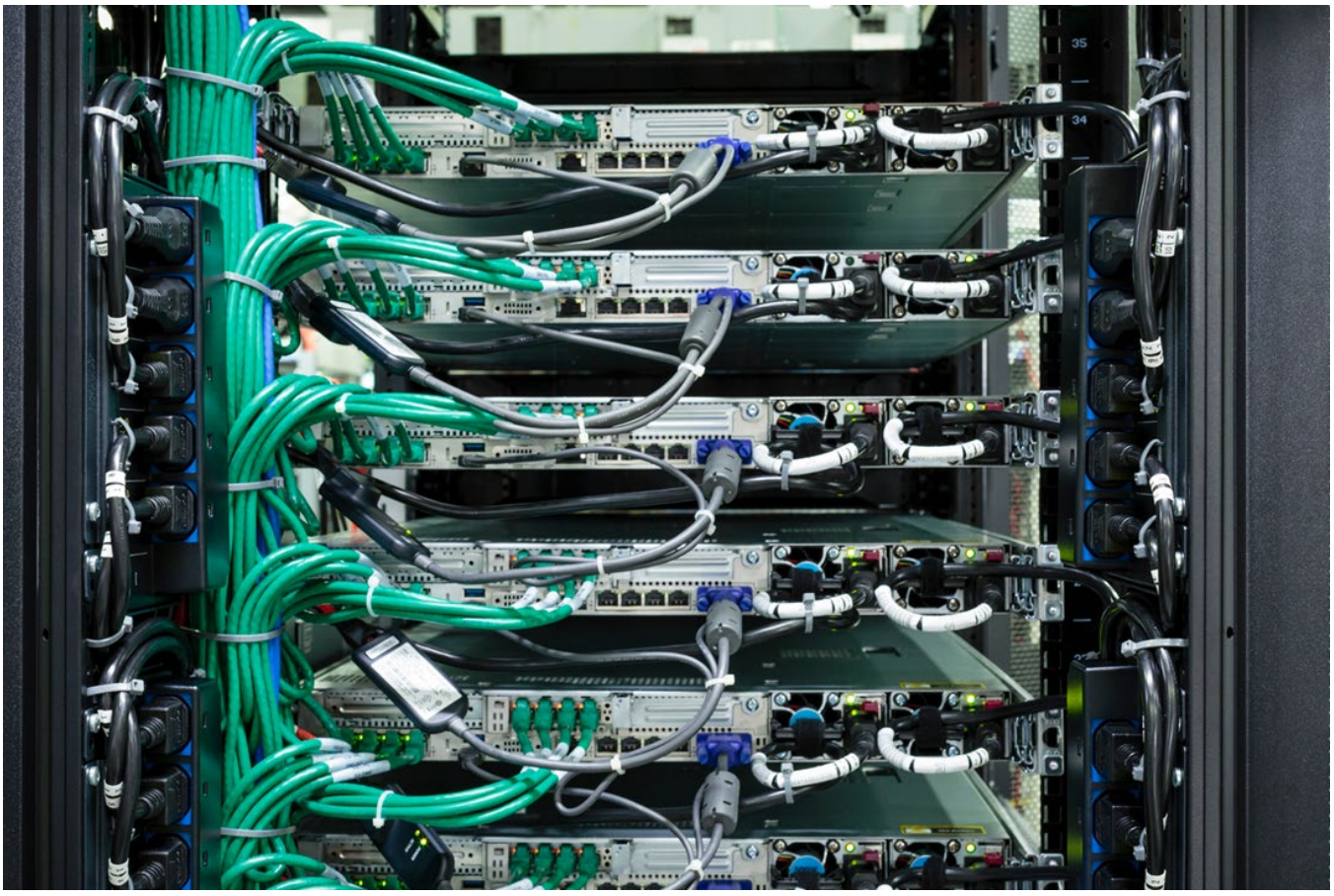
Hewlett Packard Enterprise

Technical white paper

Check if the document is available in the language of your choice.

HPE NIMBLE STORAGE DHCI SOLUTION NETWORK CONSIDERATIONS GUIDE



CONTENTS

EXECUTIVE SUMMARY

The HPE Nimble Storage dHCI solution from Hewlett Packard Enterprise is a disaggregated hyperconverged infrastructure (HCI) platform that delivers the flexibility of converged infrastructure and the simplicity of HCI. This scalable solution is designed, tested, and documented to address the business requirements, workloads, and applications of HPE customers. The solution incorporates a wide range of products into a portfolio of repeatable, scalable, and composable technologies that are supported by Hewlett Packard Enterprise.

This guide describes switch and network requirements and best practices for the HPE Nimble Storage dHCI solution. Proper network architecture and configuration are vital to ensuring that deployment is successful and that enterprise workloads gain the most value from the solution.

The solution can be deployed with either HPE switches (such as HPE Aruba, HPE M-Series, or HPE FlexFabric) or customer-provided switches (such as Cisco Nexus). This guide also describes the requirements for leveraging third-party switches.

Target audience

The target audience for this document includes HPE Authorized Partner solution engineers, distributors, and value-added resellers, as well as customers. It also includes anyone who has questions related to best practices or who is looking for additional recommendations on how to connect an HPE Nimble Storage array to an Ethernet network. Readers should be familiar with Ethernet connectivity concepts and with the needs and requirements of their organization and its network infrastructure.

Document purpose

Readers can use this guide to achieve the following goals:

- Understand the networking requirements of the HPE Nimble Storage dHCI solution
- Leverage design guidance for successful and supported network topology for their solution
- Determine which third-party switches work with and are supported by the solution

NETWORKING REQUIREMENTS FOR THE HPE NIMBLE STORAGE DHCI SOLUTION

This section provides a basic description of the networking requirements for the HPE Nimble Storage dHCI solution.

Support for HPE and third-party switches

Hewlett Packard Enterprise offers fully qualified switches in the HPE M-Series, HPE FlexFabric, and HPE Aruba product lines. For a full list of qualified switches, see the Validated Configuration Matrix or HPE Nimble Storage dHCI Quickspecs.

The solution supports any third-party switch that can be configured to meet its networking requirements. For more information about minimum requirements for third-party switches, see HPE Nimble Storage Deployment Considerations for Networking. If you use direct-attach copper (DAC) cables, Hewlett Packard Enterprise recommends limiting them to a maximum of 5 meters in length and buying them from the same manufacturer as the switches. For distances greater than 5 meters, Hewlett Packard Enterprise recommends using Optical Multimode LC cables and SFP+ transceivers.

VLANs

VLANs provide a method of segmenting a network into related groups, improving the efficiency of traffic flow, and limiting the propagation of multicast and broadcast messages. Traffic between VLANs is blocked unless the VLANs are connected by a router, which increases security.

Proper VLAN planning and implementation is key to a successful and secure deployment of the HPE Nimble Storage dHCI solution. At least five VLANs must be configured for the initial deployment.

NOTE

At the time of the writing of this document, the HPE Nimble Storage dHCI solution does not support VLAN tagging for the management or iSCSI VLANs during deployment.

Table 1 lists the correct port modes to use on various HPE and Cisco Nexus switches.

TABLE 1. Required dHCI VLANs, Port mode, and VLAN types

VLAN	Port mode	VLAN type
Management VLAN	Trunk (hybrid for HPE M-Series)	Native (PVID VLAN for HPE FlexFabric)
VM network VLAN	Trunk (hybrid for HPE M-Series)	Tagged VLAN
iSCSI VLANs	Access	N/A
iLO VLAN	Access	N/A

After deployment is complete, it is possible to change the switch configuration to enable VLAN tagging end to end. However, the change must be made manually, and the process might be disruptive.

Management VLAN

The management VLAN carries management traffic for the VMware ESXi™ hosts and the HPE Nimble Storage array. It also carries virtual machine (VM) and VMware vSphere® Storage vMotion® traffic, enabling management and communication between the components.

NOTE
The ESXi management interface and the HPE Nimble Storage array management interface must be in the same subnet. The configuration and the physical ports of these interfaces cannot be changed after deployment.

iLO VLAN

Integrated lights-out (iLO) is the embedded HPE server management technology that provides out-of-band management capabilities. iLO can operate in either shared mode (iLO served from a network interface) or dedicated mode (iLO on a dedicated iLO port). A dedicated iLO VLAN is not necessary if the iLO VLAN is on the same network as the management traffic.

NOTE
iLO connectivity is required and must be used with the HPE Nimble Storage dHCI solution.

iSCSI VLANs

The iSCSI data traffic is carried over two iSCSI VLANs. These VLANs are separate networks that provide redundant paths so that if one link goes down, iSCSI data is still served. The iSCSI VLANs are defined on both switches, but each switch carries traffic for only one iSCSI VLAN, and iSCSI traffic does not route between switches through the inter-switch links (ISLs).

VM network VLANs

The HPE Nimble Storage dHCI solution does not configure the VM network during setup. The default VM network port group will be present on **vSwitch0** at deployment, but any additional configuration must be performed manually after the initial deployment. The VM network VLAN should be labeled **trunk** and each port group should be tagged with the corresponding VLAN ID.

Dead network VLAN

A dead network VLAN is a VLAN to which unused ports are assigned. Having a dead network VLAN provides a safeguard so that if a cable is attached to an unused port, whether accidentally or intentionally, that port will not provide access to any traffic on the network.

The **network automation** feature configures a dead network VLAN as part of its workflow. If the network automation feature is not used, the HPE Nimble Storage dHCI solution itself does not require the use of a dead network VLAN. However, Hewlett Packard Enterprise recommends that you include one, and having one is considered best practice by most switch vendors.

Subnets

Subnets are required for management, iLO, and iSCSI traffic.

Management subnet

The management subnet carries management and vMotion traffic between the storage array and the compute nodes. It is possible, but not required, to include iLO traffic in this subnet.

For proper functionality of the solution's management software, the storage array management interfaces must be able to reach the compute node iLO interfaces through **port 443**.

For more information about requirements for firewall ports, see the HPE Nimble Storage Solution dHCI Security Guide – Secure Service Architecture Overview, available on the HPE InfoSight portal.

iLO subnet

The iLO subnet handles HPE ProLiant DL iLO traffic, and it must be accessible by the HPE Nimble Storage management subnet through port 443.

iSCSI subnets

The solution has two separate subnets for iSCSI traffic. These subnets do not route to each other.

Spanning tree

Each interface that is used for iSCSI traffic must be configured as **spanning tree edge ports**. Hewlett Packard Enterprise recommends that you use spanning tree mode **rpvst** if it is supported by the switches.

MTU

Many switches define **maximum transmission unit (MTU)** differently from the way the initiator or target defines it. Switches often define MTU as the frame size. End hosts almost universally define MTU as the packet size. The configured frame size on the switch might need to be larger than the packet size or the MTU value defined on the host and the array. For example, a value of 9000 on the host might require a value of 9014 or higher on the switch. When considering switch MTU size, it is important to take into consideration any VLAN tagging that might be used

This difference might vary by manufacturer. Setting the switch MTU value to a number that is higher than the MTU value on the host or initiator does not cause problems. The switch MTU setting causes problems only when the MTU value on the intermediate device (the switch) is set to a number that is lower than the MTU value on one or both end devices.

Inter-switch links

The solution does not have specific requirements for ISLs, but Hewlett Packard Enterprise recommends using them, depending on network topology. Hewlett Packard Enterprise recommends that you include at least two ISLs per switch pair at the highest possible speed to create redundancy and minimize latencies. If iSCSI traffic is carried on dedicated switches, it is not necessary to have ISLs between them.

LLDP

Link Layer Discovery Protocol (LLDP) must be enabled on each switch. The HPE Nimble Storage dHCI solution uses LLDP both during and after deployment to verify the environment.

Out-of-band management

Out-of-band management switches can be used with the HPE Nimble Storage dHCI solution to connect the iLO port.

HPE Nimble Storage management ports can be connected to out-of-band switches, but the connections must be redundant. These ports cannot be connected to only one switch.

iLO shared mode

In shared mode, iLO traffic is shared on the same port as the management traffic on the HPE ProLiant DL FlexLOM (LAN on motherboard) interfaces. The configuration change to shared mode must be made before the start of deployment and whenever a new server is added.

iLO dedicated mode

With dedicated mode, iLO traffic comes from the dedicated, built-in iLO port on the HPE ProLiant DL server. Dedicated mode is the default setting for the HPE Nimble Storage dHCI solution.

Connectivity

Table 2 lists supported speeds, with minimum requirements, for connections to the compute node and the storage array.

TABLE 2. Speed requirements for compute node and storage array connections

Component	Required number of ports	Supported speeds	Notes
Compute node	4 (for iSCSI and management)	10 Gbps or greater per port	The solution does not support hardware iSCSI. It supports only the software iSCSI provided by VMware vSphere®.
Storage array	2 per controller for iSCSI 2 per controller for management	10 Gbps or greater 1 Gbps or greater	The solution does not support aggregating the iSCSI traffic by leveraging mechanisms such as LACP.

For more information, see the Validated Configuration Matrix or HPE Nimble Storage dHCI Quickspecs for supported NICs.

PEER PERSISTENCE NETWORKING REQUIREMENTS

To leverage peer persistence with the HPE Nimble Storage dHCI solution, the storage arrays at both sites must be of the same model.

Subnet

The management and iSCSI data subnets must be the same on both sites. This configuration does not support network address translation (NAT) between site networks.

Connectivity

The management and data networks must span the two sites. For complete connectivity and configuration requirements and recommendations between sites, see HPE Nimble Storage Peer Persistence Deployment Considerations, available on the HPE InfoSight portal.

VMWARE VSPHERE NETWORKING REQUIREMENTS

vSwitch configuration

By default, dHCI stack setup configures three standard vSwitches, one for management and VM traffic (**vSwitch0**) and two for iSCSI traffic (**iSCSI1** and **iSCSI2**). These vSwitches must not be changed or renamed after the initial setup.

vSwitch0

vSwitch0 is created with two port groups: the **management network** and the **VM network**. The ESXi management interface is configured as **VMkernel network interface vmk0** on the management network port group. By default, vmk0 is enabled for **vMotion** and **management** services.

Although the vMotion service is enabled on vmk0 by default, after dHCI deployment, it can be configured to reside on a different VMkernel network interface. The management service, however, must remain on vmk0. The default setting for the VM network port group is **no VLAN**.

The following tables list the default settings for vSwitch0, the management network, and the VM network.

TABLE 3. Default vSwitch0 settings

vSwitch0 setting	Value
MTU	1500
Promiscuous mode	Reject
MAC address changes	Accept
Forged transmits	Accept
Traffic shaping	Disabled
Load balancing	Route based on originating virtual port
Network failure detection	Link status only
Notify switches	Yes
Failback	Yes
Active adapters	Management interface 1 and management interface 2
Standby adapters	None

TABLE 4. Default management network port group settings

Management network port group setting	Value
VLAN ID	None (0)
Promiscuous mode	Reject
MAC address changes	Accept
Forged transmits	Accept
Traffic shaping	Disabled
Load balancing	Route based on originating virtual port
Network failure detection	Link status only
Notify switches	Yes
Failback	Yes
Active adapters	Management interface 1
Standby adapters	Management interface 2

TABLE 5. Default VM network port group settings

VM network port group setting	Value
VLAN ID	None (0)
Promiscuous mode	Reject
MAC address changes	Accept
Forged transmits	Accept
Traffic shaping	Disabled
Load balancing	Route based on originating virtual port
Network failure detection	Link status only
Notify switches	Yes
Failback	Yes
Active adapters	Management interface 1 and management interface 2
Standby adapters	None

iSCSI vSwitches

To service iSCSI traffic, two dedicated vSwitches are created: **iSCSI1** and **iSCSI2**. Each dedicated vSwitch contains one port group with one active adapter, corresponding to its iSCSI network. The dHCI stack setup process automatically configures MTU at 9000 or 1500 to match the MTU that was specified for the array's iSCSI interfaces at the time of deployment.

The iSCSI1 and iSCSI2 vSwitches are configured with port group **iSCSI network 1** and VMkernel network interface **vmk1** and port group **iSCSI network 2** and VMkernel network interface **vmk2**, respectively.

TABLE 6. Default iSCSI1 and iSCSI2 vSwitch and port group settings

iSCSI1 and iSCSI2 vSwitch setting	Value
MTU	9000 or 1500
Promiscuous mode	Reject
MAC address changes	Accept
Forged transmits	Accept
Traffic shaping	Disabled
Load balancing	Route based on originating virtual port
Network failure detection	Link status only
Notify switches	Yes
Failback	Yes
Active adapters	1 × iSCSI interface
Standby adapters	None

vMotion

By default, the vMotion service is configured on **VMkernel network interface vmk0** in the **management network port group**. If additional NICs and VMkernel interfaces are configured, the vMotion service can be configured on them as well.

VM network

By default, the VM network resides on vSwitch0. Additional VM port groups can be created and used as needed for the environment's architectural requirements.

The physical switch ports might require changes to enable them to accept more VLANs.

Additional network card

If additional NICs are needed for VM workloads on the compute nodes, or if there are requirements for distributed virtual switches, NICs can be added before or after deployment. These additional NICs are not configured during stack setup and will require post-deployment configuration.

These NICs reside outside the dHCI stack's management. Configurations supported by VMware® for distributed virtual switches, including the VMware NSX® network virtualization platform, are supported on these additional NICs. For more information about configuring and administering distributed virtual switches and NSX, see the relevant VMware documentation.

TESTED NETWORK DESIGNS FOR THE HPE NIMBLE STORAGE DHCI SOLUTION

The network designs described in this section offer examples of topologies that have been tested and are supported, but these designs are not a full representation of all potential supported configurations.

NOTE

For the sake of simplicity, the diagrams represent compute nodes by a single entity. They also show iLO connected to only the first management switch, although in real-world environments iLO can be connected to any of the management switches.

Design 1 – Configuration with two Ethernet switches

This is the default design that implementation services would configure for the HPE Nimble Storage dHCI solution. It is currently the only topology supported by the **network automation** feature.

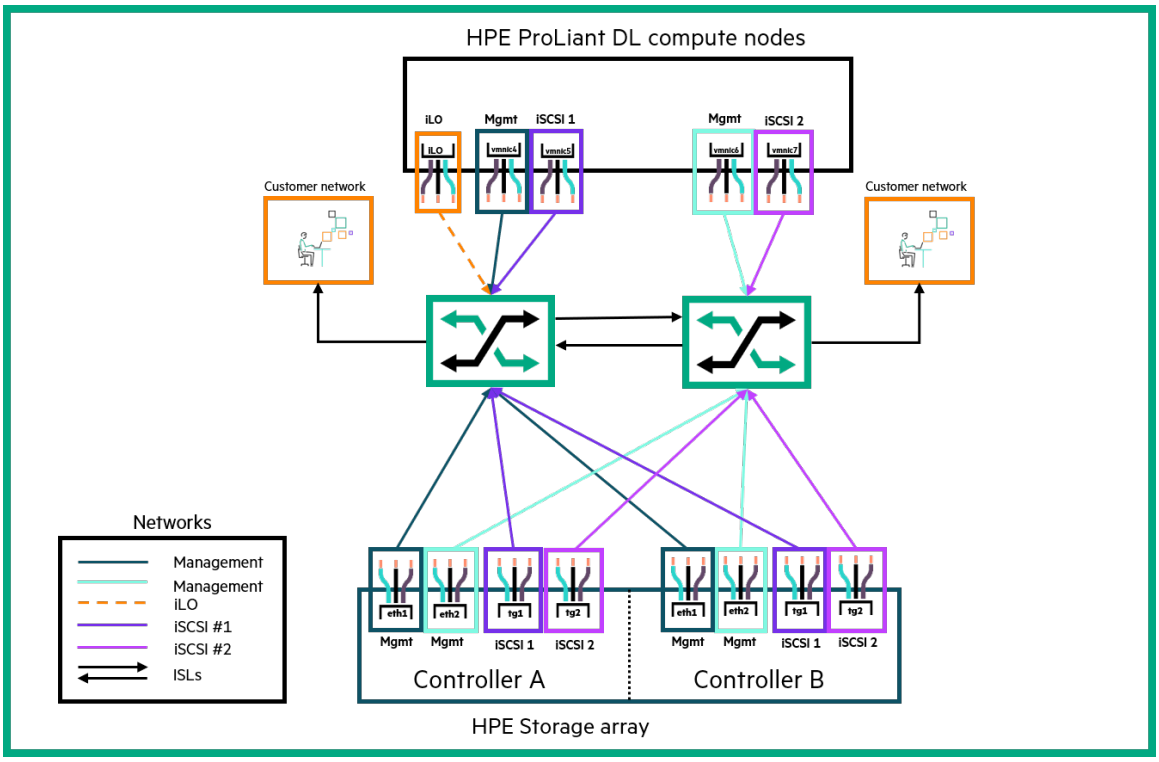


FIGURE 1. Configuration with two Ethernet switches

Design 2 – Configuration with two Ethernet switches and out-of-band management

Optionally, iLO traffic can be moved to an out-of-band switch. Note that it is necessary for the iLO switches to be connected back to the switches that carry management traffic to ensure proper deployment and operation of the HPE Nimble Storage dHCI solution.

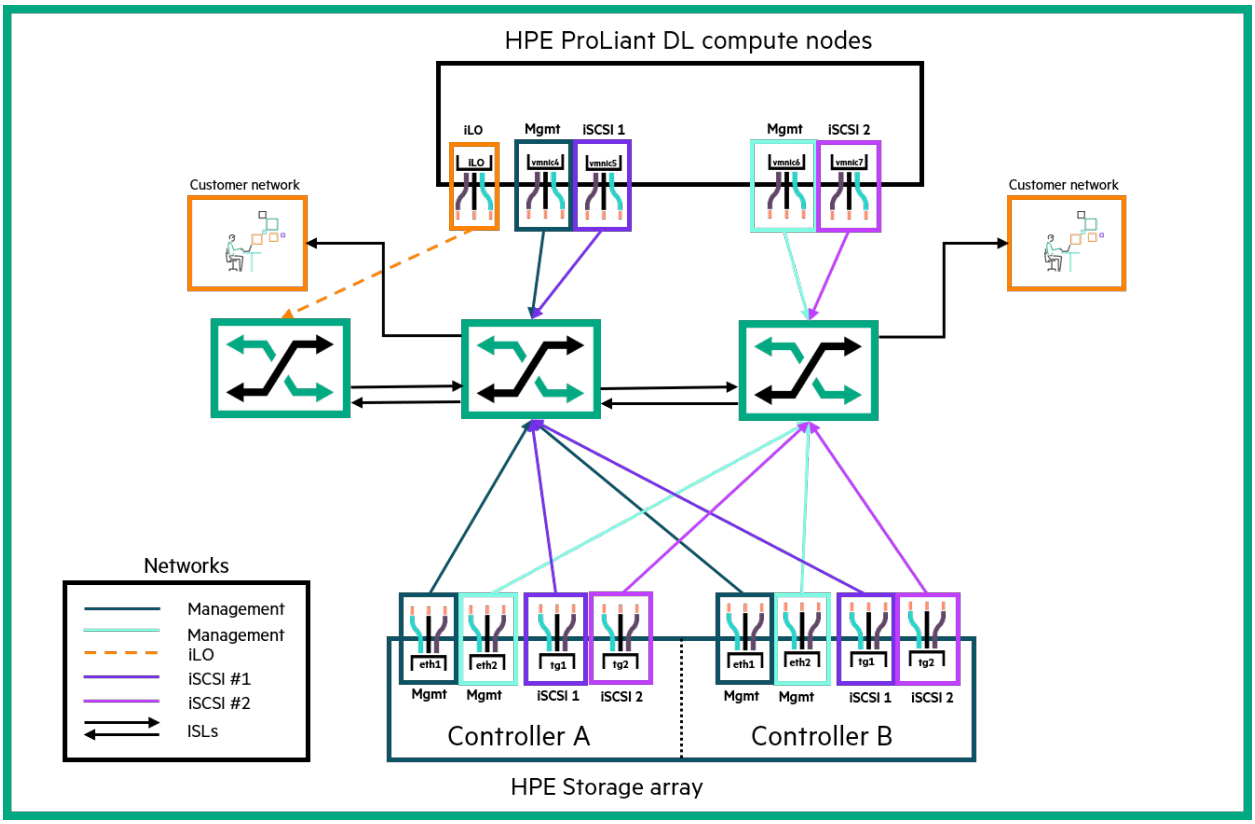


FIGURE 2. Configuration with two Ethernet switches and out-of-band management

Design 3 – Configuration with four Ethernet switches

If a configuration uses dedicated iSCSI switches, then—unless the array is also serving storage through iSCSI to clients outside the solution environment—it is not necessary to uplink to other switches. If the only traffic on these switches is dHCI iSCSI traffic, ISLs are not needed.

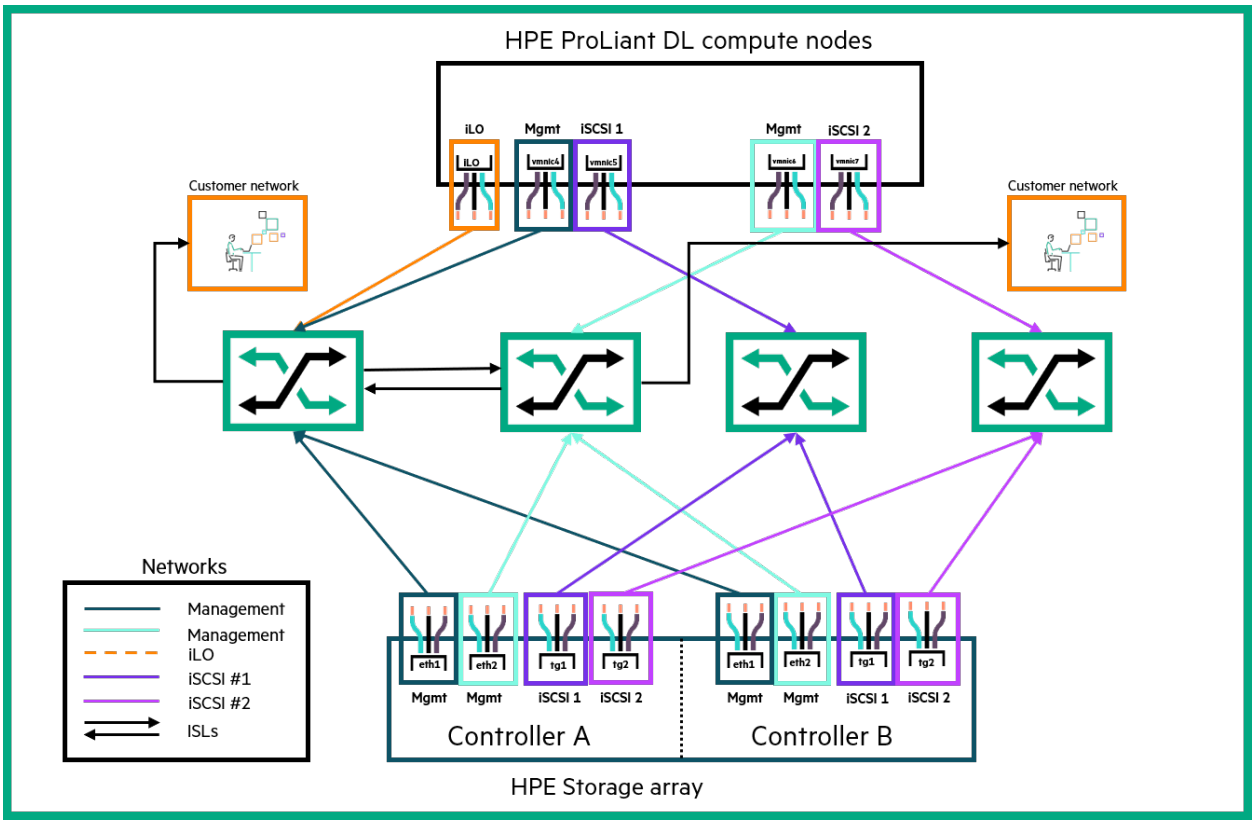


FIGURE 3. Configuration with four Ethernet switches

Design 4 – Configuration with multiple Ethernet switches and a spine-and-leaf topology

Hewlett Packard Enterprise fully supports deploying the HPE Nimble Storage dHCI solution on a spine-and-leaf topology network. This arrangement provides additional flexibility in relation to scale-out considerations and redundancy.

If the iSCSI VLANs will reside on nondedicated, multiprotocol switches, Hewlett Packard Enterprise recommends implementing **priority-based flow control (PFC)** on the iSCSI uplink ports to give iSCSI traffic high priority and lossless flows. If the interfaces carry more than a single protocol (for example, iSCSI and NFS), then PFC can be used on those interfaces.

NOTE

Do not enable PFC on any interfaces that are connected to HPE Nimble Storage arrays because it is not currently supported.

For more information about configuring PFC, see the relevant switch vendor's documentation.

Design 5 – Configuration with peer persistence

For simplicity of illustration, the diagram in Figure 4 uses the default two-switch configuration at each site. However, any of the previously described topologies (four-switch, two-switch with out-of-band management, and so forth) would be supported with peer persistence. Each switch at **site A** must have ISLs to its corresponding switch at **site B**.

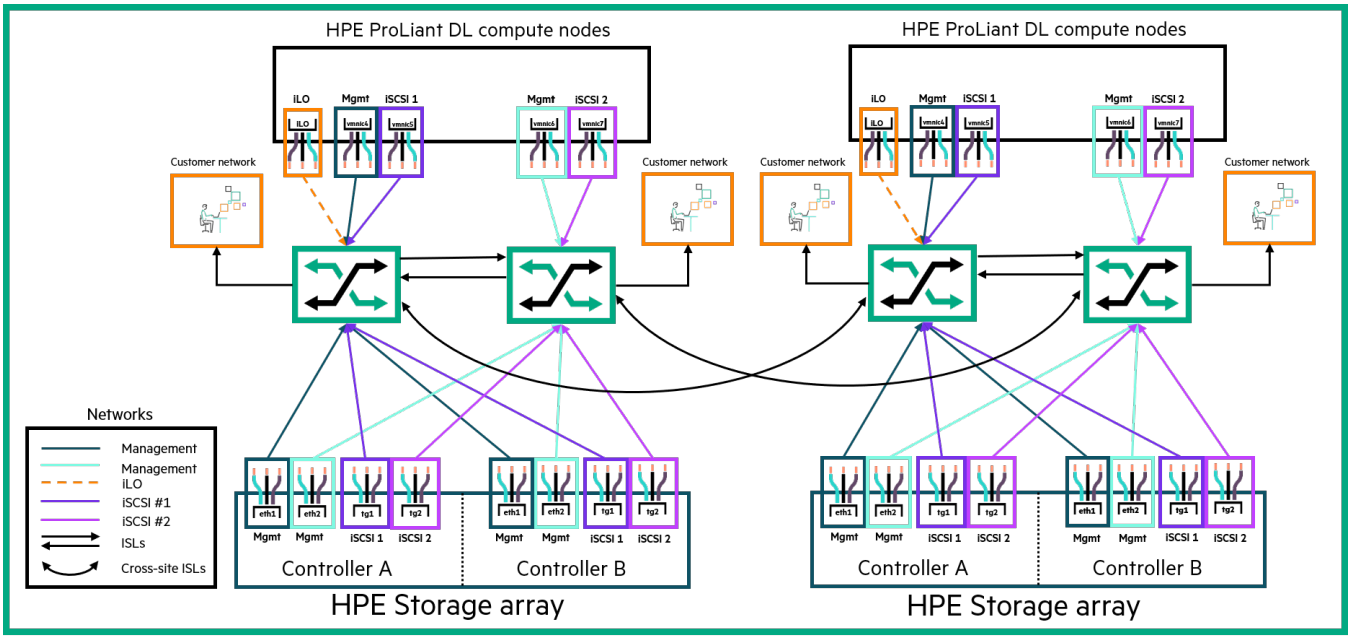


FIGURE 4. Configuration with peer persistence

SUMMARY

The HPE Nimble Storage dHCI solution provides extraordinary scalability and flexibility, coupled with ease of management. To ensure successful deployment, resiliency, and performance, it is important to understand the solution's networking requirements. Proper network planning and implementation lay the foundation for success in deployment and day-to-day operations.

GLOSSARY OF TERMS

- **Access VLAN**—The untagged VLAN on an access port. A switch port that is configured to associate a port with an access VLAN sends and receives all frames within the same VLAN. 802.1Q VLAN tags are not needed because all traffic is associated with the access VLAN.
- **Flow control**—A mechanism for temporarily pausing the transmission of data on Ethernet networks when a sending node transmits data faster than the receiving node can accept it.
- **Interface**—The point of interconnection between a computer and a private or public network. A network interface is generally a network interface card (NIC), but it does not have to have a physical form.
- **Inter-switch link (ISL)**—A connection that joins two switches and maintains VLAN information as traffic flows between switches and routers. The term ISL may refer to the proprietary Cisco Inter-Switch Link standard or to the general concept of connecting nonstacked switches through one or more Ethernet cables.
- **Jumbo frames**—An Ethernet frame with a payload greater than the standard MTU of 1500 bytes. Jumbo frames can be as large as 9000 bytes.
- **Link aggregation**—A way of bundling many individual Ethernet links together so that they act like a single logical link.
- **Multipath I/O (MPIO)**—A framework designed to mitigate the effects of adapter failure by providing an alternative data path between storage devices and hosts.
- **Maximum transmission unit (MTU)**—The largest size frame, specified in octets (eight-bit bytes), that can be sent in any Ethernet-based network. Network devices use the configured MTU setting to determine the maximum size of each frame in any transmission.
- **Native VLAN**—The VLAN ID to which untagged packets are assigned.
- **Spanning Tree Protocol (STP)**—A network protocol that builds a loop-free logical topology for Ethernet networks to prevent packet storms.
- **Stack/stacked switches/multiswitch stack**—A set of switches or switch modules that are connected through dedicated stacking ports or ports that are defined in the software to be stacking ports. Stacked switches typically behave as if they were one physical switch.
- **Virtual local area network (VLAN)**—A custom network that is used to segment broadcast domains.
- **VLAN tagging**—A method used by networked devices to identify which VLAN each individual packet should be associated with.

Resources, contacts, or additional links

HPE Nimble Storage Deployment Considerations for Networking
https://infosight.hpe.com/InfoSight/media/cms/active/public/HPE_Nimble_Storage_Deployment_Considerations_for_Networking.pdf

HPE Nimble Storage dHCI Solution Security Guide – Secure Service Architecture Overview
https://infosight.hpe.com/InfoSight/media/cms/active/public/HPE_Nimble_Storage_dHCI_Solution_Security_Guide.pdf

HPE Nimble Storage dHCI Quickspecs
https://h20195.www2.hpe.com/v2/getdocument.aspx?docname=a00067739enw

HPE Nimble Storage Peer Persistence Deployment Considerations
https://infosight.hpe.com/InfoSight/media/cms/active/public/HPE_Nimble_Storage_Peer_Persistence_Deployment_Considerations.pdf

Validated Configuration Matrix
https://infosight.hpe.com/resources/nimble/validated-configuration-matrix

LEARN MORE AT

hpe.com/storage

a00114771ENW, May 2021

Hewlett Packard Enterprise