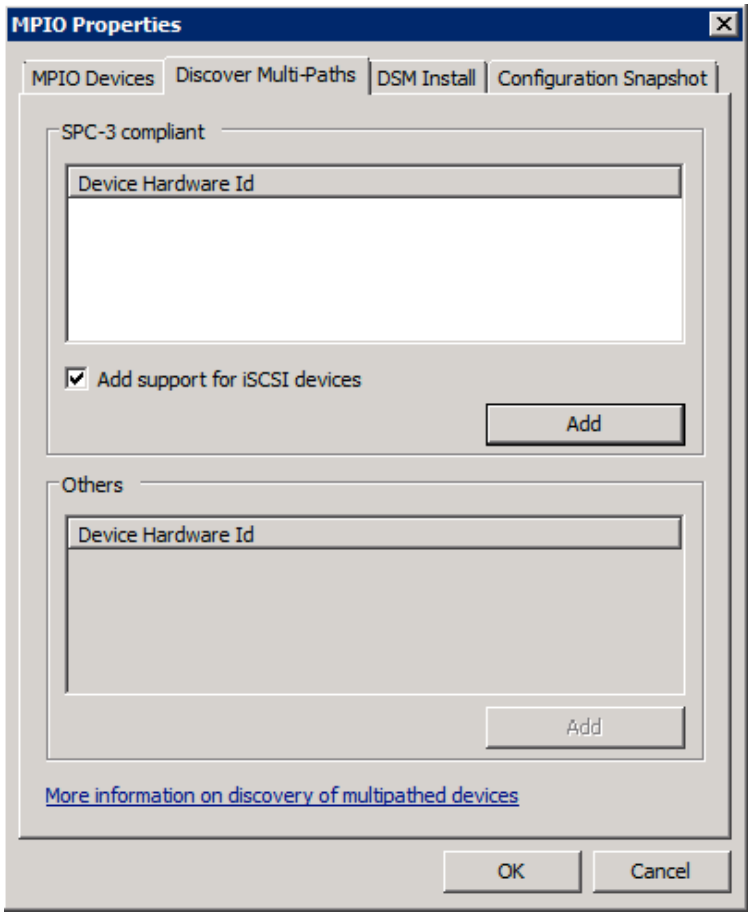
Operating System	Version	Host Utilities Kit Version
IBM AIX	6.1, 7.1	6.0
Solaris	10u8, 11	6.1
Red Hat Enterprise Linux	5.8, 6.1	6.0
HP-UX	11i v3	6.0
Microsoft Windows	2003, 2008, 2011	6.0

4.2 Microsoft Windows and Clustered Data ONTAP

Microsoft Windows and Native MPIO

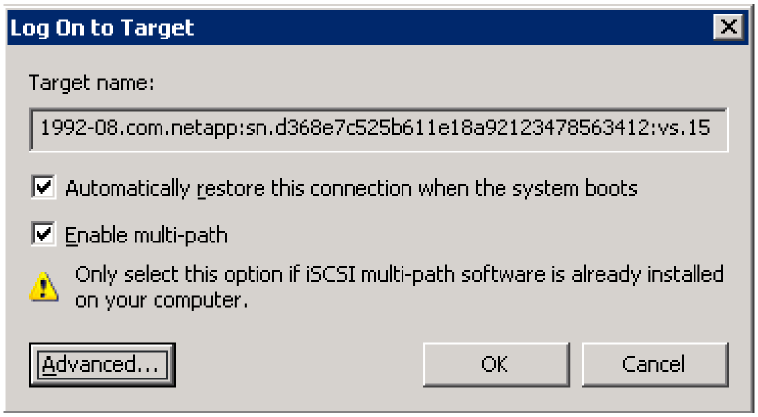
To operate as intended, clustered Data ONTAP requires MPIO and ALUA. In the case of Microsoft® Windows® 2008 and Windows 2012, these are natively supported whenever the Multipath I/O feature is installed.

When using the iSCSI protocol, it’s necessary to tell Windows to apply multipathing support to iSCSI devices in the MPIO Properties administrative application: Navigate to the Discover Multi-Paths pane, select the “Add support for iSCSI devices” checkbox, and click Add.



☑ Add support for iSCSI devices

It’s also necessary to create multiple sessions from the host initiators to the target iSCSI LIFs on the clustered Data ONTAP system. This can be accomplished using the native iSCSI initiator: Select the “Enable multi-path” checkbox when logging on to a target.



Target name:

1992-08.com.netapp:sn.d368e7c525b611e18a92123478563412:vs.15

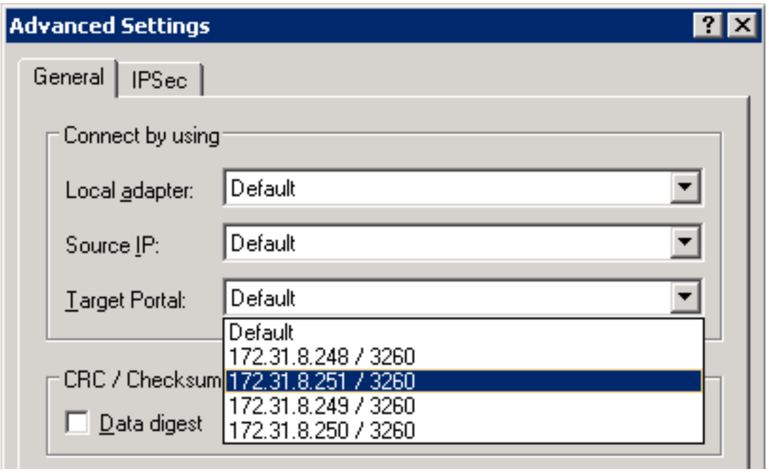
☑ Automatically restore this connection when the system boots

☑ Enable multi-path

Only select this option if iSCSI multi-path software is already installed on your computer.

To manually create additional sessions, highlight the corresponding target in the Targets pane of the iSCSI initiator and click Log on. Make sure that the session is automatically restored after the next reboot and that the new session is identified as a multipath session by selecting both checkboxes.

Click Advanced. From the Target Portal drop-down list, select the IP address of the logical interface that will be the target of the new iSCSI session.



Connect by using

Local adapter: Default

Source IP: Default

Target Portal: Default

Default
172.31.8.248 / 3260
172.31.8.251 / 3260
172.31.8.249 / 3260
172.31.8.250 / 3260

CRC / Checksum

☐ Data digest

Sessions can also be managed by using the NetApp SnapDrive iSCSI management pane. This is the preferred method, because SnapDrive remembers which target logical interfaces already have an established session and preselects an unused target portal.

4.3 Data ONTAP DSM

The Data ONTAP DSM supports attaching to a Data ONTAP cluster beginning with version 3.5. Consult the NetApp Interoperability Matrix Tool for current information on supported configurations.

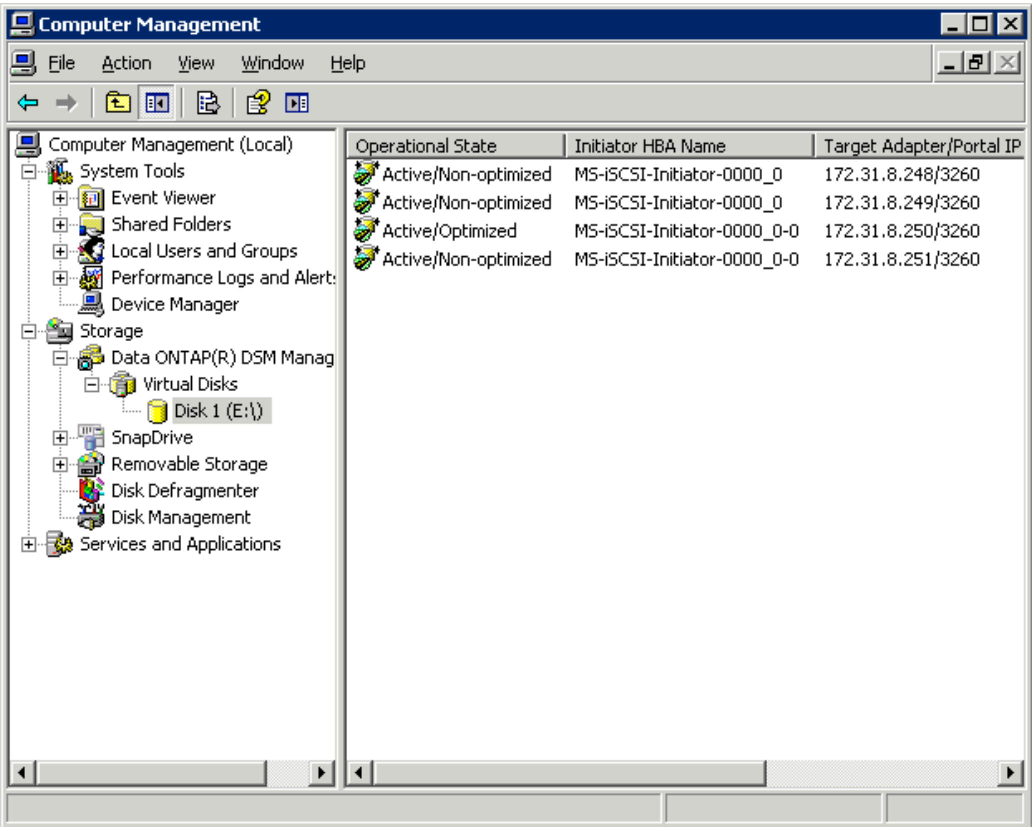
Because Windows 2003 does not include native MPIO and ALUA capability, the Data ONTAP DSM is required when using clustered Data ONTAP storage.

Although Windows 2008 and Windows 2012 support MPIO and ALUA natively, the Data ONTAP DSM can still be used; it take priority over the native MPIO implementation when accessing NetApp LUNs.

Both native MPIO and Data ONTAP DSM discover which paths are direct and indirect and route traffic appropriately, but the Data ONTAP DSM has the advantage of including a GUI in the Microsoft Management Console that correctly displays logical interface and SVM names, making management simpler and more intuitive.

During installation, the Data ONTAP DSM sets a number of Windows registry values to optimize performance and provide correct behavior during the failover scenarios covered in section 2.5, “Path Selection Changes.” For a complete list of registry values changed by the Data ONTAP DSM, refer to "Timeout and Turning Parameters Overview" in the Data ONTAP DSM Installation and Administration Guide.

Figure 9) DSM Manager pane.



Operational State	Initiator HBA Name	Target Adapter/Portal IP
Active/Non-optimized	MS-iSCSI-Initiator-0000_0	172.31.8.248/3260
Active/Non-optimized	MS-iSCSI-Initiator-0000_0	172.31.8.249/3260
Active/Optimized	MS-iSCSI-Initiator-0000_0-0	172.31.8.250/3260
Active/Non-optimized	MS-iSCSI-Initiator-0000_0-0	172.31.8.251/3260

Host Utilities Kit

The NetApp Host Utilities Kit can also be installed. As with the Data ONTAP DSM, the appropriate values are changed in the Windows registry to optimize performance and provide correct operation during failover events. However, if the Data ONTAP DSM is already installed, the Host Utilities Kit does not change the Windows registry, instead relying on the Data ONTAP DSM to make sure that the correct values are set.

4.3.1 NetApp SnapDrive

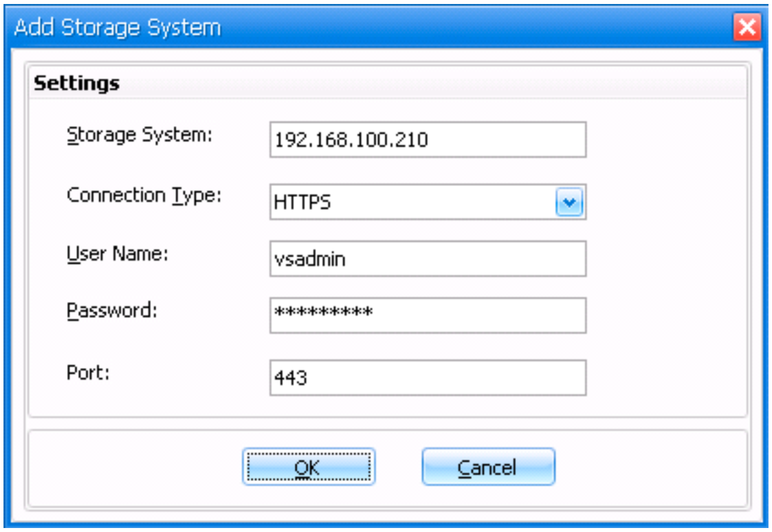
NetApp SnapDrive for Windows performs the same essential functions for clustered Data ONTAP as it does for Data ONTAP operating in 7-Mode:

- Crash-consistent LUN Snapshot copies provided by integration with Windows Volume Shadow Copy Services
- Management of iSCSI sessions by using the iSCSI management pane (highly useful now that multiple sessions are required in order to allow selection of direct paths during failover and volume move scenarios)
- On-the-fly LUN and resizing and cloning

Note: NetApp SnapDrive can be used in conjunction with clustered Data ONTAP beginning with version 6.4.

However, there is an important difference: The unit of management when using NetApp SnapDrive with clustered Data ONTAP is at the level of individual SVMs, not at the node or cluster level. As virtual storage servers, SVMs are meant to provide a secure multi-tenancy environment in which specific volumes, network interfaces, target ports, and management users are logically grouped. Therefore, a host connected to multiple SVMs that are part of the same cluster, accessing the same physical disks and network interfaces, would not have visibility into the fact that they are logical entities distributed across physically separate but interconnected nodes.

When first accessing an SVM by using NetApp SnapDrive for Windows, you must access the Transport Protocol settings under the server list in the SnapDrive management pane. From there, a management logical interface and management user as described in section 2.10, “Management Interfaces,” can be configured to enable access to the resources on a given SVM.



Storage System: 192.168.100.210

Connection Type: HTTPS

User Name: vsadmin

Password: *********

Port: 443

4.4 UNIX or Linux and Clustered Data ONTAP

Host Utilities Kit

The NetApp Host Utilities Kit contains utilities that are useful for viewing LUN configuration at the SVM level. Using the Host Utilities Kit can provide extended information about the SVM to which an attached LUN belongs, in addition to its volume and pathname in the SVM context.

```
# sanlun lun show all
controller(7mode)/              device       host                   lun
vserver(Cmode) lun-pathname     filename     adapter    protocol    size    mode
-------------------------------------------------------------------------------
vs             /vol/vol1/linux1 /dev/sdcx    host1      FCP         25g     C
vs             /vol/vol2/linux2 /dev/sdcw    host1      FCP         25g     C
vs             /vol/vol3/linux3 /dev/sdck    host1      FCP         25g     C
```

Additionally, the Host Utilities Kit can be used to display which of an SVM’s logical interfaces are providing the direct and indirect paths for a given LUN, labeled here as primary for direct paths and secondary for indirect paths.

```
# sanlun lun show -p
                    ONTAP Path: vs:/vol/vol1/linux1
                           LUN: 0
                      LUN Size: 25g
                          Mode: C
                   Host Device: 3600a09803246664c422b2d51674f7470
              Multipath Policy: round-robin 0
            Multipath Provider: Native
--------- ---------- ------- ------------ -------------------------------------
host      vserver
path      path       /dev/   host         vserver
state     type       node    adapter      LIF
--------- ---------- ------- ------------ -------------------------------------
up        primary    sdfo    host0        fcoe_lif_1
up        primary    sdfk    host1        fcoe_lif_2
up        secondary  sdga    host0        fcoe_lif_3
up        secondary  sdge    host1        fcoe_lif_4
up        secondary  sdgm    host1        fcoe_lif_5
up        secondary  sdgj    host0        fcoe_lif_6
up        secondary  sdfw    host0        fcoe_lif_7
up        secondary  sdgq    host1        fcoe_lif_8
```

As with Data ONTAP operating in 7-Mode, NetApp recommends using the values in the Recommended Host Settings for Linux Host Utilities guide when configuring the host to access LUNs on a clustered Data ONTAP cluster.

NetApp SnapDrive

As with Microsoft Windows, the unit of management when using SnapDrive and clustered Data ONTAP is at the individual SVM rather than at the node or cluster level. The **snapdrive config set** command must be used in conjunction with a management logical interface and an SVM administrative user, as described in section 2.10, “Management Interfaces,” in order to administer LUNs attached by using iSCSI or Fibre Channel from an attached host.

```
# snapdrive config set vsadmin vs
Password for vsadmin:
Retype password:

# snapdrive config list
username    appliance name   appliance type
-------------------------------------------
vsadmin     vs               StorageSystem
```

4.5 IBM AIX and Clustered Data ONTAP

IBM AIX using Fibre Channel to access data on a cluster is supported beginning with clustered Data ONTAP 8.2. For the supported AIX technology levels and service packs, refer to the NetApp Interoperability Matrix.

The iSCSI protocol is not supported with AIX and clustered Data ONTAP.

Version History

Version	Date	Document Version History
Version 1.0	June 2012	Initial release. Covers Data ONTAP 8.1.
Version 2.0	June 2013	Second release. Covers Data ONTAP 8.2.

References

Note: All linked product documentation is for versions compatible with clustered Data ONTAP 8.2.

- Interoperability Matrix Tool
  http://support.netapp.com/matrix/
- Clustered Data ONTAP SAN Administration Guide
  https://library.netapp.com/ecm/ecm_download_file/ECMP1155469
- Clustered Data ONTAP SAN Configuration Guide
  https://library.netapp.com/ecm/ecm_download_file/ECMP1155472
- Data ONTAP DSM for Windows MPIO Documentation
  http://support.netapp.com/documentation/docweb/index.html?productID=61500
- Host Utilities for AIX
  http://support.netapp.com/documentation/docweb/index.html?productID=61370&platformID=30455
- Host Utilities for HP-UX
  http://support.netapp.com/documentation/docweb/index.html?productID=61370&platformID=30455
- Host Utilities for Solaris
  http://support.netapp.com/documentation/docweb/index.html?productID=61477&platformID=30488
- Host Utilities for Linux
  http://support.netapp.com/documentation/docweb/index.html?productID=61477&platformID=30480
- Host Utilities for Windows
  http://support.netapp.com/documentation/docweb/index.html?productID=61608&platformID=30462
- SnapDrive for Windows
  http://support.netapp.com/documentation/productlibrary/index.html?productID=30049
- SnapDrive for UNIX
  http://support.netapp.com/documentation/productlibrary/index.html?productID=30050
- TR-4017: Fibre Channel SAN Best Practices
  www.netapp.com/us/library/technical-reports/tr-4017.html

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.

Go further, faster®

NetApp®

www.netapp.com