White Paper

NetApp HCI Theory of Operations

Version 1.7

Daniel Elder, NetApp May 2020 | WP-7261

Abstract

Hyper converged infrastructure (HCI) has promised the industry simplicity of deployment, ease of operation, and scale. However, the inherent design of most modern HCI systems limits their ability to deliver on one or more of these promises.

This white paper describes how NetApp® HCI brings simplicity of configuration, efficiency of operation, and elasticity of scale together in a single product.



TABLE OF CONTENTS

1	HCI	and the Theory of Everything	5
	1.1	Performance Guarantee	5
	1.2	Enterprise Scale	5
	1.3	Streamline Operations	5
	1.4	Configuration	6
2	Sys	tem Design and Architecture	8
	2.1	Overview	
	2.2	Networking	9
	2.3	Storage Components	.12
	2.4	vSphere Components	.17
3	Exp	anding NetApp HCI	28
	3.1	Adding Nodes	.28
4	Mix	ed Workloads	30
	4.1	Test Configuration	.30
	4.2	Results	.32
5	Ext	ending NetApp HCI	33
	5.1	NetApp ONTAP Select	.33
	5.2	NetApp SnapMirror	.34
	5.3	vRealize Orchestrator Plug-In	.35
	5.4	NetApp HCI and SolidFire vRealize Operations Management Pack	.36
	5.5	External Host Access	.37
	5.6	HCI Collector Project	.38
6	Cor	nclusion	39
Wł	nere	to Find Additional Information	39
Ve	rsior	ı History	39
LIS	ST OI	FTABLES	
	-	NetApp HCI storage nodes.	
	-	NetApp HCI compute nodes	
		Node interface requirements.	
		vSphere components.	
Tal	ole 5)	Mongo replica set cluster VMs.	.31

Table 6) Yahoo! Cloud Serving Benchmark (YCSB) client-server VM	31
Table 7) SQL VMs	31
Table 8) VDI VMs.	31
LIST OF FIGURES	
Figure 1) Minimum configuration.	6
Figure 2) NetApp HCI with H610S storage nodes.	8
Figure 3) HCI components.	8
Figure 4) Compute node ports	10
Figure 5) Six-cable compute option.	10
Figure 6) Two-cable compute option.	11
Figure 7) ZeroConf.	11
Figure 8) HCI protection domains	12
Figure 9) NetApp vCenter Plug-In for HCI	13
Figure 10) NMA options	14
Figure 11) NetApp Monitoring Agent.	15
Figure 12) Active IQ capacity forecasting.	15
Figure 13) Active IQ vCenter reporting.	16
Figure 14) Active IQ QoS recommendations.	16
Figure 15) Compute node physical-to-logical port map (VSS)	18
Figure 16) VMkernel binding.	19
Figure 17) Management port groups configuration on VMware standard switches	19
Figure 18) vMotion and VM_Network configuration on VMware standard switches	20
Figure 19) iSCSI-A and iSCSI-B configuration on VMware standard switches	20
Figure 20) VDS uplink names	21
Figure 21) Compute node physical-to-logical port map (VDS).	21
Figure 22) Management port groups configuration on VMware distributed switches	22
Figure 23) vMotion and VM_Network configuration on VMware distributed switches	22
Figure 24) iSCSI-A configuration on VMware distributed switches	23
Figure 25) iSCSI-B configuration on VMware distributed switches	23
Figure 26) Compute node physical-to-logical port map (VDS 2 cable)	24
Figure 27) Example configuration for management, vMotion, and user VM port groups	25
Figure 28) Host IQN: iSCSI name.	25
Figure 29) Volume Access Group.	26
Figure 30) Default datastores.	27
Figure 31) Create datastore.	27
Figure 32) Adding nodes to NetApp HCI.	29
Figure 33) Mixed workloads.	31
Figure 34) VDI boot storm with QoS disabled.	32

Figure 35) VDI boot storm with QoS enabled	33
Figure 36) ONTAP Select Deploy appliance.	34
Figure 37) ONTAP Select on HCI	34
Figure 38) SnapMirror data replication	35
Figure 39) vRO workflows.	36
Figure 40) vRealize Operations Manager Plug-In.	37
Figure 41) External host access	38
Figure 42) HCI Collector interface.	38

1 HCI and the Theory of Everything

NetApp® HCI is an enterprise-scale hybrid cloud solution that is ideally suited for customers who are looking to break free from first-generation HCI limitations.

NetApp HCI customers can run multiple applications with guaranteed performance to confidently deploy resources across their entire data center. This architecture allows you to deploy your infrastructure by simplifying management and independently scaling both compute and storage resources. NetApp HCI is Data Fabric ready out of the box for easy access to all your data across any public, private, or hybrid cloud. By moving to NetApp HCI, you can transform your data center, driving operational efficiencies and reducing costs.

The NetApp Data Fabric is a software-defined approach for data management that enables you to connect disparate data management and storage resources. NetApp HCl can streamline data management between on-premises and cloud storage for enhanced data portability, visibility, and protection.

1.1 Performance Guarantee

A common challenge for a data center is delivering predictable performance, a goal complicated by running multiple applications that share the same infrastructure. An application that interferes with other applications creates performance degradations, which forces IT administrators to spend valuable time troubleshooting the environment. Mainstream applications, such as virtual desktop infrastructure (VDI) and database applications, have unique I/O patterns that can affect one another's performance during normal operations when they are deployed in a shared environment.

The NetApp HCI quality of service (QoS) feature enables fine-grained control of performance for every application, eliminating noisy neighbors, meeting unique performance needs, allowing higher utilization of infrastructure, and satisfying performance SLAs. The storage architecture, which is part of the NetApp HCI solution, eliminates performance variance in the context of data locality because the data is distributed across all the nodes in the HCI cluster.

1.2 Enterprise Scale

Unlike previous generations of HCI, which have fixed resource ratios, NetApp HCI scales compute and storage resources independently. Independent scaling avoids costly and inefficient overprovisioning and simplifies capacity and performance planning. NetApp HCI is an enterprise-scale hybrid cloud infrastructure solution that runs on innovative NetApp SolidFire® Element® technology and is delivered on an architecture designed by NetApp. NetApp HCI comes in 2RU x 4-node building blocks (chassis) in small, medium, and large storage and compute configurations that can be mixed and matched. HCI can be scaled so that you can rapidly meet changing business needs on your terms.

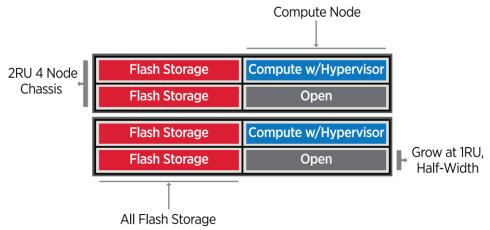
1.3 Streamline Operations

A common goal of IT organizations is to automate all routine tasks and eliminate the risk of user errors associated with manual operations. Automation allows you to focus valuable resources on higher value priorities that drive business efficiencies. The NetApp Deployment Engine (NDE) streamlines day zero installation from hours to minutes. Centralized management through the vCenter Plug-In gives you full control over your infrastructure with an intuitive UI. A robust suite of APIs allows you to seamlessly integrate higher-level management, orchestration, backup, and disaster-recovery tools.

1.4 Configuration

NetApp HCI is available with a range of configuration options for both compute and storage. A minimum starting configuration must have four storage nodes and two compute nodes (Figure 1).

Figure 1) Minimum configuration.



As shown in the configuration information in Table 1, each storage node can deploy 4.5TB to 78TB of effective capacity. From a compute-node perspective, 16 to 36 CPU cores and 384GB to 768GB of RAM are available, as shown in Table 2.

Table 1) NetApp HCl storage nodes.

Model	Form Factor	Performance (4K IOPS)	Drive Size (GB)	Raw Block Capacity (TB)	Usable Capacity (TB)	Effective Capacity (TB)
H300S	1RU, half width	50K	480	2.4	1.1	4.5
H500S	1RU, half width	50K	960	4.8	2.2	8.9
H700S	1RU, half width	100K	1920	9.6	4.5	17.9
H610S-1	1RU, full width	100K	960	10.6	4.9	19.6
H610S-2	1RU, full width	100K	1920	21.1	9.8	39.3
H610S-4	1RU, full width	100K	3840	42.2	19.6	78.6

Table 2) NetApp HCI compute nodes.

Target Use Case	Model	RAM	Per-Node CPU Metrics					
			Model	Family	Cores	Threads	Base GHz	Turbo GHz
Virtualization	H300E	384GB	E5-2620 v4	Broadwell	16	32	2.1	3.0
Virtualization	H500E	512GB	E5-2650 v4	Broadwell	24	48	2.2	2.9

Target Use Case	Model	RAM	Per-Node CPU Metrics					
			Model	Family	Cores	Threads	Base GHz	Turbo GHz
Virtualization	H600E	768GB	E5-2695 v4	Broadwell	36	72	2.1	3.3
VDI	H410C- 14020	384GB	4110	Skylake	16	32	2.1	3.0
VDI	H410C- 15020	512GB	4110	Skylake	16	32	2.1	3.0
VDI	H410C- 17020	768GB	4110	Skylake	16	32	2.1	3.0
VDI/Virtualization	H410C- 25020	512GB	5120	Skylake	28	56	2.2	3.2
VDI/Virtualization	H410C- 27020	768GB	5120	Skylake	28	56	2.2	3.2
VDI/Virtualization	H410C- 28020	1TB	5120	Skylake	28	56	2.2	3.2
Virtualization	H410C- 57020	768GB	6138	Skylake	40	80	2	3.7
Virtualization	H410C- 58020	1TB	6138	Skylake	40	80	2	3.7
Database	H410C- 35020	512GB	5122	Skylake	8	16	3.6	3.7
Database	H410C- 37020	768GB	5122	Skylake	8	16	3.6	3.7

For example, for a minimum-size starting solution, a configuration with two small compute nodes and four small storage nodes could have 32 cores with 768GB of memory and 18TB of effective capacity. As requirements change, you can add more compute or storage nodes of any size to the chassis independently of each other. The flexibility to add only compute or storage nodes provides unique scalability options for building an efficient and agile cloud in your data center for various use cases.

H410C Series

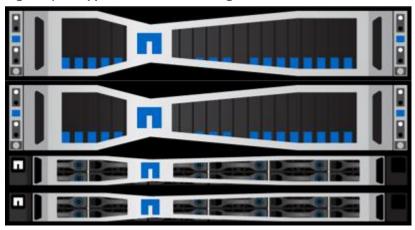
The NetApp HCl 1.4 release brings support for the H410C series of compute nodes. These nodes are based on the Skylake Scalable Processor family and offer a broad range of CPU and memory configurations to support a wide range of virtualization workloads.

H610S Series

The 1.4 release of NetApp HCI adds support for the H610S series nodes, which offer greater capacity and are more cost effective than other HCI storage-node types. The top line benefits of the new H610S nodes are as follows:

- More than four times the density of existing HCI storage nodes
- NVMe SSDs and 25GbE
- Compatible with existing HCI and SolidFire clusters

Figure 2) NetApp HCI with H610S storage nodes.

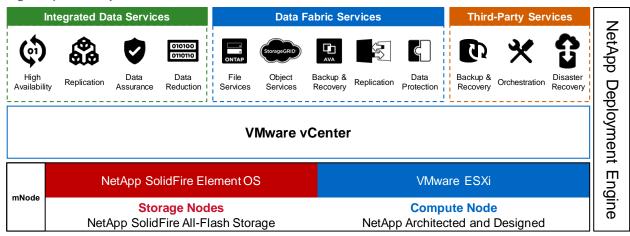


2 System Design and Architecture

2.1 Overview

Although NetApp HCI is a new product, it is built from best-in-class components available from VMware and NetApp SolidFire. The hypervisor layer is provided by vSphere 6.x, which is the premier virtualization provider in the market. The storage layer is provided by SolidFire Element version 10, which offers a robust and unique scale-out model for NetApp HCI. The management layer is provided by a vCenter Plug-In.

Figure 3) HCI components.



The NDE orchestrates the initial deployment and configuration of the HCI components. The NDE is responsible for collecting all user inputs that are required to completely install and configure the system. The NDE reduces the number of required user inputs from more than 400 to fewer than 40.

The following sections describe in detail the networking, virtualization, and storage pieces of the NetApp HCI solution.

2.2 Networking

During deployment of a NetApp HCI system, the NDE installs and configures multiple components. Therefore, it requires a properly configured and well-functioning network. In general, the network must support the following configuration:

- MTU 9000 on all 10/25GbE ports
- Separate VLANs for management, vMotion, and iSCSI (single VLAN deployments are not supported)
- All host ports should have a spanning tree configured to immediately enter a forwarding state (PortFast)

Table 3 describes the basic requirements for the networks ports of each node type.

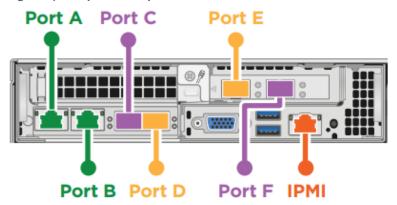
Table 3) Node interface requirements.

Node Type	Interface	Requirements
Storage node	Bond1G - management ports A and B	Ping all other Bond1G devices Dedicated management VLAN Spanning tree edge port
	Bond10G - iSCSI ports C and D	Ping all other Bond10G devices MTU set to 9000 (no fragmentation) Dedicated iSCSI VLAN Spanning tree edge port LACP (recommended but optional)
Compute node	Bond1G - management ports A and B	Ping all other Bond1G devices Dedicated management VLAN No LACP Spanning tree edge port
	Bond10G - iSCSI ports D and E	Ping all other Bond10G devices MTU set to 9000 (no fragmentation) Dedicated iSCSI VLAN No LACP Spanning tree edge port
	VM_Network and vMotion ports C and F	MTU set to 9000 (no fragmentation) vMotion and VM_Network VLANs No LACP Spanning tree edge port

Flexible Cabling for Compute Nodes

NetApp HCI 1.4 and later supports the FlexCable feature, which provides the option of configuring the NetApp HCI compute nodes with either two or six cables. When you are deploying a node with two cables, only ports D and E are configured by NDE. All traffic types are present on these two ports (management, vMotion, iSCSI, and VM traffic). VLANs are used to separate traffic types in this case.

Figure 4) Compute node ports.



This capability permits three possible deployment scenarios when running NDE:

- Connect and deploy with six cables (ports A through F)
- Connect six cables, deploy with two cables (ports D and E)
- Connect and deploy with two cables (ports D and E)

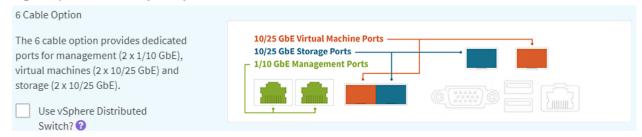
Note: Storage nodes require four cables to be connected and configured during deployment.

Connect and Deploy with Six Cables

Choose this option to get the optimum performance from your HCI deployment and the maximum number of options for separating out different traffic types.

This option requires the cabling of all network ports. The six-cable option provides dedicated ports for management (2 x 1/10GbE), virtual machines (2 x 10/25GbE), and storage (2 x 10/25GbE). VMware standard switches (VSS) are the default virtual switch type, with the vSphere distributed switch (VDS) being optional. To take advantage of the VDS option, you need a VMware Enterprise Plus license.

Figure 5) Six-cable compute option.



Connect Six Cables, Select Two-Cable Configuration

Choose this option when you want to leave ports A, B, C, and F unconfigured by NDE. This consolidates the NDE-managed port groups and traffic types to ports D and E. This scenario is beneficial for installations in which some network resources must be physically separated from the management and data plane of the core HCI infrastructure. Some possible use cases are a network DMZ and VMware NSX deployments. This option requires the use of a VDS.

Note: The VDS and VMware NSX require a VMware Enterprise Plus license.

Figure 6) Two-cable compute option.



Connect and Deploy with Two Cables

Choose this option when you must deploy with the smallest cable plant possible. Connecting and deploying with two cables is the least complicated network configuration because all NetApp HCl core infrastructure and customer network traffic share the two uplinks. This option requires the use of a VDS.

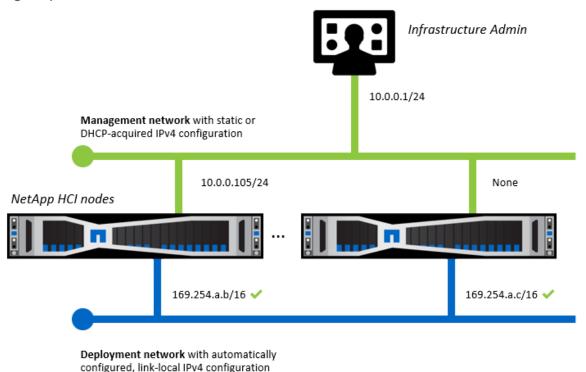
Note: The VDS requires a VMware Enterprise Plus license.

ZeroConf IP Assignment

NDE 1.3 and later includes the ZeroConf feature that autoconfigures a private nonroutable network on Bond10G interfaces. These private IPs are used to set up a temporary network for NDE to use for initial configuration. This eliminates the need to assign temporary routable IP addresses before starting the NDE.

ZeroConf requires the assignment of a single, temporary routable IP address to one of the storage nodes. This IP address is used to drive the NDE. During deployment, the permanent IP addresses are applied to all nodes in the system.

Figure 7) ZeroConf.



2.3 Storage Components

SolidFire Cluster

The first component to be configured during a system installation is the SolidFire storage cluster. The NDE uses the inputs for the desired user account, the management and storage IPs, and the nodes to include in the system to create and configure the SolidFire cluster.

After the cluster is created, the NDE adds all available drives to the cluster and then creates two 2TB volumes. These volumes are formatted with the Virtual Machine Filing System (VMFS) and receive the vCenter Server Appliance (VCSA) Open Virtual Applicance (OVA) and SolidFire mNode OVA when the NDE installs vSphere components. A volume access group is also created to hold the ESXi iSCSI qualified name (IQNs).

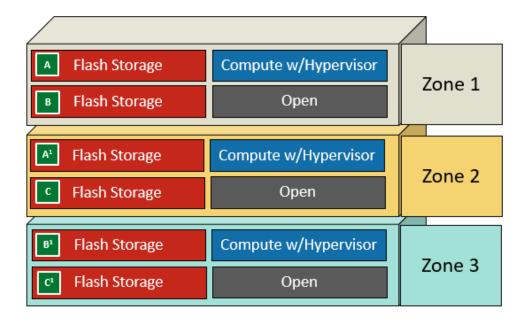
Protection Domains

Protection domains at the chassis level allow an HCI customer to scale their clusters in such a way that the failure of a chassis can be sustained even if there is more than one storage node in the chassis. The minimum allowed configuration is three chassis with two storage nodes each, for a six-node storage cluster. The system automatically lays out data correctly when the configuration is compatible with protection domains.

Protection domains provide the following benefits:

- · Reduced risk of downtime due to a chassis failure
- Automatic detection of chassis and node configuration
- Self-healing
- Domain-level capacity monitoring

Figure 8) HCI protection domains.



SolidFire mNode

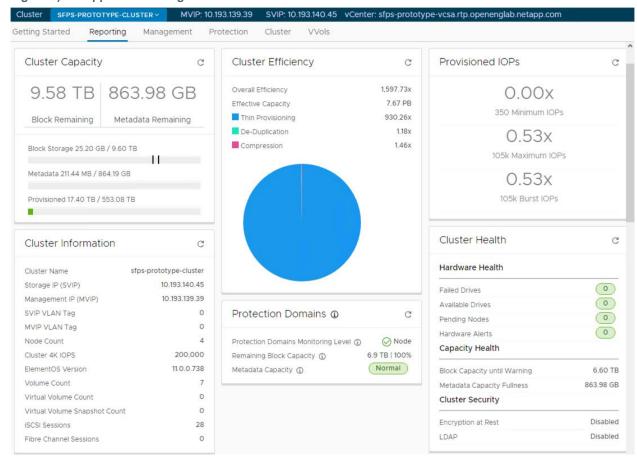
The SolidFire management node (mNode) is one of the two virtual machines automatically deployed and configured on every NetApp HCI system during a new installation. The mNode enables the following features:

- NetApp HCI vCenter Plug-In
- NetApp Monitoring Agent (NMA)
- NetApp Active IQ[®] Collector
- SolidFire support reverse VPN tunnel
- System upgrades

NetApp vCenter Plug-In for HCI

The NetApp vCenter Plug-In (VCP) is the management focal point for the HCl system. It brings day-to-day storage management tasks into vCenter and enables the NetApp Monitoring Agent.

Figure 9) NetApp vCenter Plug-In for HCI.



All common day-to-day storage management tasks are available in the VCP. The VCP has five sections related to SolidFire integration with vCenter:

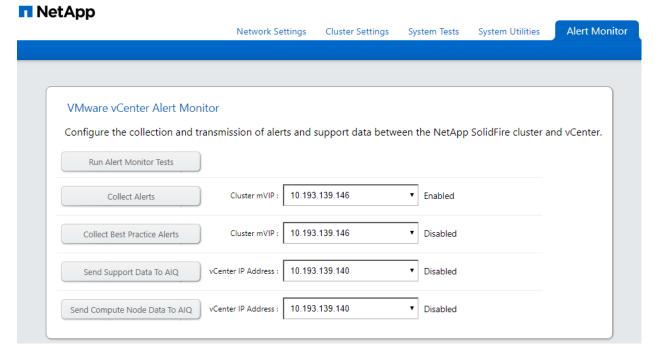
- Reporting. Includes system overview, event logs, alerts, and running tasks on the SolidFire cluster.
- Management. Tasks for creating, deleting, and managing datastores and SolidFire volumes, accounts, access groups, and ESXi initiators. QoS policies can also be configured here.

- **Data protection.** Management of SolidFire replication, NetApp Snapshot™ copies, group Snapshot copies, and Snapshot copy schedules.
- Cluster. Tasks for adding and removing SolidFire nodes and disks to and from the cluster as well as managing virtual iSCSI networks.
- VMware Virtual Volumes (VVols). Reporting and management of VVols, storage containers, and
 protocol endpoints. The VVols feature is not enabled by default. For information about enabling and
 configuring VVols, see <u>TR-4642</u>: NetApp SolidFire VMware vSphere Virtual Volumes for SolidFire
 Storage Configuration Guide.

NetApp Monitoring Agent

The NetApp Monitoring Agent (NMA) is a component of the VCP that consolidates the events from the compute and storage components into a single view that can easily be consumed in vCenter. NMA features can be enabled or disabled from the SolidFire management node (mNode) interface.

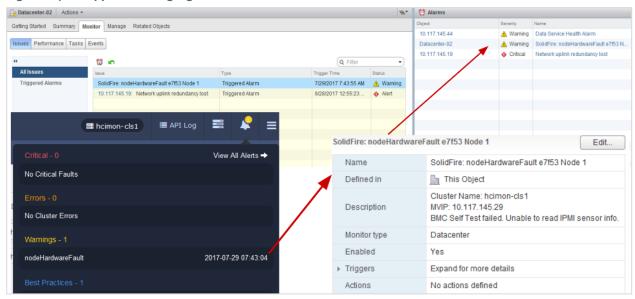
Figure 10) NMA options.



The NMA includes the following options:

- Run alert monitor tests. Tests the functionality of the NMA agent.
- Collect alerts. Toggles the collection of all alerts. Must be turned on to enable any of the alert collection options. Enabled by default.
- Collect best practice alerts. Enables reporting of best practice alerts from the SolidFire cluster to vCenter. This reporting includes network misconfigurations and other issues that indicate a suboptimal configuration. Disabled by default.
- **Send support data to Active IQ.** Sends vCenter alarms to Active IQ. Required for SolidFire to proactively engage support issues. Enabled by default.
- Send compute node data to Active IQ. Sends vCenter telemetry data and logs to Active IQ. This option enables Active IQ to collect support bundles and provide reporting specific to vCenter in Active IQ. Enabled by default.

Figure 11) NetApp Monitoring Agent.



For more information about using the VCP, see the vCenter Plug-In User Guide available under NetApp HCI Resources.

Active IQ

Active IQ is a comprehensive reporting platform for NetApp HCl and SolidFire systems that enables near-real time and historical reporting of configuration, alerting, and performance data.

Figure 12) Active IQ capacity forecasting.

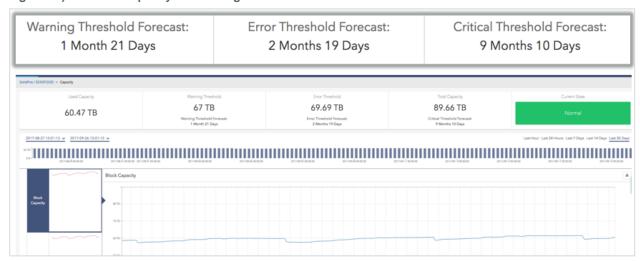
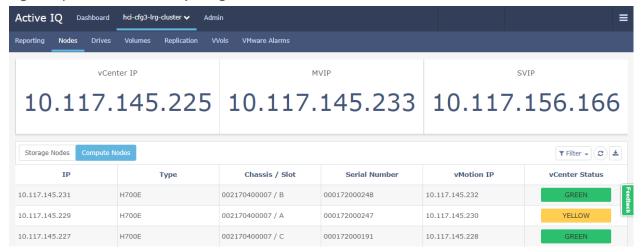
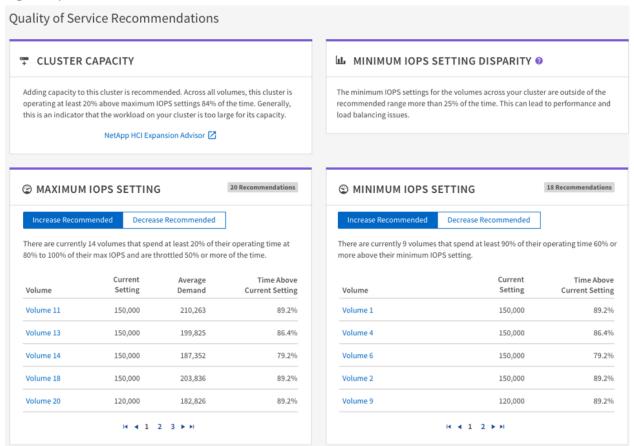


Figure 13) Active IQ vCenter reporting.



With the NDE 1.6 release, Active IQ also now supports QoS recommendations based on historical data that are updated every day.

Figure 14) Active IQ QoS recommendations.



More information about using Active IQ is available online at Active IQ Help.

2.4 vSphere Components

The NDE completely automates the installation and configuration of vSphere components and prepares the system according to established best practices.

Note: These versions do not contain the required patches for Spectre and Meltdown. For more information, see this <u>VMware KB article</u>.

Table 4) vSphere components.

Component	Automation Tasks
ESXi	 Install ESXi 6.0 U3a, ESXi 6.5 U2, or ESXi 6.7 U1. Create vSwitches and port groups. Enable the software iSCSI adapter. Configure VMkernel ports for vMotion and iSCSI. Configure VMkernel port binding for iSCSI. Add a dynamic discovery address for the SolidFire storage virtual IP (SVIP) address. Add an ESXi iSCSI IQN to the SolidFire volume access group. Provision two VMFS datastores for vCenter and the SolidFire mNode.
vCenter	 Deploy vCenter VCSA 6.5.0-9451637 or 6.7.0-10244745. Create default data center and cluster objects. Enable and configure the Distributed Resource Scheduler (DRS) and high availability (HA). Install and register the vCenter HCl Plug-In. Register the SolidFire system with the vCenter Plug-In. Configure NetApp Monitoring and Alerting (NMA).

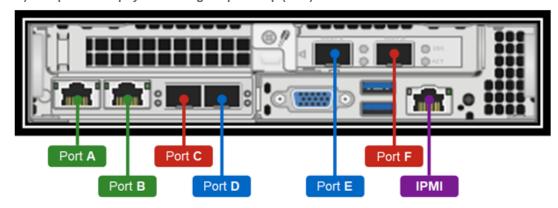
The NDE can be deployed with either VDS or VMware standard switches (VSS). VSS deployments require a six-cable configuration. For two-cable configurations, a VDS is required.

Compute Node Physical-to-Logical Port Map (VSS – Six Cable)

Deployments based on VSS have a set of three logical switches configured:

- vSwitch0: management VMs and ESXi management interfaces
- vSwitch1: vMotion and customer VM network traffic
- vSwitch2: iSCSI interfaces

Figure 15) Compute node physical-to-logical port map (VSS).



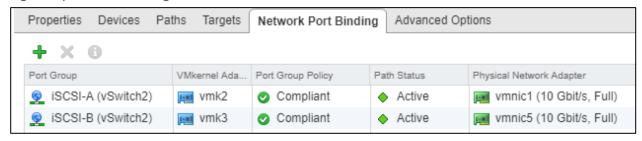
Port	Α	A B		F C		E	
Interface	eth2	eth3	eth5	eth0	eth1	eth4	
vmnic	vmnic2	vmnic2 vmnic3 vmnic0		vmnic4	vmnic1	vmnic5	
vSwitch	vSwi	itch0	vSw	itch1	vSwitch2		
Portgroup	HCI_Internal_vCenter_Network HCI_Internal_mNode_Network HCI_Internal_OTS_Network Management Network				iSCSI-A	iSCSI-B	
VLAN	Managem	ent VLANs	User VM	s/vMotion	iscsi vlan		
vmkernel	vm	nk0	vn	nk1	vmk2	vmk3	

For NetApp HCI compute nodes, the 1GbE interfaces eth2 and eth3 are bound to vSwitch0. These interfaces are used for VMware ESXi management traffic and the VCSA, mNode, and ONTAP Select management interfaces. NetApp recommends that the management VLAN be native on these interfaces.

10GbE interfaces eth5 and eth0 are assigned to vmnic0 and vmnic4, respectively, and bound to vSwitch1. The port groups for vMotion and VM_Network are assigned so that each port group has one active adapter and one standby adapter.

10GbE interfaces eth1 and eth4 are assigned to vmnic1 and vmnic5, respectively. These ports are configured for VMkernel port binding exclusively for storage connectivity and are configured in a manner to support it.

Figure 16) VMkernel binding.



Any additional customer VMs or iSCSI clients that require direct access to iSCSI storage hosted on the storage nodes should have their network ports mapped to this vSwitch through an additional VM port group.

Figures 16, 17, and 18 show the NetApp HCl compute node virtual switch configuration when using VMware standard switches.

Figure 17) Management port groups configuration on VMware standard switches.

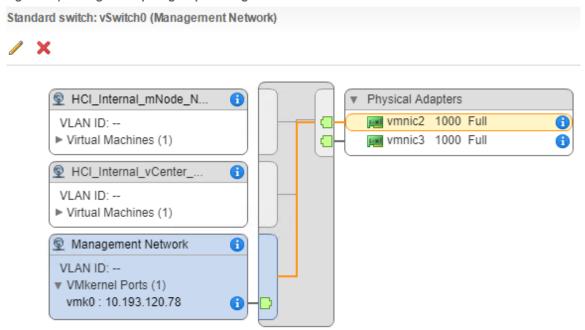


Figure 18) vMotion and VM_Network configuration on VMware standard switches.

Standard switch: vSwitch1 (VM_Network)



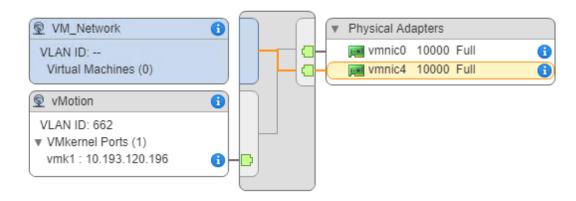
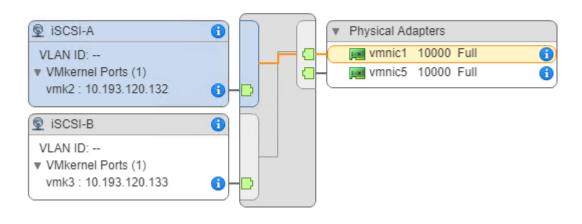


Figure 19) iSCSI-A and iSCSI-B configuration on VMware standard switches.

Standard switch: vSwitch2 (iSC SI-A)





Compute Node Physical-to-Logical Port Map (VDS Six-Cable)

In VDS-based deployments, a single vSphere distributed switch is configured that contains all available host uplinks. The uplinks are named according to their role in the port groups configured on the VDS.

Figure 20) VDS uplink names.

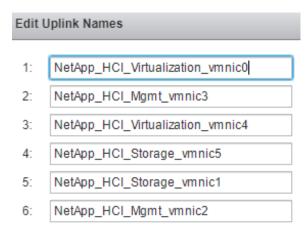
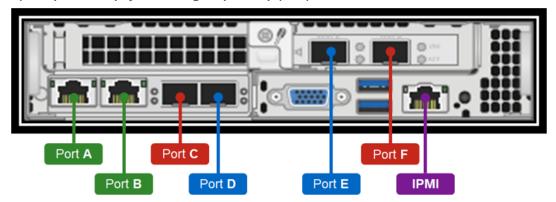


Figure 21) Compute node physical-to-logical port map (VDS).



Port	Α	В	F	С	D	Е
Interface	eth2	eth3	eth5	eth0	eth1	eth4
vmnic	vmnic2	vmnic3	vmnic0 vmnic4		vmnic5	vmnic1
vSwitch	NetApp HCI VDS					
Portgroup	HCI_Internal_vCenter_Network HCI_Internal_mNode_Network HCI_Internal_OTS_Network Management Network			otion etwork	iSCSI-A	iSCSI-B
VLAN	Managem	ent VLANs	User VM:	s/vMotion	iSCSI	VLAN
vmkernel	vm	nk0	vn	nk3	vmk1	vmk2

For NetApp HCI compute nodes, the 1GbE interfaces eth2 and eth3 are bound to uplinks NetApp_HCI_Mgmt_vmnic2 and NetApp_HCI_Mgmt_vmnic3. These interfaces are used for VMware ESXi management traffic and the VCSA, mNode, and ONTAP Select management interfaces. The management VLAN should be native on these interfaces.

10GbE interfaces eth5 and eth0 are bound to uplinks NetApp_HCI_Virtualization_vmnic0 and NetApp_HCI_Virtualization_vmnic4 and back port groups for vMotion and VM_Network.

10GbE interfaces eth1 and eth4 are assigned to NetApp_HCI_Storage_vmnic5 and NetApp_HCI_Storage_vmnic1 respectively. These ports are configured for VMkernel port binding exclusively for storage connectivity and are configured in a manner to support it.

The following figures show the NetApp HCI compute node virtual switch configuration when using VMware distributed switches.

Figure 22) Management port groups configuration on VMware distributed switches.

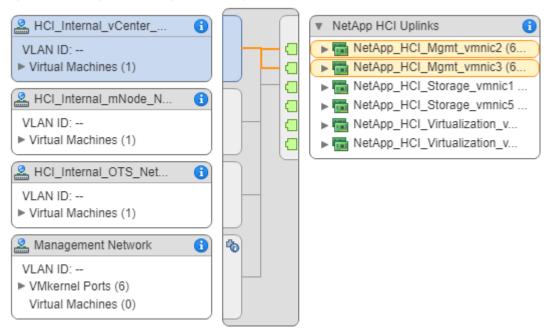


Figure 23) vMotion and VM_Network configuration on VMware distributed switches.

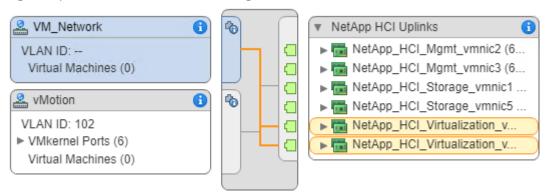


Figure 24) iSCSI-A configuration on VMware distributed switches.

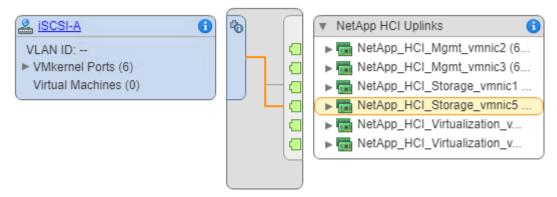
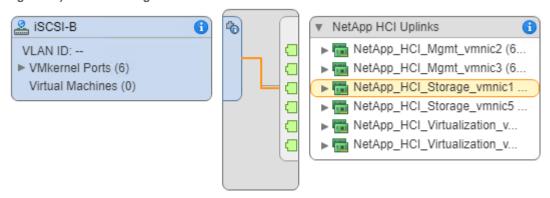


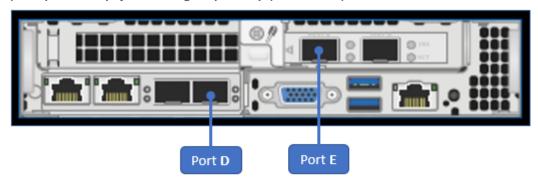
Figure 25) iSCSI-B configuration on VMware distributed switches.



Compute Node Physical-to-Logical Port Map (VDS Two-Cable)

In VDS-based deployments using a two-cable configuration, a single vSphere distributed switch is configured that contains the NetApp_HCI_Storage_vmnic1 and NetApp_HCI_Storage_vmnic5 uplinks. All management and data plane traffic for core infrastructure deployed by NDE traverses only these two uplinks.

Figure 26) Compute node physical-to-logical port map (VDS 2 cable).



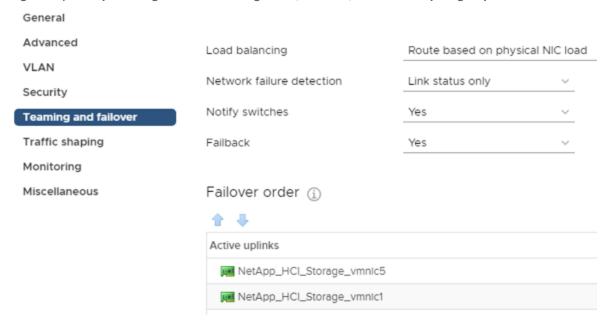
Port	D	E			
Interface	eth1	eth4			
vmnic	vmnic5	vmnic1			
vSwitch	NetApp HCI VDS				
Portgroup	HCI_Internal_vCenter_Network HCI_Internal_mNode_Network HCI_Internal_OTS_Network Management Network		otion letwork	iSCSI-A	iSCSI-B
VLAN	Management VLANs	User VM:	s/vMotion	iscsi	VLAN
vmkernel	vmk0	vn	nk3	vmk1	vmk2

There are three primary benefits to this arrangement:

- A reduction in cables for compute nodes from six to two
- Dedicated interfaces for customer workloads such as a DMZ or for software-defined networking solutions like VMware NSX
- Streamlined network configuration

All port groups except those in use for iSCSI are active/active across both uplinks. iSCSI port groups are configured with a single active uplink with the default load-balancing policy. iSCSI VMkernel port binding is also configured.

Figure 27) Example configuration for management, vMotion, and user VM port groups.



iSCSI Configuration

During the installation of ESXi, the software iSCSI IQN is extracted and automatically added to a default volume access group on the SolidFire cluster to provide access to storage. Any additional volumes created after the system configuration that are added to this default access group are visible to the ESXi hosts.

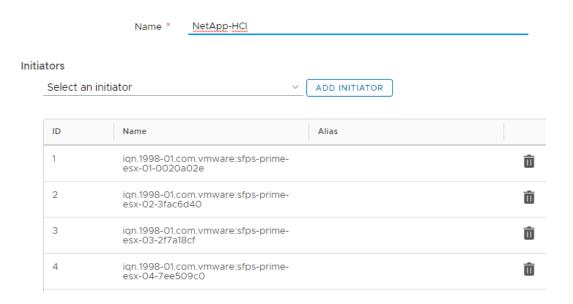
Figure 28) Host IQN: iSCSI name.



Additional ESXi hosts that are added to the HCl system should be added to the default access group NetApp-HCI to facilitate access to the existing storage volumes in use by the vCenter hosts. This process can be performed by editing the access group located under NetApp SolidFire Management > Management > Access Groups in the VCP.

Figure 29) Volume Access Group.

NetApp-HCI - Edit Access Group



Custom Storage Array Type Plug-In Rule

ESXi servers deployed by NDE have a custom storage array type plug-in (SATP) rule created for SolidFire devices. The actual rule is created with the following command:

```
esxcli storage nmp satp rule add -s VMW_SATP_DEFAULT_AA -P VMW_PSP_RR -O iops="10" -V "SolidFir" -M "SSD SAN" -e "SolidFire custom SATP rule"
```

This rule performs the following tasks:

- Claims any volumes that match a vendor name of SolidFir and model of SSD SAN
- Sets the SATP to VMW SATP DEFAULT AA
- Sets the path selection plug-in (PSP) to VMW PSP RR
- Sets the IOOperationsLimit (IOPS) to 10

This configuration creates better multipathing performance under higher queue depth situations. Any new SolidFire volumes created have this rule applied automatically. There is no need to reapply the rule.

Datastores

NDE configures two datastores that are automatically mapped to all ESXi hosts. The vCenter VCSA and mNode are both deployed to the NetApp-HCI-Datastore-01 datastore. The NetApp-HCI-Datastore-02 datastore is available for customer VMs. A local datastore is created from the leftover capacity on each ESXi host's boot device. These datastores are on local devices and are not protected by any form of RAID, so NetApp does not recommend using them for storing customer production VMs.

Figure 30) Default datastores.

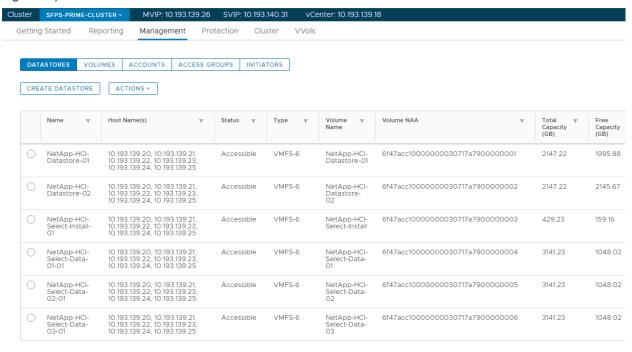
Name	~	Type v	Capacity
NetApp-HCI-Datastore-01		VMFS 6	1.95 TB
■ NetApp-HCI-Datastore-02		VMFS 6	1.95 TB
sfps-prime-esx-01-local-datastore-01		VMFS 6	216 GB
sfps-prime-esx-02-local-datastore-01		VMFS 6	216 GB
sfps-prime-esx-03-local-datastore-01		VMFS 6	216 GB
sfps-prime-esx-04-local-datastore-01		VMFS 6	216 GB
sfps-prime-esx-05-local-datastore-01		VMFS 6	216 GB

More datastores should be created as needed when additional VMs are deployed on the system. You should create a datastore when either of the two following conditions is true:

- There are more than 50 VMs per existing datastore.
- The total IOPS required in an existing datastore is greater than 15,000.

Additional datastores can be created through the VCP by navigating to NetApp SolidFire Management > Management > Datastores > Create Datastore.

Figure 31) Create datastore.



For more information about working with datastores, see the Datastore Management section of the NetApp SolidFire Plug-In for VMware vCenter Server Web Client User Guide, available on the NetApp HCI Resources page at the NetApp Support Site.

DRS and HA

Both DRS and HA are enabled with the VMware default configuration. NetApp supports customers changing the default DRS and HA rules to meet their specific requirements. Full information about configuring HA and DRS can be found at VMware Docs:

- Creating and Using vSphere HA Clusters
- Creating a DRS Cluster

vSphere Licensing

NetApp HCI comes with a 60-day, fully featured evaluation licenses for vCenter and ESXi. The customer can license the vSphere environment with any valid vSphere license. Depending on the license version applied, some automatically configured features, such as DRS or VDS, might not be available and must be unconfigured.

For more information about which features are available for each license level, see the VMware vSphere with Operations Management and VMware vSphere paper available on the VMware website.

3 Expanding NetApp HCI

The lack of ability to scale has been a consistent frustration with HCI. Nearly all early entrants into the HCI market had good solutions for point applications in which the footprint and number of workloads on the system were small. However, scaling these systems to larger configurations or mixing workloads was typically impractical for several reasons:

- Reliance on homogenous workloads. Early HCI systems worked well if the ratios of consumption for CPU, RAM, and storage were well understood and relatively constant. However, heterogeneous workloads could easily cause system hot spots.
- Cost to scale the systems. HCl systems used a core-sharing model in which compute and storage
 were included in each node. This approach led to extraneous costs when scaling one vector (CPU,
 RAM, or storage) was impossible or impractical.
- System overhead from a controller VM on every node. Sometimes this configuration led to the
 consumption of a significant amount of resources that were therefore not available for customer
 workloads.
- Nonlinear performance due to reliance on data locality.

NetApp HCI avoids these issues by employing the strength of the SolidFire scale-out architecture and by avoiding a core-sharing model:

- The combination of QoS and the SolidFire scale-out, node-based architecture enables true mixed workload capability. Every volume has a guaranteed performance level, which is further enhanced with VMware VVols.
- NetApp HCl does not use a core-sharing model. This approach enables the you to scale compute and storage resources separately, providing a finer-grained scaling model without extraneous costs.
- Compute nodes do not run a storage stack, so all hypervisor resources are available to run customer workloads.
- The SolidFire cluster dynamically spreads data over the entire cluster to eliminate hot spots.

3.1 Adding Nodes

The smallest configuration that can be deployed with NDE is two compute nodes and four storage nodes. The largest configuration is 64 compute nodes and 40 storage nodes. Usually, the installation time ranges from approximately 35 minutes for small configurations to approximately 1 hour for the largest configurations.

The expansion feature enables you to expand your NetApp HCI system by adding compute and/or storage nodes to the cluster up to the maximum configuration. You can use any combination of compute and storage nodes as long as basic requirements are met.

The workflow to expand a NetApp HCI system can be accessed by navigating to the management IP (MIP) for any of the storage nodes in the existing HCI system except for the cluster master.

Note: A list of cluster node MIPs can be found in the vCenter Plug-In under Cluster > Nodes.

Adding Nodes

Before adding new nodes to the cluster, the nodes must be loaded with the code version of the existing cluster and be on a network reachable from the existing cluster.

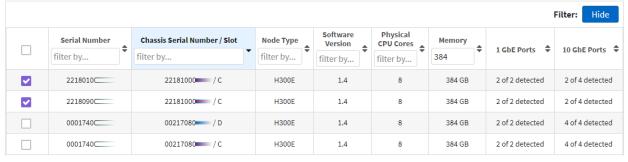
Figure 32) Adding nodes to NetApp HCI.

Available Inventory

Verify the available nodes and select at least one compute node or one storage node to add to your HCI installation.

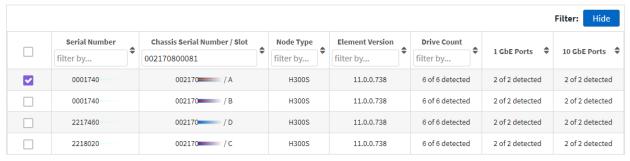
Refresh Inventory

Compute Nodes



2 compute nodes selected

Storage Nodes



1 storage nodes selected

Note: When expanding a NetApp HCI system, the new nodes must have the same cable configuration as the cluster resource you are expanding.

The number of compute nodes can be expanded from the minimum configuration to a maximum of 64 nodes.

The number of storage nodes can be expanded from the minimum to a maximum of 40 nodes. Mixing storage node types is supported as long as a single node does not occupy more than one-third of the available capacity of the cluster.

For example, consider an initial system built with four H300S (small) storage nodes, each of which has a capacity of 2.4TB. The total cluster capacity is 9.6TB (2.4TB * 4). Suppose that you want to expand the cluster with H700S storage nodes with a capacity of 9.6TB per node.

Example 1

Adding a single H700S (large) storage node with a capacity of 9.6TB results in an invalid configuration.

4x H300S and 1x H700S:

- Total cluster capacity = (2.4TB * 4) + 9.6TB = 19.2TB
- H300S = 2.4TB = ~13% of cluster capacity per node
- H700S = 9.6TB = ~50% of cluster capacity per node (not valid)

Example 2

Adding two H700S nodes results in a valid configuration.

4x H300S and 2x H700S:

- Total cluster capacity = (2.4TB * 4) + (9.6TB * 2) = 28.8TB
- H300S = 2.4TB = ~8% of cluster capacity per node
- H700S = 9.6TB = ~33% of cluster capacity per node (valid configuration)

VMware EVC and Mixed Node Types

When expanding NetApp HCI compute resources using nodes with differing CPU models, you must configure VMware Enhanced vMotion Capability (EVC) to support the live migration of machines between nodes with different CPU families.

If you are adding a previous generation of nodes (H300E, 500E, or 700E) to a cluster composed of H410C nodes, you must enable EVC before attempting to expand the vSphere cluster. However, if you are adding H410C nodes to an existing Hx00E cluster, EVC is enabled automatically during the workflow to expand the vSphere cluster.

4 Mixed Workloads

One of the core strengths of the NetApp HCI platform is its ability to run true mixed workloads in production on the same system. Mixed workloads are enabled by the QoS and scale-out features of SolidFire systems.

An application that interferes with other applications creates performance degradation, causing IT administrators to spend valuable time troubleshooting the environment. Mainstream applications, such as VDI and database applications, have unique I/O patterns that can affect one another's performance when deployed in a shared environment during normal operations. The NetApp HCI QoS feature allows fine-grained control of performance for every application, eliminating noisy neighbors, meeting unique performance needs, enabling higher utilization of infrastructure, and satisfying performance SLAs.

4.1 Test Configuration

As an example, a mixed workload test was conducted using Microsoft SQL Server, MongoDB, and VDI desktops running on a single HCI system.

Figure 33) Mixed workloads.

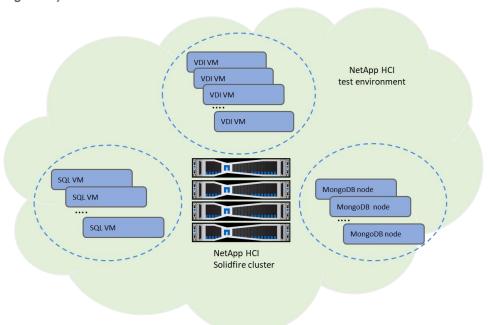


Table 5, Table 6, Table 7, and Table 8 show the VMs configured for the application workloads.

Table 5) Mongo replica set cluster VMs.

MongoDB Cluster	Qty	MongoDB Version	71	Cores/ vCPUs	RAM	os
Cluster 1	3	3.4	VMware	8	32GB	Red Hat Linux
Cluster 2	3	3.4	VMware	8	32GB	Red Hat Linux

Table 6) Yahoo! Cloud Serving Benchmark (YCSB) client-server VM.

YCSB Client	Qty	YCSB Version	7.1	Cores/ vCPUs	RAM	os
YCSB	1	0.12.0	VMware	8	32GB	Red Hat Linux

Table 7) SQL VMs.

SQL Cluster	Qty	Туре	Cores/ vCPUs	RAM	os
Hosts	5	VMware	8	16GB	Red Hat Linux

Table 8) VDI VMs.

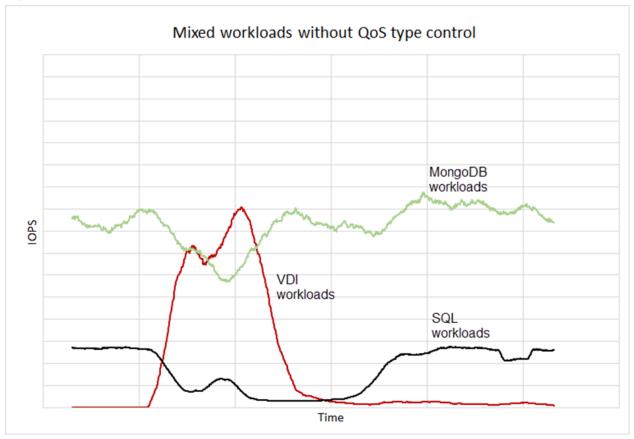
VM Type	Qty	Туре	vCPUs	RAM	os
Virtual desktops	400	VMware linked clones	2	4GB	Windows 10

4.2 Results

Results with QoS Disabled

The graph in Figure 34 shows the SQL Server and MongoDB workloads in steady state followed by a boot storm from the 400 VDI desktops.

Figure 34) VDI boot storm with QoS disabled.

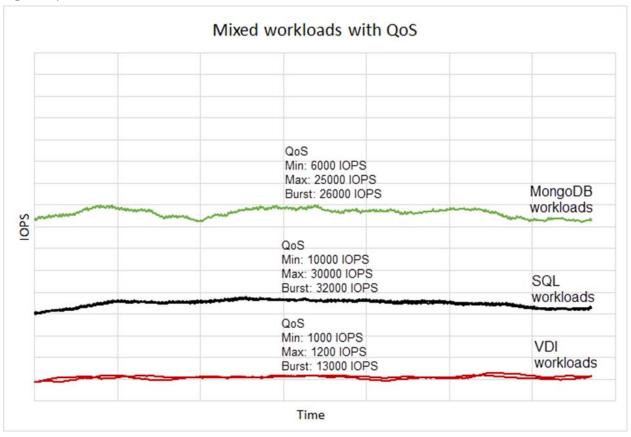


The VDI boot storm significantly affected both the SQL and MongoDB workloads. This result exemplifies the performance characteristics of most traditional HCI solutions and shows why they are not typically used for mixed workloads at scale.

Results with QoS Enabled

The graph in Figure 35 shows the same workload, this time with QoS enabled.

Figure 35) VDI boot storm with QoS enabled.



The VDI boot storm had no impact on the SQL and MongoDB workloads on the system. This example shows the strength of the QoS capabilities of the platform. QoS controls enable true mixed workload applications for NetApp HCI.

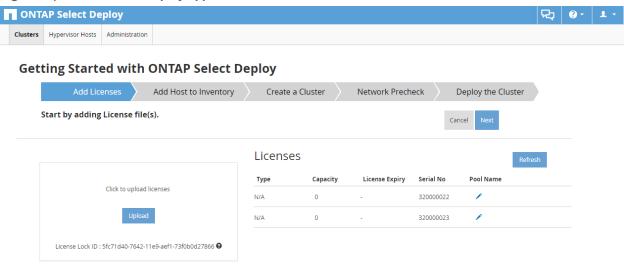
For more detailed information about the mixed workload testing, see <u>TR-4632</u>: <u>NetApp HCI QoS and Mixed Workloads</u>.

5 Extending NetApp HCI

5.1 NetApp ONTAP Select

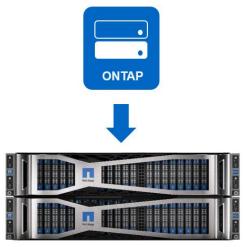
The SolidFire platform provides an enterprise-grade scalable architecture for NetApp HCI. However, SolidFire is a block-only solution. NetApp ONTAP Select 9.5 or later can be deployed as a customer-installed feature after system installation using the ONTAP Select Deploy appliance available from support.netapp.com.

Figure 36) ONTAP Select Deploy appliance.



For full information about configuring ONTAP Select for NetApp HCI systems, see <u>TR-4669: HCI File Services Powered by ONTAP.</u>

Figure 37) ONTAP Select on HCI.



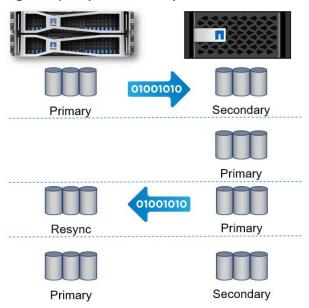
5.2 NetApp SnapMirror

The release of Element OS 10.1 and NetApp ONTAP 9.3 brought NetApp SnapMirror® support to NetApp HCI. SnapMirror relationships are managed through the Element UI at initial release.

Replication and restoration of volumes is a four-step process:

- 1. Replicate HCI volumes to a FAS system licensed for SnapMirror.
- 2. In a disaster recovery scenario, promote the FAS target volume to primary and run production.
- 3. After primary site recovery, resync the primary FAS volume back to HCI.
- 4. To resume production at the primary site, fail back the HCl volumes.

Figure 38) SnapMirror data replication.



The FAS system chosen as the replication target must be licensed for SnapMirror. The HCI system requires no additional licenses.

The 1.4 release of NetApp HCI added the following enhancements:

- You can now create SnapMirror relationships with ONTAP Cloud instances hosted in public clouds.
 This enables disaster recovery to the cloud for NetApp HCl systems.
- ONTAP LUNs can now be replicated to HCI.
- The fan-in ratio increased to up to 32 HCl systems.
- SnapMirror to ONTAP Select is now supported.

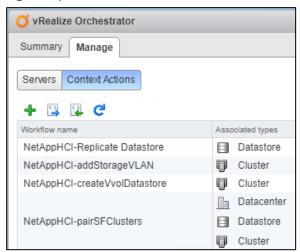
For information about configuring SnapMirror for HCI systems, see <u>TR-4641: NetApp HCI Data</u> Protection.

5.3 vRealize Orchestrator Plug-In

The SolidFire vRealize Orchestrator Plug-In enables further automation of the NetApp HCl system. The plug-in and user guide are available at the NetApp SolidFire vRealize Orchestrator Plug-In page.

To further simplify the management of the NetApp HCI platform, NetApp has also provided a few example workflows for vRealize Orchestrator (vRO) that simplify repetitive storage-related tasks.

Figure 39) vRO workflows.



These workflows can be downloaded from this source.

5.4 NetApp HCI and SolidFire vRealize Operations Management Pack

The Blue Medora management pack enhances VMware vRealize Operations (vROps) with at-a-glance KPI dashboards for navigating the entire environment to understand relationships and discover dependencies. This unique capability dramatically reduces mean time to recover (MTTR) and automates remediation. The Blue Medora out-of-the-box dashboards provide relational visibility into the following functions:

- NetApp HCI Cluster Overview
- NetApp HCI Health Investigation
- NetApp HCI Volume Overview
- NetApp HCI VVol Overview

NetApp HCl Cluster Overview Actions - All Dashboards -Cluster Overview 1.62 Deduplication Factor Compression Factor 2.08 • Overall Efficiency Factor 47.32 13.99 7.24 Space Used Space Provisioned 332.7 Capacity Utilization 0.0022 Active Block Space 318.9 Relationships Collection Status Badge|Health (%) 1 - 1 of 1 items Ш Cluster IOPS Cluster Throughput Cluster Latency 3.84 53,934.72 Read Why 127.27 Write 1,649,161.38 Write Write 132.23 Total 1,634,737.5 Total Total Storage Node Overview bluemedora-stg-03 CPU Usage Memory Used 39.38 0 bluemedora-stg-04 bluemedora-stg-02 100 m

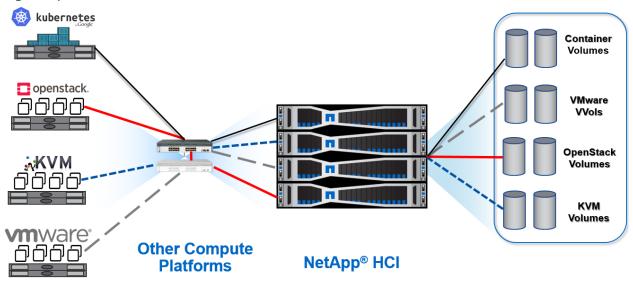
Figure 40) vRealize Operations Manager Plug-In.

The Blue Medora Plug-In is available here.

5.5 External Host Access

As a scale-out system, NetApp supports extending the storage resources of the HCI system to external hosts that are supported by the SolidFire platform. This approach allows additional consolidation of workloads onto the HCI platform.

Figure 41) External host access.



5.6 HCI Collector Project

The HCI Collector is an open-source project for graphing time series data from NetApp HCI and SolidFire systems. The HCI Collector solution uses Grafana for graphing performance data, Docker for containerizing the application stack, the Trident for Docker Plug-In for persistent storage of metrics and container state, and Graphite for storing the time-series data.

Figure 42) HCI Collector interface.



For information about deploying and using the HCl Collector, see <u>TR-4694: Visualizing NetApp HCl</u> Performance.

6 Conclusion

NetApp HCI is the embodiment of an API-driven, scale-out, multiworkload platform for the next-generation data center and hybrid multi-cloud environments. This combination enables several key capabilities, including the following:

- Making automation and orchestration first-class citizens in the data center
- Scaling the scarcest resources without overprovisioning the entire stack
- Driving true consolidation of workloads by pushing better system utilization
- Reducing go-forward capex and opex costs
- Integrating into the NetApp Data Fabric to leverage all NetApp products, increase data mobility, and reduce data silos

The solution yields significant efficiency and agility when deploying applications, helping the business to consolidate workloads with confidence.

Where to Find Additional Information

To learn more about the information described in this document, see the following documents and/or websites:

 NetApp HCI Documentation Resources http://mysupport.netapp.com/netapphci/resources

Version History

Version	Date	Document Version History		
Version 1.0	October 2017	Initial release.		
Version 1.1	December 2017	Updates for NetApp HCl v1.1		
Version 1.2	March 2018	Updates for NetApp HCI v1.2		
Version 1.3	June 2018	Updates for NetApp HCI v1.3		
Version 1.4	November 2018	Updates for NetApp HCI v1.4		
Version 1.6	July 2019	Updates for NetApp HCI v1.6 and NKS		
Version 1.7	May 2020	Updates for NKS		

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.

Copyright Information

Copyright © 2019–2020 NetApp, Inc. All rights reserved. Printed in the U.S. No part of this document covered by copyright may be reproduced in any form or by any means—graphic, electronic, or mechanical, including photocopying, recording, taping, or storage in an electronic retrieval system—without prior written permission of the copyright owner.

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

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

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

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

Data contained herein pertains to a commercial item (as defined in FAR 2.101) and is proprietary to NetApp, Inc. The U.S. Government has a non-exclusive, non-transferrable, non-sublicensable, worldwide, limited irrevocable license to use the Data only in connection with and in support of the U.S. Government contract under which the Data was delivered. Except as provided herein, the Data may not be used, disclosed, reproduced, modified, performed, or displayed without the prior written approval of NetApp, Inc. United States Government license rights for the Department of Defense are limited to those rights identified in DFARS clause 252.227-7015(b).

Trademark Information

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

WP-7261-0520

