WHITE PAPER

Hitachi Content Platform on vSAN

Best Practices Guide

By Tim Darnell

June 2018

Feedback

Hitachi Vantara welcomes your feedback. Please share your thoughts by sending an email message to SolutionLab@HitachiVantara.com. To assist the routing of this message, use the paper number in the subject and the title of this white paper in the text.

Revision History

Revision	Changes	Date
MK-SL-065-00	Initial release.	June 1, 2018

Table of Contents

Hitachi Content Platform on vSAN

Best Practices Guide

Hitachi Content Platform (HCP) is a component of the Hitachi Vantara Cloud Object Platform that allows users to archive, search, present, report on, and secure structured and unstructured data. HCP is the core object storage component which enables many of the Cloud Object Platform features and functionalities, and it can be used as a highly-available enterprise-class object storage system in a standalone environment as well.

Hyper-converged infrastructure is becoming more prevalent in today's datacenters, increasingly sitting alongside or displacing more traditional converged infrastructure designs. Appliances such as Hitachi Vantara's Unified Compute Platform Hyper-Converged (UCP HC) and Unified Compute Platform RS (UCP RS) allow customers to purchase low-cost, commodity hardware to take advantage of software-defined storage platforms such as VMware vSAN.

This paper describes the best practice recommendations for deploying HCP on VMware vSAN clusters to ensure availability, performance, and reliability of the HCP system running on a hyper-converged infrastructure.

Hitachi Vantara Cloud Object Platform Software

Hitachi provides best-of-breed storage options for private, public, and hybrid cloud services. Building upon the foundation of Hitachi Content Platform (HCP) object storage, Hitachi software can enable control and mobility of data while providing big data analytics capabilities. For more information on Hitachi Vantara Cloud Object Platform software, please visit https://www.hitachivantara.com/en-us/products/cloud-object-platform.html.

Hitachi Content Platform

Hitachi Content Platform is an object storage solution that enables IT organizations and cloud service providers to store, share, sync, protect, preserve, analyze and retrieve file data from a single system. It is more efficient, easier to use, and capable of handling much more data than traditional file storage solutions. HCP automates day-to-day IT operations like data protection and readily evolves to changes in scale, scope, applications, storage, server and cloud technologies over the life of data. In IT environments where data grows quickly or must live for years, decades or even indefinitely, these capabilities are invaluable.

Hitachi Content Platform eliminates the need for a siloed approach to storing unstructured content. The platform provides massive scale, multiple storage tiers, powerful security, Hitachi Vantara reliability, cloud capabilities, broad protocol support, multitenancy, and configurable attributes for each tenant. It can support a wide range of applications on a single cluster and is backed by a thriving community of third-party software partners. With access to a robust ecosystem of cloud applications, Hitachi Content Platform can solve a wide range of current problems and adapt to meet future needs.

Hitachi Unified Compute Platform Appliances

Hitachi provides appliance-based platforms for virtualization within the enterprise, from simple hyper-converged systems to fully managed and automated scale-out converged systems. For more information on Hitachi Unified Compute Platform, please visit https://www.hitachivantara.com/en-us/products/converged-systems.html.

Hitachi Unified Compute Platform Hyper-Converged

Hitachi Unified Compute Platform HC (UCP HC) family of hybrid and all-flash models combines compute, storage and virtualization into a hyper-converged infrastructure appliance. UCP HC appliances create simple, easy-to-deploy, all-in-one solutions powered by software-defined storage (VMware vSAN) and Hitachi software to extend the agility and simplicity of the UCP family.

UCP HC family provides a comprehensive dashboard to view virtual machines. Compute and storage are viewable, with health monitors for CPU, memory, storage and virtual machine usage for entire clusters, individual appliances and individual nodes. Minimal IT experience is required to deploy, configure, and manage the UCP HC appliances. As UCP HC leverages VMware's core products, administrators can apply existing VMware knowledge, best practices, and processes.

Ensure powerful performance and availability for traditional and cloud-native application deployment. UCP HC solutions deliver a reliable platform for business-critical applications, databases, virtual desktop infrastructure (VDI), DevOps, containers, and remote and branch office (ROBO) deployments, among other use cases.

Hitachi Unified Compute Platform Rack Scale

Hitachi Unified Compute Platform RS (UCP RS) is a turnkey solution completely supported End-to-End by Hitachi Vantara and powered by VMware Cloud Foundation. As a hyper-converged infrastructure (HCI), UCP RS enables the convergence of compute, storage, networking, and management onto industry-standard x86 servers, providing a building-block approach with scale-out capabilities. By taking advantage of UCP RS, all main datacenter functions run as software on the VMware hypervisor in an integrated management software layer. In today's modern datacenters, HCI is a stepping stone to the Software-Defined Data Center (SDDC), but lifecycle management and automation still need to be tightly coupled.

SDDC Manager, a component of VMware Cloud Foundation, automates the entire provisioning process for ESXi, vCenter, vSAN, and NSX on top of Hitachi UCP RS. VMware Cloud Foundation enables deployment of a private cloud environment based on VMware's Software Defined Data Center architecture. A VMware Cloud Foundation installation is a turnkey private cloud instance that is easily deployed in a corporate network. Within VMware Could Foundation, SDDC Manager enables the ability for streamlined and automated data center operations and delivery of service offerings, such as virtual infrastructure (VI) and virtual desktop infrastructure (VDI) environments, based on a VMware SDDC architecture on top of Hitachi UCP RS.

Hitachi UCP RS is the answer for today's IT demands to deliver agile, flexible, and reliable services based on organization requests, and is a cornerstone in a new and modern datacenter that provides an easy and fast way to deploy SDDC.

Solution Components

This section lists the hardware and software components used during development of the best practices documented in this guide. While a Hitachi V240 HC appliance was used, the best practices in this guide may be applied to any VMware vSAN ReadyNode unless explicitly noted.

TABLE 1. TESTED HARDWARE SOLUTION COMPONENTS

Hardware	Description	Version	Quantity
UCP HC vSAN ReadyNode	HY-4 Series VMware vSAN ReadyNode 192 GB Memory per node Intel Xeon E5-2620 v3 CPUs 1 × Intel 400 GB SSD per node (Flash Tier) 3 × 1.2 TB Seagate HDDs per node (Capacity Tier) 1 × 300 GB Hitachi HDD per node (Boot Device) 1 × Dual-port 10 GigE Base-T Intel X540 OCP Mezzanine Card	V240	4 Nodes

TABLE 2. SOFTWARE SOLUTION COMPONENTS

Software	Description	Version
Hitachi Content Platform	Object-based storage platform software	8.0.0.9
VMware vSphere	Virtualization platform	6.5 Update 1
VMware vSAN	Software defined storage platform for virtual machines	6.6.1

Best Practice Recommendations

Many of the necessary best practices for an available, performant, and reliable virtualized HCP infrastructure are similar to generic VMware vSphere best practices for a hyper-converged infrastructure. The sections below describe recommended configurations for vSphere, vSAN, and HCP.

Standard vSphere Best Practices

Both vSAN and HCP are distributed object storage systems, therefore time synchronization is critical. Choose reliable, physical time sources to use for all vSphere-based components (vCenter, PSC, ESXi, etc.) and use the same time sources when configuring HCP for NTP services.

DNS services are also critical to reliable vSphere and HCP functionality. Ensure you are using fully qualified domain names (FQDN) throughout your vSphere environment with redundant DNS servers, and consider using the same DNS servers to host the stub zones for HCP resolution in a multi-node HCP cluster deployment.

vSAN Storage Recommendations

vSAN is a flexible software-defined storage solution with many configuration options that can be used to modify the basic behaviors and configurations of vSAN. The recommendations below may help increase HCP performance in certain situations and use cases.

Multiple Disk Groups

For Hybrid VMware vSAN ReadyNodes that contain high numbers of drive bays, consider using multiple disk groups to increase overall vSAN performance and distribute physical disk fault domains. For example, newer models of Hitachi Vantara Skylake-based UCP HC VMware vSAN ReadyNodes have up to twelve drive bays – for example, three disk groups could be created across these drives with one flash device and three capacity devices per disk group. Consider using multiple disk groups on the underlying vSAN for HCP to reduce performance impact during vSAN resync operations and to increase the overall storage throughput and capacity of the vSAN datastore.

VM Storage Policies – Flash Read Cache Reservation

In hybrid vSAN configurations, the flash tier may be configured as a flash-based read cache for blocks that are commonly accessed. Within a VM Storage Policy, you can configure the “Flash read cache reservation percentage” and assign the policy to HCP VM VMDKs used for object storage to enable this feature. Consider implementing the flash read cache feature on HCP VM virtual disks that may host objects that are commonly read by applications that utilize HCP object storage.

VMDK Disk Type

Since HCP is an object-based storage system, thick provisioned VMDKs are recommended for all HCP VM VMDKs that will be used for object storage. Ensure that if additional VMDKs are provisioned to an HCP node that they are created using this provisioning method in vSphere.

Network Best Practices

Since vSAN relies on the network for presentation and replication of storage, network configuration is important to ensure availability, performance, and reliability of the vSAN system and the virtual machines running on top of it. The sections below describe recommendations to consider for configuring the network on the vSAN cluster hosting the HCP infrastructure to ensure predictable behavior of the vSphere and HCP infrastructures.

Physical Network Adapters and VLANs

A minimum of two 10 GbE ports should be used on the VMware ESXi hosts comprising the vSAN cluster. If necessary, more 10 GbE or higher speed ports can be added to each host (preferably in pairs) for additional bandwidth. Cable each adapter port to separate switching infrastructure to ensure network connectivity in the case of a switch failure. Utilize VLAN tagging to isolate typical vSphere guest and host network traffic onto separate VLANs to support virtual distributed switch port group segregation inside the vSphere infrastructure.

Virtual Distributed Switches

Utilize virtual distributed switches (vDS) throughout the vSphere infrastructure instead of virtual standard switches (VSS). Using a vDS allows you to easily configure cluster-wide failover policies to network adapters to guarantee isolation for traffic such as vSAN and HCP. Configuration is simplified by ensuring consistent network configuration across the ESXi nodes in a cluster for fewer errors and misconfigurations of ESXi networking.

Distributed Port Groups and Teaming/Failover

Utilize distributed port groups within a vDS to map VLANs or traffic types to a specific group of ports on the vDS. At a minimum, typical deployments will have separate and dedicated VLANs for vSphere Management, vMotion, vSAN, and virtual machine (guest) networks.

By creating dedicated distributed port groups and configuring teaming/failover policies for typical vSphere traffic types, you can allow specific traffic types to traverse specific physical network adapters. This is important with vSAN-based infrastructures, as VMware recommends that a dedicated 10 GbE adapter be used for vSAN traffic. Table 3 shows an example of a 2-port 10 GbE configuration that could be used when running HCP on vSAN.

TABLE 3. EXAMPLE OF 2-PORT 10 GBE DISTRIBUTED PORT GROUP CONFIGURATION

Traffic Type	VLAN ID	Teaming Policy	Active Uplinks	Standby Uplinks
vSphere Management	10	Active/Standby	Uplink 1	Uplink 2
vMotion	20	Active/Standby	Uplink 1	Uplink 2
vSAN	30	Active/Standby	Uplink 2	Uplink 1
HCP Frontend	100	Active/Standby	Uplink 1	Uplink 2
HCP Backend	101	Active/Standby	Uplink 1	Uplink 2
VM Network	102	Active/Standby	Uplink 1	Uplink 2

If sharing network adapters with other virtual machines and management traffic is not desired for HCP frontend and backend traffic, four 10 GbE NICs per ESXi host could be used and two of them dedicated and teamed for HCP-only traffic. Table 4 shows an example of a 4-port 10 GbE configuration that could be used when running HCP on vSAN.

TABLE 4. EXAMPLE OF 4-PORT 10 GBE DISTRIBUTED PORT GROUP CONFIGURATION

Traffic Type	VLAN ID	Teaming Policy	Active Uplinks	Standby Uplinks
vSphere Management	10	Active/Standby	Uplink 1	Uplink 2
vMotion	20	Active/Standby	Uplink 1	Uplink 2
vSAN	30	Active/Standby	Uplink 2	Uplink 1
HCP Frontend	100	Active/Standby	Uplink 3	Uplink 4
HCP Backend	101	Active/Standby	Uplink 3	Uplink 4
VM Network	102	Active/Standby	Uplink 1	Uplink 2

The previous configurations are only examples, yet they illustrate the flexibility you have in dedicating physical adapters to specific traffic types. While any traffic type can be dedicated to a physical adapter, the only hard requirement is that vSAN traffic be dedicated to a minimum of a single 10 GbE network adapter in an optimal (i.e., no active failure) scenario.

Network I/O Control

For even higher resiliency and greater control of network bandwidth resources, consider implementing VMware Network I/O Control (NIOC). NIOC allows for bandwidth reservations at all times for specific traffic types and implements a share-based system for guaranteeing bandwidth during times that the network adapter may be saturated (typically during adapter failover events or vSAN resync operations).

With version 3 of NIOC, individual virtual machines can now be allocated bandwidth guarantees on the vDS. This can be especially useful for HCP clusters that may be in a replication topology and need specific bandwidth requirements to stay synchronized.

NSX

VMware Cloud Foundation appliances such as UCP RS utilize NSX as part of the built-in SDDC infrastructure and already provide the capability for network configurations described previously. NSX allows for the addition of distributed firewalling, policy-based control of services, L4-L7 load-balancing, and more. Consider implementation of NSX if you require strict isolation of the HCP infrastructure without the cost or effort of adding physical firewalls and routing infrastructure.

HCP Network Requirements

Create two dedicated VLAN-backed distributed port groups - one for the HCP frontend network and one for the HCP backend network. When deploying the HCP VM nodes, ensure that the proper distributed port group is assigned to the respective frontend and backend virtual network adapter for the HCP node.

Use the information in the previous sections to determine the necessary network configuration measures for your environment and the expected usage of HCP. Take into consideration HCP replication topologies, location of the application(s) consuming HCP services, and co-resident virtual machine network requirements that will reside on the same cluster when designing the network configurations to support HCP on vSAN-based infrastructures.

Node Placement and Data Protection Level Best Practices

HCP is a Redundant Array of Independent Nodes (RAIN) based system, so it is important to ensure that fault domains are identified and designed to, and data protection levels understood. These can impact data availability during an outage as well as HCP cluster capacity and vSAN storage consumption. Both vSAN and HCP are object storage systems, and each has its own data protection level configuration capabilities. The sections below describe recommendations for node placement and data protection levels to consider when running HCP on vSAN.

vMotion

Live vMotion of HCP nodes is currently unsupported. Powered-off vMotion and Storage vMotion are currently acceptable methods of moving HCP VM nodes between ESXi hosts or underlying vSphere storage.

Single HCP Node Deployment

Note - Single HCP Node deployments should be used for demonstration and non-production uses only. Use a multi-node HCP cluster deployment combined with appropriate data protection levels within HCP and vSAN to ensure availability of metadata and objects within the HCP system.

With a single HCP node deployment, consider disabling DRS automation for the virtual machine by creating a VM Override and setting the automation level for the virtual machine to Disabled. Note that certain operations, such as placing the ESXi host into maintenance mode when the HCP node is running on it, could be impacted and may require manual intervention by an administrator.

With a single node, HCP Data Protection Level (DPL) does not need to be addressed since the node is running as a standalone entity and is not part of a multi-node HCP cluster. The vSAN Default Storage Policy setting of Failures To Tolerate of one (FTT=1) is an acceptable level of protection for a single HCP node since VMDK data will be written to two hosts within the vSAN cluster.

Consideration of physical fault domains with a single HCP node deployment is limited to the ESXi host on which the HCP node resides. If this ESXi host becomes unavailable, so will the HCP VM node, unless it is re-registered on another host and powered back on. Consider setting up a secondary HCP node on another ESXi host in a separate vSAN cluster and configure HCP replication between the two for increased availability.

Multi-Node HCP Cluster Deployment

With a multi-node HCP cluster deployment, no more than one node in the HCP cluster should reside on a single ESXi host. For example, in a four-node HCP cluster, four ESXi hosts would be required in the vSAN cluster and one HCP node assigned to each ESXi host. Use VM/Host Groups along with VM/Host Rules that use the “Must run on hosts in group” option to ensure the HCP nodes only run on their respective ESXi hosts.

In the example above with a four-node HCP cluster, DPL could be configured from one to four. Using the vSAN Default Storage Policy setting of FTT=1, the following can be observed:

- DPL=1 would result in one copy of data and metadata stored in the HCP cluster. If an HCP node becomes unavailable, all of the objects on that node would become unavailable until the node returns to service. HCP capacity is at the highest possible level since only one copy of the data and metadata is stored. vSAN consumption is twice the VMDK size since vSAN FTT=1. If availability is required with DPL=1, replication to another HCP cluster is recommended to ensure object and metadata availability in the case of an HCP node failure.

- DPL=2 through DPL=4 would result in two through four copies of data and metadata to be stored in the HCP cluster respectively. This means that if DPL=n, n-1 HCP nodes could become unavailable before objects or metadata become unavailable. HCP capacity is reduced for each increase in DPL since multiple copies of the data and metadata are stored. vSAN consumption is twice the VMDK size since vSAN FTT=1.

Custom vSAN Storage Policies with increasing FTT levels could be created and applied to only the HCP nodes to increase physical fault domain resiliency. For example, a four-node HCP cluster could have a DPL=4 setting, and a custom vSAN Storage Policy created with FTT=3 and be applied to the HCP nodes. This would result in HCP data still being available even with three ESXi host failures since HCP data and metadata are written to each HCP node, and the custom vSAN Storage Policy is configured to write copies of the VMDK to every node in the four-node vSAN cluster. This is an extreme example, as setting DPL=4 and FTT=3 in a four-node HCP and vSAN cluster would result in maximum availability and resiliency, yet a dramatic capacity reduction within HCP and highest possible storage consumption on vSAN. Table 5 shows the difference in HCP capacity and vSAN storage consumption by a four-node cluster of HCP VMs running on a four-node vSAN cluster with different FTT and DPL levels.

TABLE 5. HCP CAPACITY, VSAN STORAGE CONSUMPTION, AND RESILIENCY BASED ON PROTECTION LEVELS

FTT/DPL Settings	Total HCP Capacity Consumed (10,000 1 kb objects)	vSAN Storage Consumption per HCP node (1 TB VMDK)	Resiliency
FTT=1 DPL=1 (Default vSAN and HCP Settings)	39.06 MB	2 TB	No single-fault resiliency at HCP object level, single-fault resiliency at vSAN layer
FTT=1 DPL=2	78.12 MB	2 TB	Full single-fault resiliency at HCP and vSAN layers
FTT=2 DPL=3	117.19 MB	3 TB	Full double-fault resiliency at HCP and vSAN layers
FTT=3 DPL=4	156.25 MB	4 TB	Full triple-fault resiliency at HCP and vSAN layers

Note that using the default levels for FTT and DPL will not result in full single-fault resiliency throughout the system. To resolve this, consider deploying another HCP instance or cluster for replication on a separate vSAN cluster, or increase the default DPL level in HCP to have a FTT=1/DPL=2 configuration. HCP object ingest performance will reduce with each increase in DPL due to multiple copies of the same object being written; however, the performance decay for ingest is not exponential.

Be sure to take into consideration failure domains with respect to power and physical network when selecting ESXi hosts to place the HCP nodes onto, especially in multi-rack deployment scenarios.

Additional Recommendations

While not required, the recommendations below may help increase HCP performance and/or resiliency when running on vSAN storage if needed.

Virtual Machine Sizing

HCP virtual machines are typically deployed in either a "standard" or "small" configuration. For more information on CPU, memory, and virtual disk sizing recommendations for these configurations, please reference "Deploying an HCP-VM System on ESXi".

AF-Series vSAN ReadyNodes

Use of all-flash VMware vSAN ReadyNodes can increase storage performance considerably. Consider using an AF-series VMware vSAN ReadyNode based cluster if high-performance, ultra-low latency storage infrastructure is necessary.

For More Information

Hitachi Vantara Global Services offers experienced storage consultants, proven methodologies and a comprehensive services portfolio to assist you in implementing Hitachi products and solutions in your environment. For more information, see the Services website.

Demonstrations and other resources are available for many Hitachi products. To schedule a live demonstration, contact a sales representative or partner. To view on-line informational resources, see the Resources website.

Hitachi Academy is your education destination to acquire valuable knowledge and skills on Hitachi products and solutions. Our Hitachi Certified Professional program establishes your credibility and increases your value in the IT marketplace. For more information, see the Hitachi Vantana Training and Certification website.

For more information about Hitachi products and services, contact your sales representative, partner, or visit the Hitachi Vantara website.

Hitachi Vantara

Corporate Headquarters
2845 Lafayette Street
Santa Clara, CA 96050-2639 USA
www.HitachiVantara.com | community.HitachiVantara.com

Regional Contact Information
Americas: +1 408 970 1000 or info@hitachivantara.com
Europe, Middle East and Africa: +44 (0) 1753 618000 or info.emea@hitachivantara.com
Asia Pacific: +852 3189 7900 or APACMarketing@hitachivantara.com

MK-SL-065-00 June 2018.