vSAN 6.7 Update 1 Technical Overview

First Published On: 08-22-2018
Last Updated On: 03-15-2019

Table Of Contents

1. Introduction

1.1 Introduction

We are excited to introduce you to our eighth generation of vSAN, version 6.7 Update 1 (vSAN 6.7 U1), which makes it even easier to adopt HCI and build a Digital Foundation.

VMware HCI supports six important hybrid cloud use cases today:

- Backup and Disaster Recovery
- Test & Dev
- Multicloud workloads
- App Development
- Multicloud management
- Mode 1 applications.

VMware HCI is uniquely able to support these use cases due to our mature, Full Stack HCI. All HCI vendors offer storage virtualization, a few offer compute and network virtualization as well; however, many of these solutions are not mature enough to be hybrid cloud ready. VMware provides a feature-rich full stack, composed of vSphere, vSAN, NSX and the vRealize Suite – the components of the VMware Digital Foundation.

First, we are simplifying operations day one and two operations by streamlining the deployment process, improving lifecycle management, reducing disruptions during maintenance operations, and improving capacity reporting. These updates help administrators more quickly and easily deploy and extend infrastructure while minimizing disruptions while keeping the environment up to date.

Next, vSAN is an even more efficient infrastructure choice. The ability to automatically reclaim capacity that is no longer used, reduces the capacity needed for popular workloads without administrator interaction. More efficient networking is now available with Stretched Cluster and 2 Node deployments that have independent networking requirements for Witness and Data sites.

Updates to the HCI Assessment and vSAN Sizer tools work even better together to provide a more streamlined and flexible sizing and infrastructure selection process to ensure the most efficient configuration for new deployments.

Faster resolution, quicker diagnosis, and simplified self-help make vSAN Supportability even better. vSAN ReadyCare has simplified the support process by reducing requirements of customers, and speeding time to resolution through faster insight and integrated self-help in vSAN.

Native health checks, more self-help tools, better reactive support with vSAN Support Insight, and participation in the VMware Customer Experience Improvement program, provide for an overall enhanced support experience. VMware can rapidly understand a customer's environment, perform root cause analysis to identify the cause of the problem, and deliver proactive support based on trends and analytics.

vSAN Runs on standard x86 servers from more than 15 OEMs. Deployment options include over 500 vSAN ReadyNode™ choices, integrated systems such as Dell EMC VxRail or, Dell EMC VxRack SDDC systems, and build-your-own using validated hardware on the VMware Compatibility List. A great fit for large and small deployments with options ranging from a 2-node cluster for small implementations to multiple clusters each with as many as 64 nodes—all centrally managed by vCenter Server.

2. Architecture

2.1 Servers with Local Storage

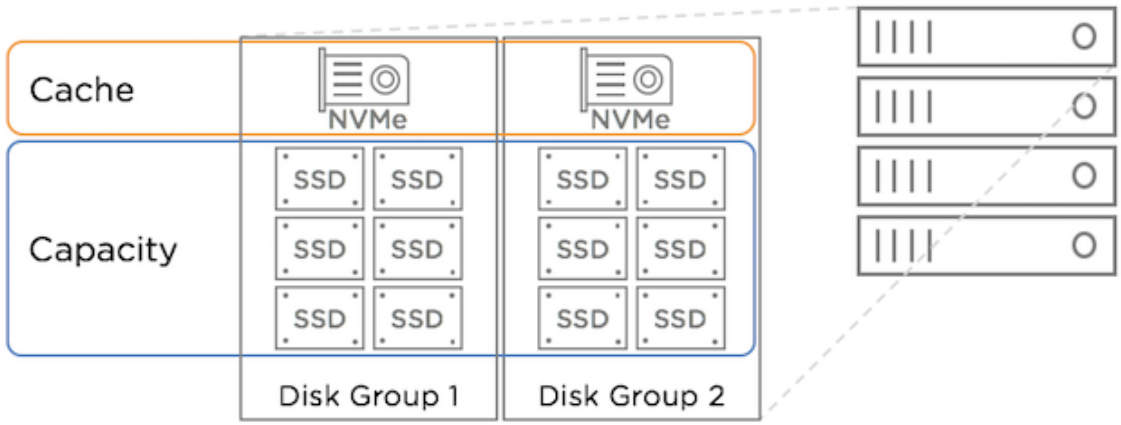
A wide variety of deployment and configuration options make vSAN a flexible and highly scalable HCI storage solution. A single vSAN cluster consists of any number of physical server hosts from two to 64.

Servers with Local Storage

Organizations can start with what is needed today and implement a “just-in-time” provisioning model for additional compute and storage capacity. Additional hosts can easily be added to an existing cluster in a matter of minutes. This method of purchasing, expanding, and refreshing an IT infrastructure is more efficient and less costly than provisioning large, monolithic “blocks” of capacity every few years.

Each host contains flash devices (all flash configuration) or a combination of magnetic disks and flash devices (hybrid configuration) that contribute cache and capacity to the vSAN distributed datastore.

Each host has one to five disk groups. Each disk group contains one cache device and one to seven capacity devices.

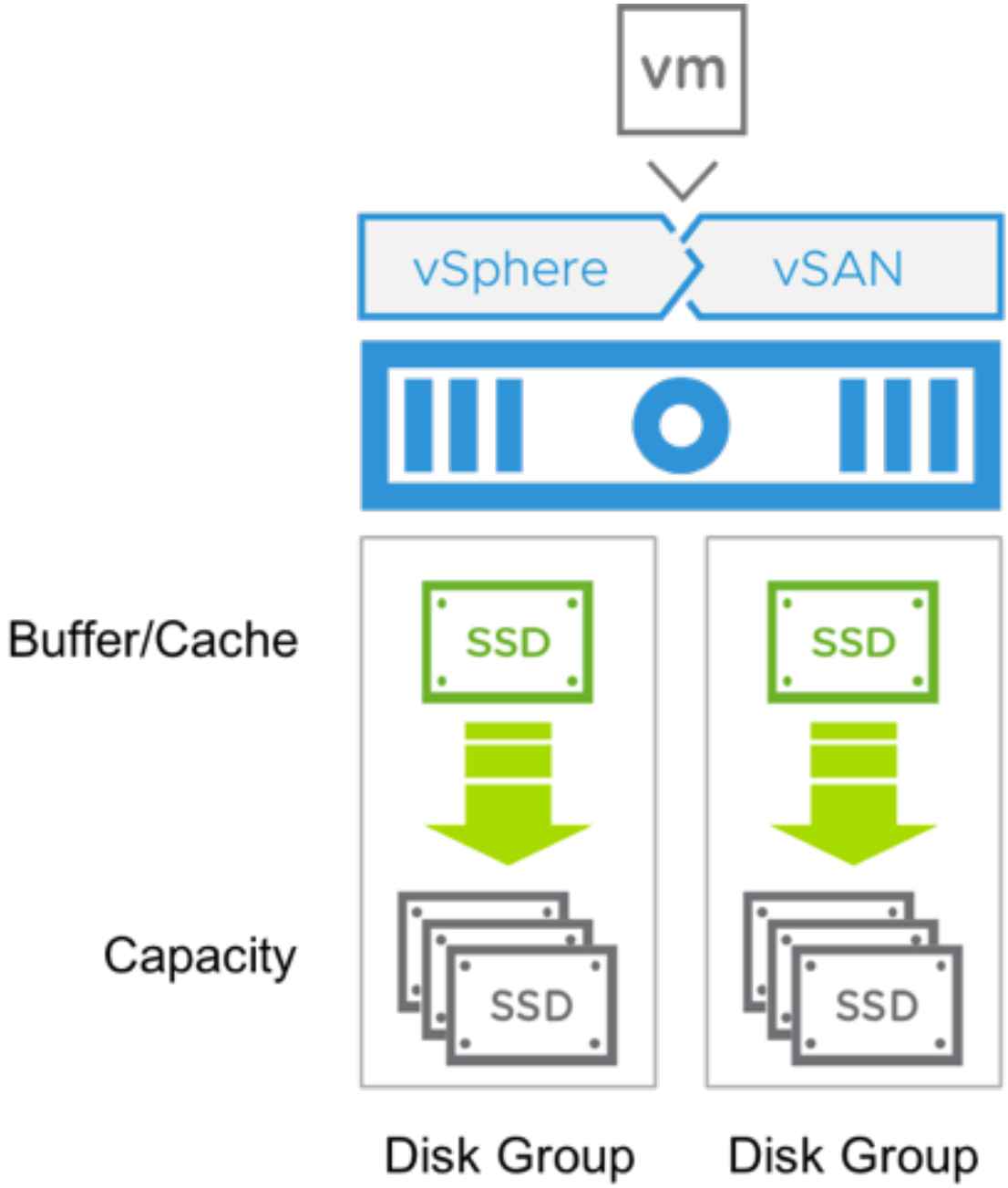


In all flash configurations, the flash devices in the cache tier are used for buffering writes. There is no need for read cache as performance from the capacity flash devices is more than sufficient. Two grades of flash devices are commonly used in an all flash vSAN configuration: Lower capacity, higher endurance devices for the cache layer and more cost effective, higher capacity, lower endurance devices for the capacity layer. Writes are performed at the cache layer and then de-staged to the capacity layer, as needed. This helps maintain performance while extending the usable life of the lower endurance flash devices in the capacity layer.

In hybrid configurations, one flash device and one or more magnetic drives are configured as a disk group. A disk group can have up to seven drives for capacity. One or more disk groups are used in a vSphere host depending on the number of flash devices and magnetic drives contained in the host.

Flash devices serve as read cache and write buffer for the vSAN datastore while magnetic drives make up the capacity of the datastore.

vSAN uses 70% of the flash capacity as read cache and 30% as write cache.



VMware is always looking for ways to not only improve the performance of vSAN but improve the consistency of its performance so that applications can meet their service level requirements. vSAN 6.7 continues this drive for better performance and consistency through optimizations made in the data path.

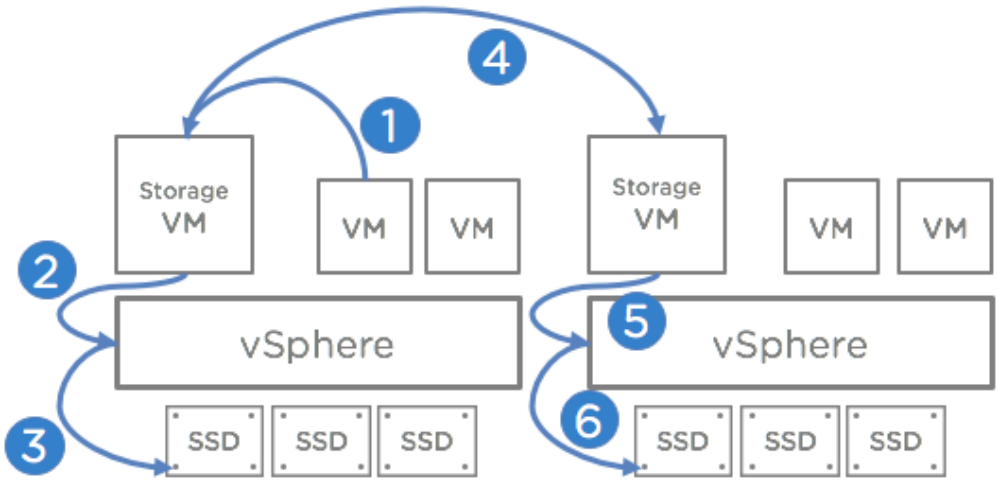
vSAN 6.7 optimizes the destaging mechanism, resulting in data that drains more quickly from the write buffer, to the capacity tier. The ability to destage this data more quickly allows for the buffer tier to be available to accept new incoming I/O's, which will reduce periods of congestion.

Storage Controller Virtual Appliance Disadvantages

Storage in a HyperConverged Infrastructure (HCI) requires computing resources that have been traditionally offloaded to dedicated storage arrays. Nearly all other HCI solutions require the deployment of storage virtual appliances to some or all hosts in the cluster. These appliances provide storage services to each host. Storage virtual appliances typically require dedicated CPU and/or memory to avoid resource contention with other virtual machines.

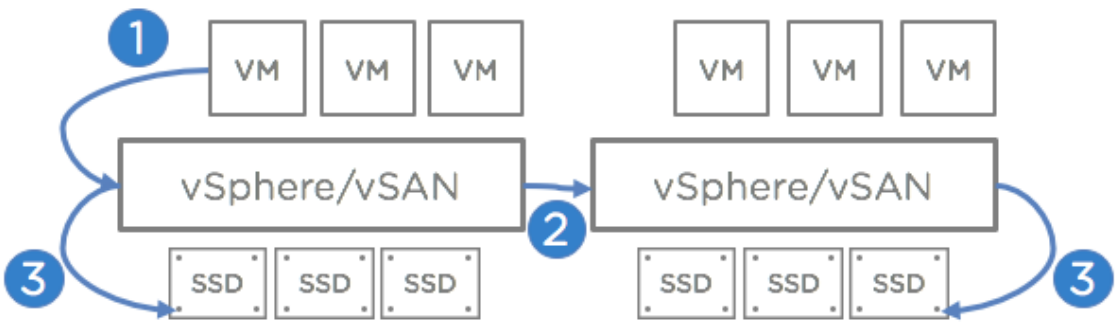
Running a storage virtual appliance on every host in the cluster reduces the overall amount of computing resources available to run regular virtual machine workloads. Consolidation ratios are lower and total cost of ownership rises when these storage virtual appliances are present and competing for the same resources as regular virtual machine workloads.

Storage virtual appliances can also introduce additional latency, which negatively affects performance. This is due to the number of steps required to handle and replicate write operations as shown in the figure below.



vSAN is Native in the vSphere Hypervisor

vSAN does not require the deployment of storage virtual appliances or the installation of a vSphere Installation Bundle (VIB) on every host in the cluster. vSAN is native in the vSphere hypervisor and typically consumes less than 10% of the computing resources on each host. vSAN does not compete with other virtual machines for resources and the I/O path is shorter.

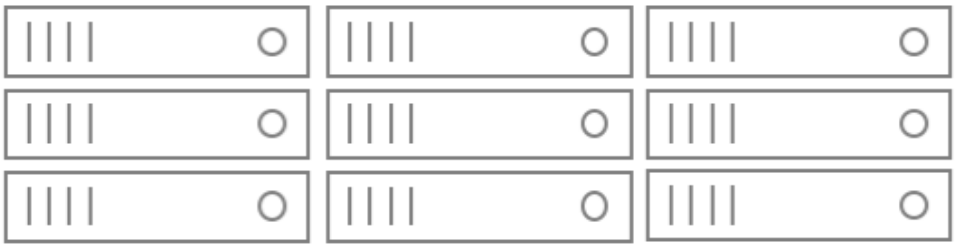


A shorter I/O path and the absence of resource-intensive storage virtual appliances enables vSAN to provide excellent performance with minimal overhead. Higher virtual machine consolidation ratios translate into lower total costs of ownership.

2.2 Cluster Types

Standard Cluster

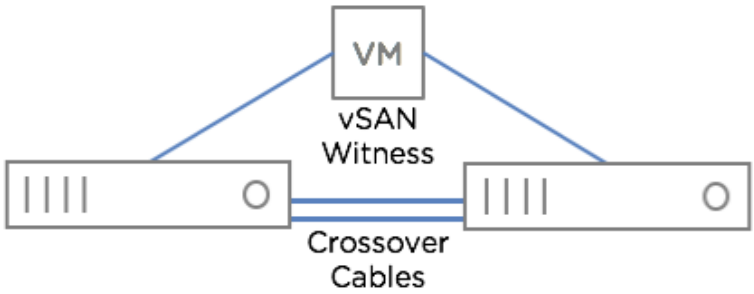
A standard vSAN cluster consists of a minimum of three physical nodes and can be scaled to 64 nodes. All the hosts in a standard cluster are commonly located at a single location and are well-connected on the same Layer-2 network. 10Gb network connections are required for all-flash configurations and highly recommended for hybrid configurations.



2 Node Cluster

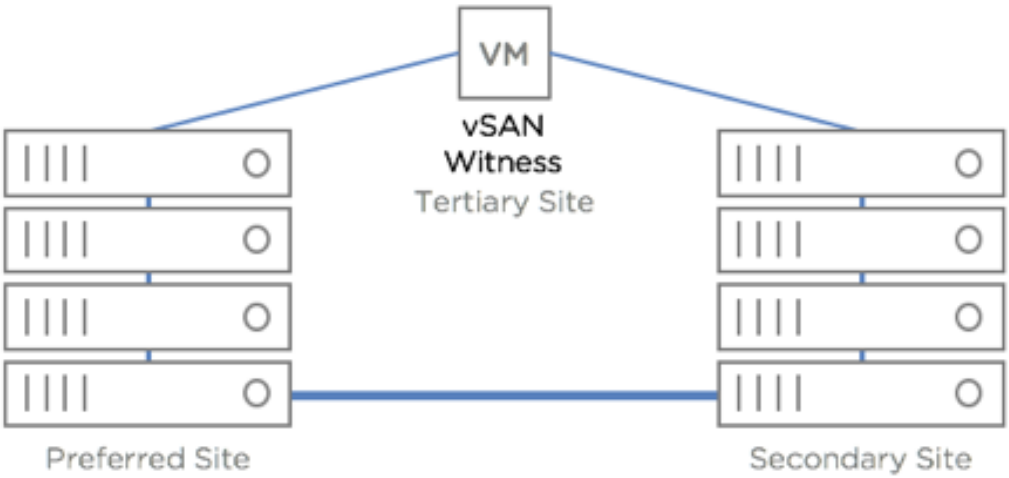
A 2-node cluster consists of two physical nodes in the same location. These hosts are usually connected to the same network switch or are directly connected. Direct connections between hosts eliminate the need to procure and manage an expensive network switch for a 2-node cluster, which lowers costs—especially in scenarios such as remote office deployments. *While 10Gbps connections may be directly connected, 1Gbps connections will require a crossover cable.

A third “vSAN Witness Host” is required for a 2-node configuration to avoid “split-brain” issues when network connectivity is lost between the two physical nodes. We will discuss the vSAN Witness Host in more detail shortly.



Stretched Cluster

A vSAN Stretched Cluster provides resiliency against the loss of an entire site. The hosts in a Stretched Cluster are distributed evenly across two sites. The two sites are well-connected from a network perspective with a round trip time (RTT) latency of no more than five milliseconds (5ms). A vSAN Witness Host is placed at a third site to avoid “split-brain” issues if connectivity is lost between the two Stretched Cluster sites. A vSAN Stretched Cluster may have a maximum of 30 hosts in the cluster and can be distributed proportionally or disproportionately. In cases where there is a need for more hosts across sites, additional vSAN Stretched Clusters may be used.



vSAN Witness Host

While not a cluster type, it is important to understand the use of a vSAN Witness Host in 2 Node and Stretched Cluster vSAN deployments. This "Witness" stores metadata commonly called "witness components" for vSAN objects. Virtual machine data such as virtual disks and virtual machine configuration files are not stored on the vSAN Witness Host. The purpose of the vSAN Witness Host is to serve as a "tie-breaker" in cases where sites are network isolated or disconnected.

A vSAN Witness Host may be a physical vSphere host, or a VMware provided virtual appliance, which can be easily deployed from an OVA. When using a physical host as a vSAN Witness Host, additional licensing is required, and the host must meet some general configuration requirements. When using a vSAN Witness Appliance as the vSAN Witness Host, it can easily reside on other/existing vSphere infrastructure, with no additional need for licensing.

When using 2 Node clusters for deployments such as remote office branch office (ROBO) locations, it is a common practice for vSAN Witness Appliances to reside at a primary datacenter. They may be run at the same ROBO site but would require additional infrastructure at the ROBO site.

vSAN Witness Hosts providing quorum for Stretched Clusters may only be located in a tertiary site that is independent of the Preferred and Secondary Stretched Cluster sites.

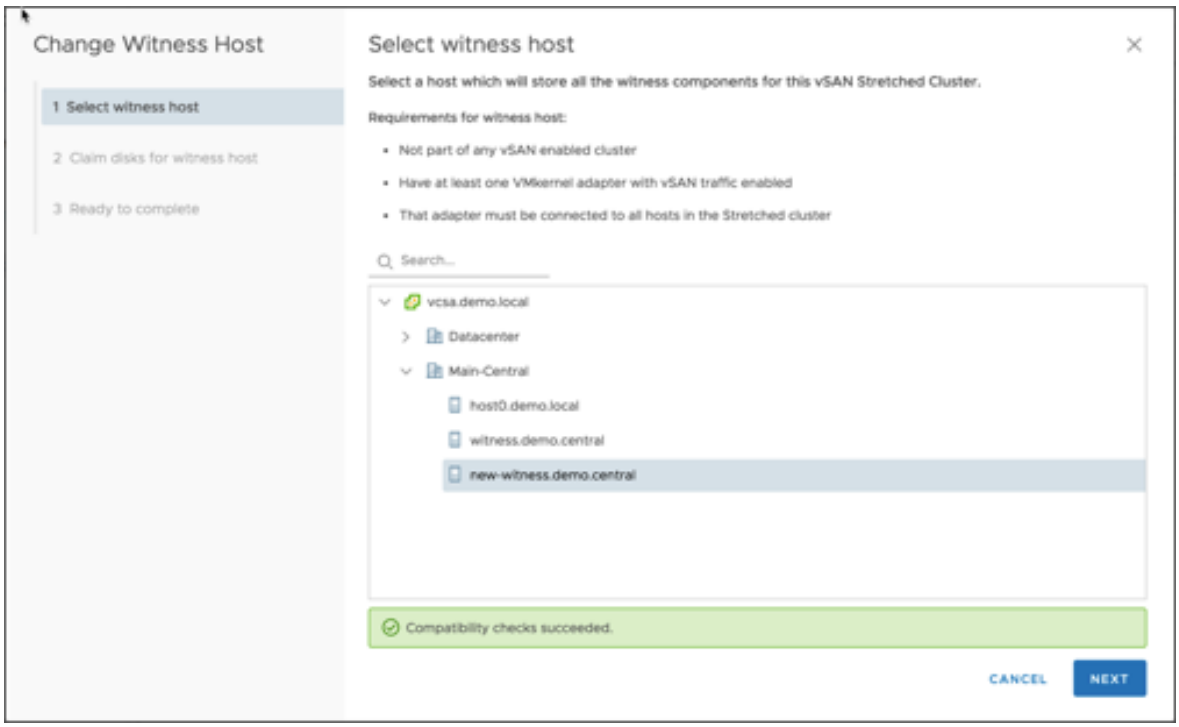
One vSAN Witness Host is required for each 2 Node or Stretched Cluster vSAN deployment.

Bandwidth requirements to the vSAN Witness Host are determined by the number of vSAN components on a cluster. During failover scenarios, ownership of vSAN components must be moved to the surviving site over a five second (5s) window. The rule of thumb is 2Mbps for every 1000 vSAN components. Maximum latency requirements to/from the vSAN Witness Host depend on the number of hosts in the cluster. 2 Node configurations are allowed up to five hundred milliseconds (500ms) and Stretched Clusters are allowed two hundred milliseconds (200ms) or one hundred milliseconds (100ms) depending on the number of hosts in the Stretched Cluster.

Using the VMware provided vSAN Witness Appliance is generally recommended as a better option for the vSAN Witness Host than using a physical vSphere host. The utilization of a vSAN Witness Appliance is relatively low during normal operations. It is not until a failover process occurs that a vSAN Witness Host will have any significant utilization. Because of this, especially in large 2 Node deployments to ROBO sites, multiple vSAN Witness Appliances may be run on the same shared vSphere infrastructure. VMware supports running the vSAN Witness Appliance on any vSphere 5.5 or higher infrastructure, which can include a standalone ESXi host, a typical vSphere infrastructure, in OVH (the service formally known as vCloud Air), any vCloud Air Network Partner, or any Service Provider/Shared/Co-Location where vSphere is used.

When using a vSAN Witness Appliance, it is patched in the same fashion as any other ESXi host. It is the last host updated when performing 2 Node and Stretched Cluster upgrades and should not be

backed up. Should it become corrupted or deleted, it should be redeployed. vSAN 6.6 introduced a quick and easy wizard to change the associated vSAN Witness Host.



The Change Witness Host is available in the vSphere Web Client and the vSphere Client above based on the Clarity UI framework.

2.3 Hardware Support

Hardware Support - Compute

Compute hosts that are on the VMware Hardware Compatibility List (HCL) are supported with their approved versions of vSphere and as a result, vSAN.

The is no distinction between what is required for compute with vSphere and vSAN. This is because vSAN is a native part of vSphere. Because of this, customers who already have an investment in vSphere supported hosts can easily add vSAN as a storage platform.

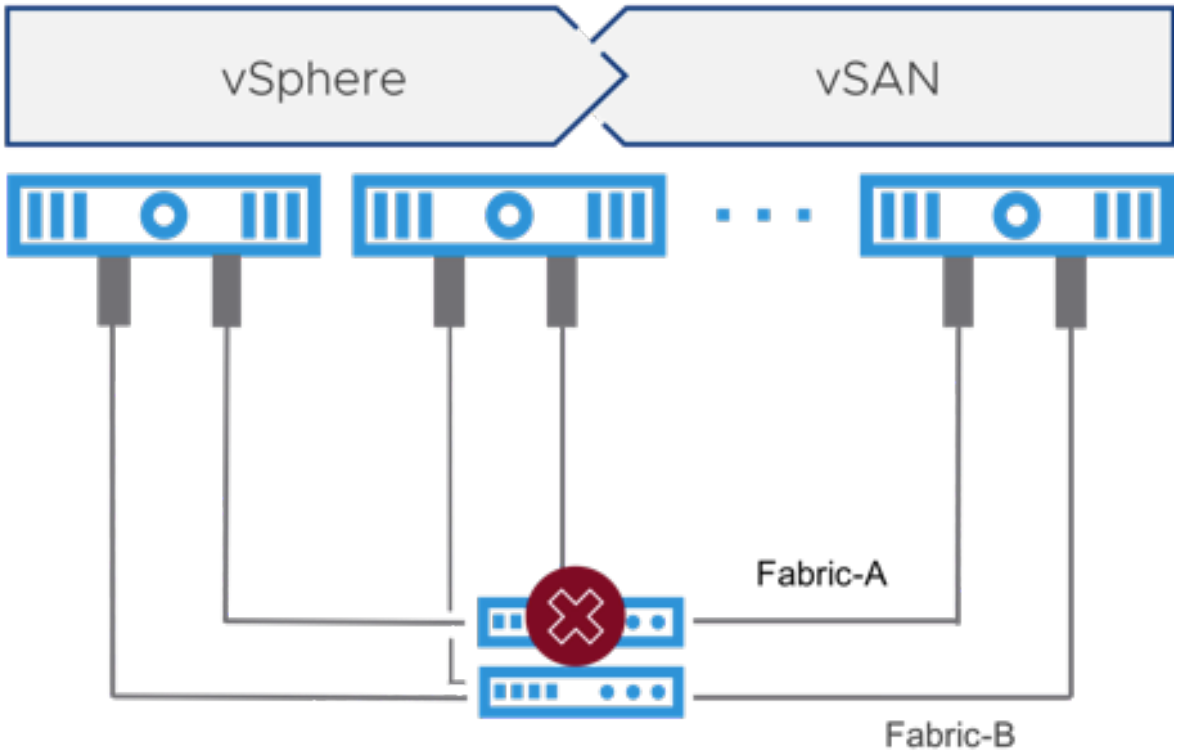
Hardware Support - Networking

Networking requirements for vSAN include both hardware and connectivity. vSphere Host networking devices that are on the VMware Hardware Compatibility List (HCL) are supported for use with their approved versions of vSphere, and again also with vSAN.

There are also some minimal network requirements for using vSAN. As of version 6.6, VMware removed the requirement for network infrastructure to support Multicast. vSAN 6.6 and higher versions require Unicast. Hosts participating in a vSAN cluster must be connected to a Layer 2 or Layer 3 network, and either IPv4 or IPv6 may be used.

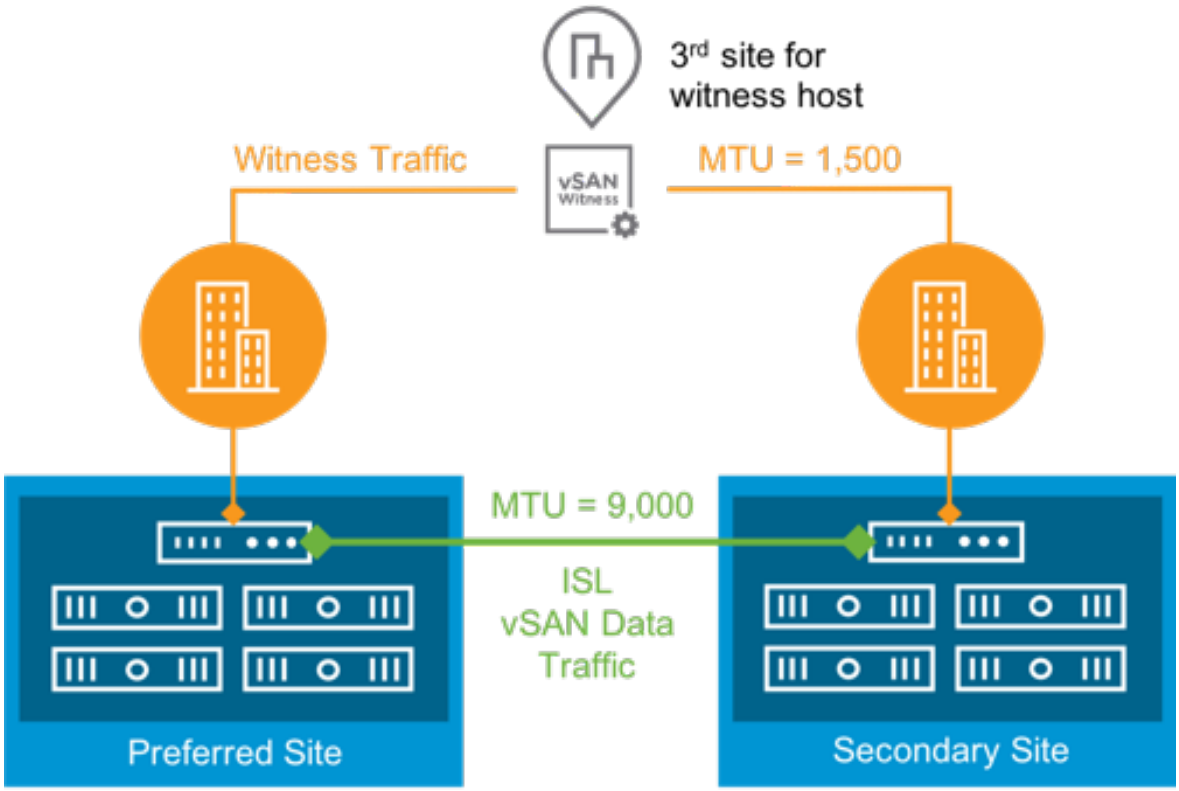
Host bandwidth to the vSAN network must be at least 1Gbps for a Hybrid configuration or shared (or dedicated) 10Gbps bandwidth for All-Flash configurations. The network infrastructure for a vSAN environment should be designed with the same levels of redundancy as any other storage fabric, without the requirement for a costly, specialized storage fabric. This helps ensure desired performance and availability requirements are met for the workloads running on vSAN.

In most vSAN environments, a single VMkernel interface is backed by redundant physical NICs and network connectivity. VMware also supports the use of multiple vSAN VMkernel interfaces in a multi-fabric approach with vSAN.



vSAN 6.7 improved upon the performance of the failover process from one vSAN network, or "fabric", to an alternate/redundant fabric.

vSAN 6.7 also introduced support for Witness Traffic Separation for Stretched Cluster configurations just as 2 Node Clusters have since vSAN 6.5. Witness Traffic Separation aligns connectivity requirements relative to type of traffic. This reduces network complexity in these configurations and isolates traffic types.

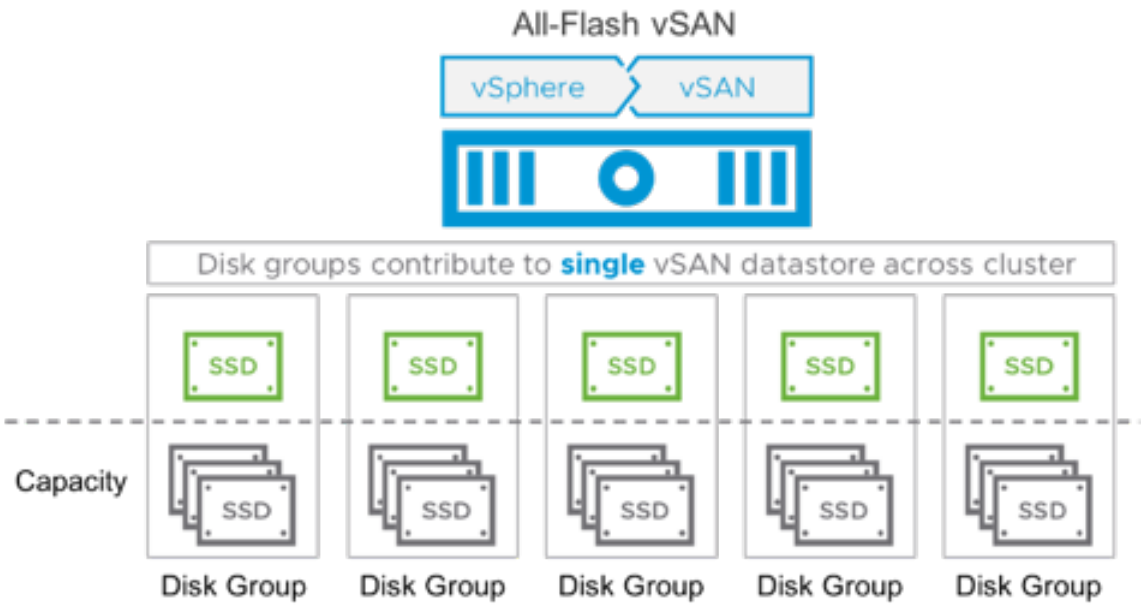


New in vSAN 6.7 U1, independent MTU sizes can be configured when using Witness Traffic Separation. This gives flexibility when administrators desire to use large frame sizes between sites and smaller frame sizes when connecting to the vSAN Witness Host, which has a significantly smaller traffic requirement.

More information on vSAN network requirements and configuration recommendations can be found in the VMware vSAN Network Design guide.

Hardware Support - Storage

vSAN hosts that contribute storage can be configured with between one (1) and five (5) Disk Groups for the storage of vSAN components. Disk Groups have requirement of a single flash device used for Cache, and between one (1) and seven (7) devices for Capacity.



In all Disk Group configurations, a flash device is used for Cache. In Hybrid configurations, the capacity devices are comprised of SAS or NL-SAS magnetic disks. In All-Flash configurations, the capacity devices may be flash SATA, SAS, PCIe, or NVMe.

Devices such as SAS, NL-SAS, or SATA are attached to a Host Bus Adapter (HBA) or RAID controller for consumption of vSAN. These devices may be connected in pass-through or RAID0 mode, depending

on the HBA/RAID controller. For controllers that do not support pass-through, each device must be presented as an individual RAID0 device. While RAID controllers may support drive mirroring, striping, or erasure coding, these are not supported, nor required by vSAN. vSAN accommodates data protection and performance properties using the Storage Policy Based Management (SPBM) framework instead.

Just as compute and networking must be on the VMware Hardware Compatibility List (HCL), vSAN storage devices, such as Host Bus Adapters (HBA), RAID Controllers, and storage devices must be on the VMware Compatibility Guide for vSAN to be supported.

Industry standard storage devices have been using a native (physical) 512 bytes sector size. Due to the increasing demand for drive with larger capacities, the storage industry introduced drive formats that use a 4KB physical sector sizes.

512e Drives

512e is the advanced format in which the physical sector size is 4,096 bytes, but the logical sector size emulates 512 bytes sector size. The purpose of 512e is for the new devices to be used with OSs that do not support 4Kn sectors yet. However, inherently, 512-byte emulation involves a read-modify-write process in the device firmware for every write operation that is not 4KB aligned.

With 512e drives, if a workload's I/O is either not aligned at a 4KB offset from the start of the disk or are not 4KB in length will get an alignment penalty caused by a read-modify-write process for every write operation. This is more pronounced for smaller operations, and for larger operations, the per operation latency is dominated by transfer times. In many cases, 512e drives are slightly faster than older 512n drives and this can be largely cancelled out for larger operations.

vSphere 6.5 and vSAN 6.5 introduced the ability to support 512e drives.

4Kn Drives

4Kn is the format in which the physical sector size is 4,096 bytes, and the logical sector size is also 4,096 as well. This 4,096-byte sector size requires less overall disk capacity required for metadata, thereby increasing a disk's payload capacity utilization.

vSphere 6.7 and vSAN 6.7 introduce support for 4Kn drives, offering better compatibility for today and more device options for tomorrow. vSAN's data plane I/O is optimized for 4Kn. Disk Groups in vSAN 6.7 can be comprised of 512n, 512e, and 4Kn disks. Virtual disks will continue to utilize a 512-byte disk format and can still be easily moved to or from supported vSphere datastores that do not use 4Kn devices. Non-vSphere datastores will be required to use VMFS 6 to support 4Kn devices. VMware vSAN 4Kn drives are supported for vSAN capacity only.

2.4 Deployment Options

Custom configurations using jointly validated components from all the major OEM vendors is an option. The vSAN Hardware Quick Reference Guide provides some sample server configurations as directional guidance and all components should be validated using the VMware Compatibility Guide for vSAN .

If the task of validating compute, networking, and storage components is daunting, there are some pre-validated solutions that make this process much easier.

Available from many of the leading server vendors, a vSAN ReadyNode™ is an x86 server, which is configured, tested, and certified for vSAN. vSAN ReadyNodes provide an open, flexible approach when considering deployment methods. Organizations can continue to use their server vendor(s) of choice. Each ReadyNode is optimally configured for vSAN with an optimal amount of CPU, memory, network, I/O controllers, and storage devices for a variety of workload types.

Alternatively, turn-key appliances such as Dell EMC VxRail™ provide a fully integrated VMware HyperConverged Infrastructure solution for a variety of applications and workloads. Simple deployment enables customers to be up and running in as little as 15 minutes. The turn-key appliance approach offers complete validation of interoperability among all of the components the solution is comprised of.

For customers looking to deploy vSAN at a much larger scale, Dell EMC VxRack™ SDDC system powered by VMware Cloud Foundation™ is available. These solutions provide an agile and flexible infrastructure foundation that IT organizations can use as part of their transformation into an IT-as-a-Service operating model.

3. Enabling vSAN

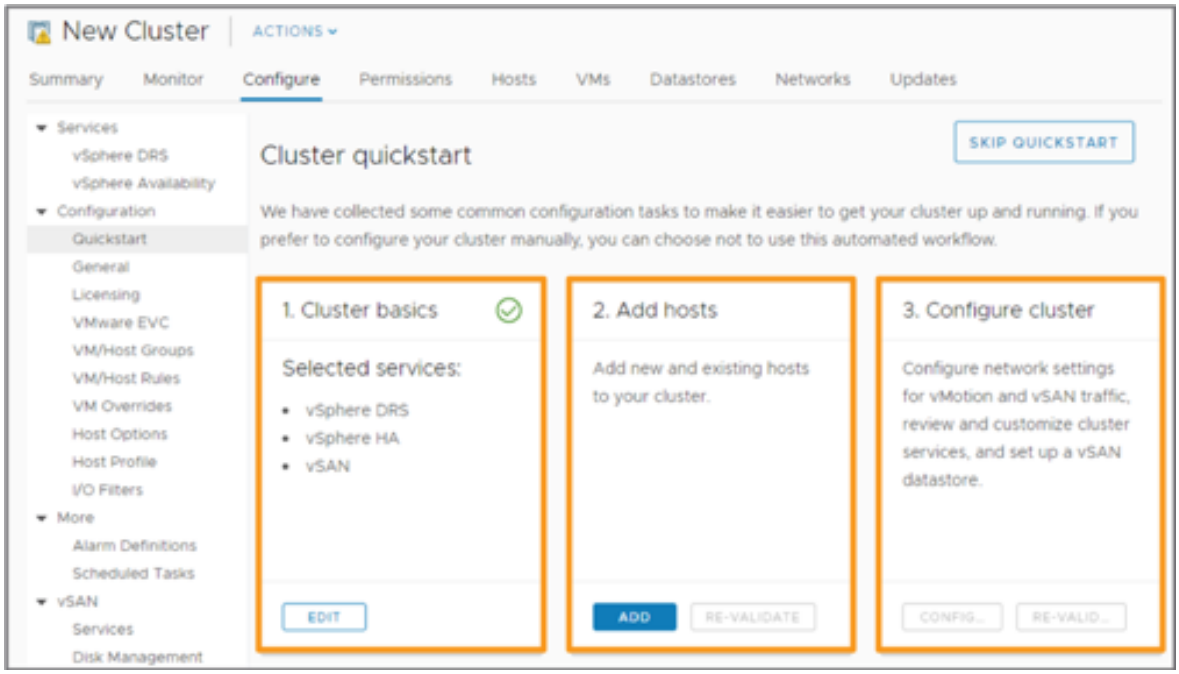
3.1 Cluster Quickstart

For existing vSphere clusters that meet the requirements, vSAN can be enabled with just a few mouse clicks. Because vSAN is part of vSphere, there is no requirement to install additional software or deploy any virtual storage appliances.

Cluster Quickstart

vSphere 6.7 U1 has introduced a new "Quickstart" guided cluster creation wizard that guides a user through the deployment process of vSAN, and non-vSAN, clusters.

It is easy to use, step-by-step configuration wizard that makes it even easier to create a production ready vSAN cluster. Cluster Quickstart handles not only the initial deployment, but also the process of expanding the cluster as needs change.



To enable vSAN, simply click the "Configure" option in the Configure tab of the vSphere cluster. This will start the process.

The Cluster Quickstart wizard workflow includes each of these to ease the deployment process:

- Cluster basics – Selection of services like vSphere DRS, vSphere HA and vSAN
- Adding hosts – Add multiple hosts simultaneously
- Cluster configuration
- Cluster type – Normal, 2 Node, Stretched
- Disk Group configuration
- Networking, including vSphere Distributed Switch Creation
- vMotion Networking
- vSAN Networking
- Services
- Deduplication & Compression
- Encryption
- vSphere HA Settings
- vSphere DRS Settings
- VMware EVC

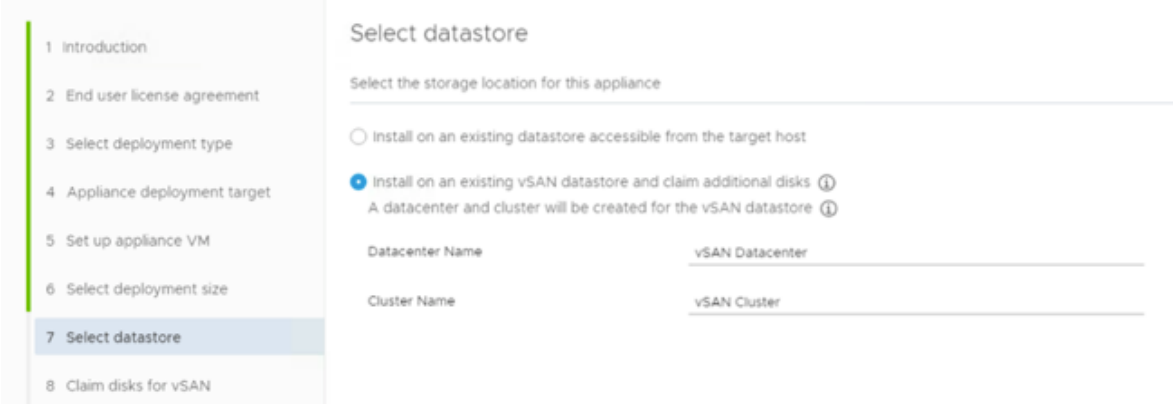
The Cluster Quickstart wizard works great for configuring vSAN clusters added to an existing vCenter 6.7 U1 deployment, as well as a great next step after using vSAN Easy Install to bootstrap a new vCenter onto the first host in a new vSAN cluster.

3.2 Starting from Scratch with Easy Install

Deployment of a vSAN cluster from scratch is easier than ever before. The vCenter Server Appliance (VCSA) installation wizard enables administrators to install vSphere on a single host, configure a vSAN datastore, and deploy a VCSA to this datastore. This is especially useful when deploying a new cluster where there is no existing infrastructure to host the VCSA.

After launching the VCSA installer and choosing installation parameters that pertain to the VCSA deployment like the deployment type, where the VCSA will be deployed to, and what size the VCSA will be, an option to select the datastore will be presented.

The vSAN Easy Install portion of the VCSA Installation will prompt for a datacenter name, a vSAN cluster name, claim disks on the host the VCSA is being installed to, and deploy the VCSA to that single node vSAN cluster.



The disk claiming process is very easy with an intuitive interface. Disks can easily be selected for Cache or Capacity use. Devices that are not properly represented, such as flash devices attached as RAID0 and are displayed as HDD, can be easily “marked” for vSAN to treat them as the media type they really are.

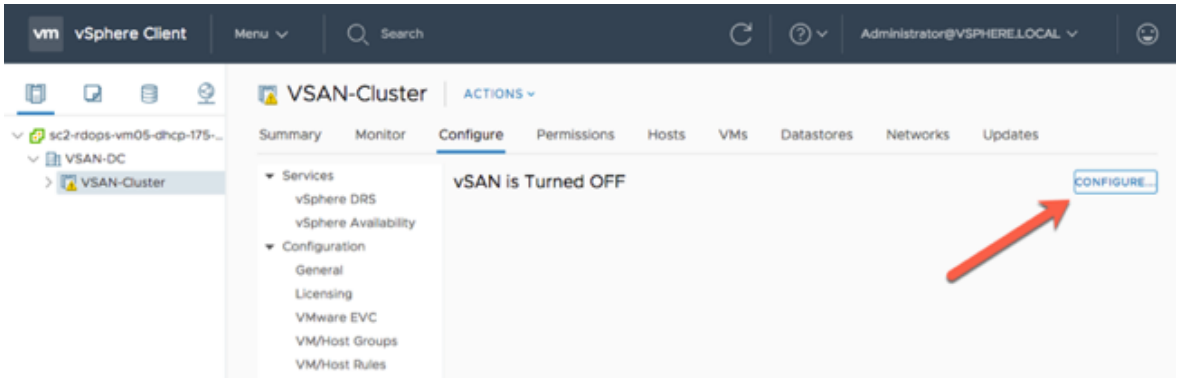
The VCSA installer will continue and request network settings before completing Stage 1 of the deployment process. When the VCSA installer begins Stage 2, a single node vSAN cluster has been created and the VCSA is deployed to that vSAN host. After completing Stage 2, the vCenter interface will be available for management of the environment.

Easy Install only deploys the VCSA to a single host. Mentioned earlier, vSAN requires 3 Nodes (or 2 Nodes and a vSAN Witness Host) for a supported configuration. Additional hosts will need to be added from the vSphere Client and the VCSA will need to have a vSAN Storage Policy applied to it.

Previously the process of starting from scratch required a manual bootstrapping process that required a specific order of operations. The Easy Install feature streamlines this process ensuring consistency and repeatability.

3.3 Using the vSAN Cluster Wizard

For administrators that have not yet decided to Deploy 6.7 Update 1, simply click the “Configure” option in the Configure tab of the vSphere cluster. to enable vSAN. This will start the process.



The Configure vSAN Wizard begins the process to easily configure vSAN in the cluster. This can be a single site cluster, 2 Node cluster, or a Stretched Cluster. After 60 days, a valid vSAN license will be required. *Stretched Clusters require a vSAN Enterprise license.

Additional services such as Deduplication and Compression as well as Encryption can be selected when enabling vSAN.

Deduplication and Compression will require a vSAN Advanced license and All-Flash hardware

Encryption will require a vSAN Enterprise license and may be used with either Hybrid or All-Flash hardware

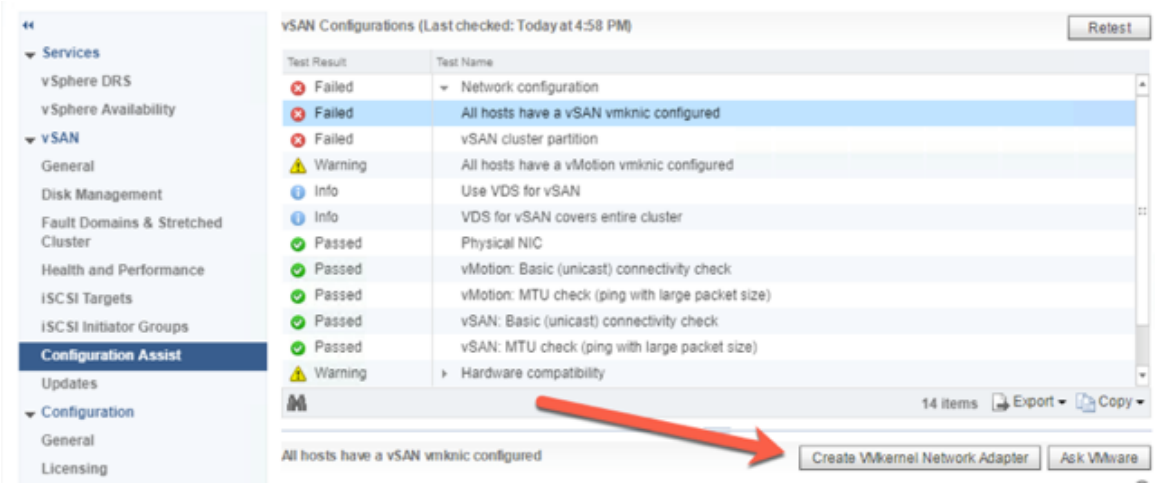
Local disks that are eligible to be used by each host in vSAN need to be claimed by vSAN for either the Cache or Capacity tiers. As mentioned previously, hosts may have between 1 and 5 Disk Groups, and each Disk Group may have between 1 and 7 capacity devices.

The vSAN Cluster Wizard easily addresses each of the tasks of configuring a vSAN Cluster.

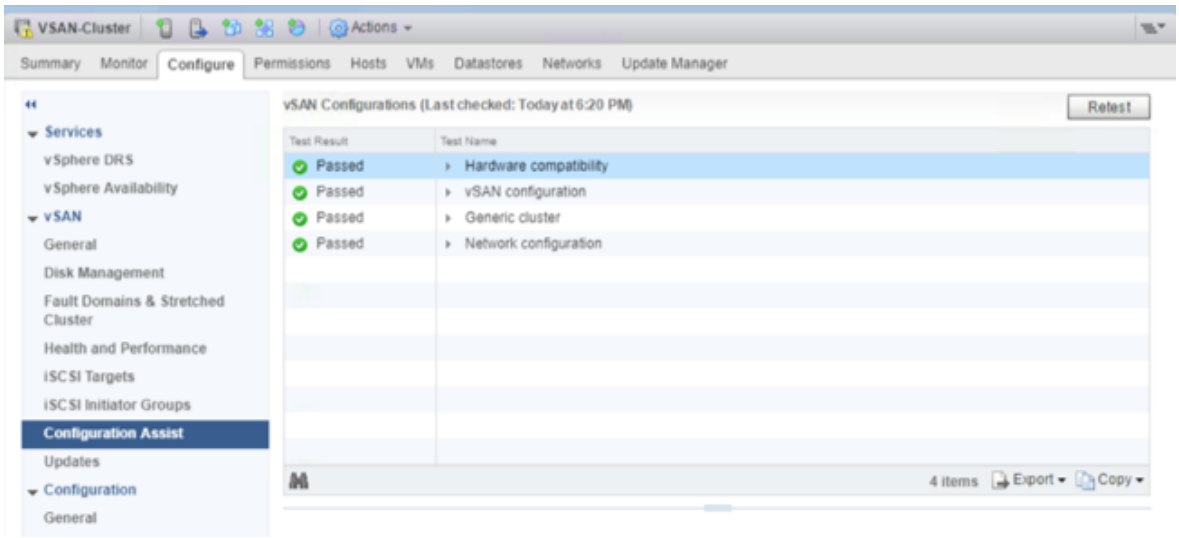
3.4 Configuration Assist to Complete the Deployment

The Configuration Assist tool introduced in vSAN 6.6 for vSAN can round out the configuration using the vSphere Web Client (Flex) for those that don't yet have Cluster Quickstart available.

From the vSAN Cluster's Configure tab, Configuration Assist can be selected to see the necessary tasks to complete the vSAN configuration. Tasks such as creating of a vSphere Distributed Switch, VMkernel ports for vSAN and vMotion, are easily completed.



Disks can be claimed as well as completing the general recommendations of enabling vSphere Availability (HA) and vSphere Distributed Resource Scheduler (DRS). When all warning items have been accomplished, the Configuration Assist results will be green.



Configuration Assist is not reserved for use with Easy Install based deployments and may be used when configuring an existing vSAN cluster. With deep integration into vSAN, Configuration Assist makes it easy to ensure a healthy vSAN cluster.

4. Availability

4.1 vSAN Objects and Component Placement

vSAN is an object datastore with a mostly flat hierarchy of objects and containers (folders). Items that make up a virtual machine are represented by objects. These are the most prevalent object types found on a vSAN datastore:

- VM Home, which contains virtual machine configuration files & logs
- Virtual machine swap
- Virtual disk (VMDK)
- Delta disk (snapshot)
- Performance database

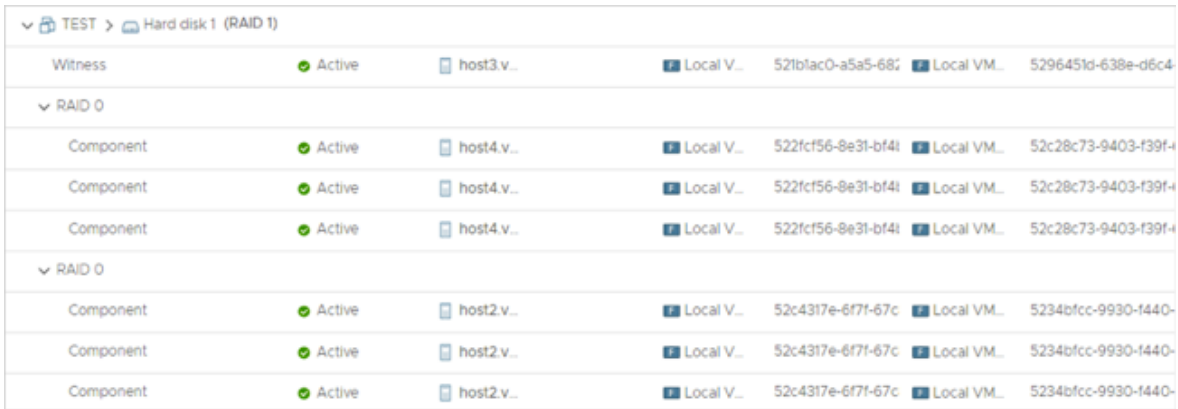
There are a few other objects that are commonly found on a vSAN datastore such as the vSAN performance service database, memory snapshot deltas, and VMDKs that belong to iSCSI targets.

vSAN Objects and Component Placement

Each object consists of one or more components. The number of components that make up an object depends primarily on a couple things: The size of the objects and the storage policy assigned to the object. The maximum size of a component is 255GB. If an object is larger than 255GB, it is split up into multiple components. The image below shows a 600GB virtual disk split up into three components.

vSAN will break down a large component into smaller components in certain cases to help balance capacity consumption across disks, optimize rebuild and resynchronize activities, and improve overall efficiency in the environment.

In most cases, a VM will have a storage policy assigned that contains availability rules such as Number of Failures to Tolerate and Failure Tolerance Method. These rules affect the number of components that make up an object. As an example, let's take that same 600GB virtual disk and apply the vSAN Default Storage Policy, which uses the RAID-1 mirroring failure tolerance method and has the number of failures to tolerate set to one. The 600GB object with three components will be mirrored on another host. This provides two full copies of the data distributed across two hosts so that the loss of a disk or an entire host can be tolerated. The figure below shows the six components (three on each host). A seventh component, a witness component, is created by vSAN to "break the tie" and achieve quorum in the event of a network partition between the hosts. The witness object is placed on a third host.



In this last example of component placement, we apply storage policy with RAID-5 erasure coding (Failures to Tolerate = 1) to an object. The object consists of four components—three data components

and a parity component. These components are distributed across four hosts in the cluster. If disk or host containing any one of these components is offline, the data is still accessible. If one of these components is permanently lost, vSAN can rebuild the lost data or parity component from the other three surviving components.

vSAN requires a minimum number of hosts depending on the failure tolerance method and number of failures to tolerate (FTT) configuration. For example, a minimum of three hosts is needed for FTT=1 with RAID-1 mirroring. A minimum of four hosts is required for FTT=1 with RAID-5 erasure coding.

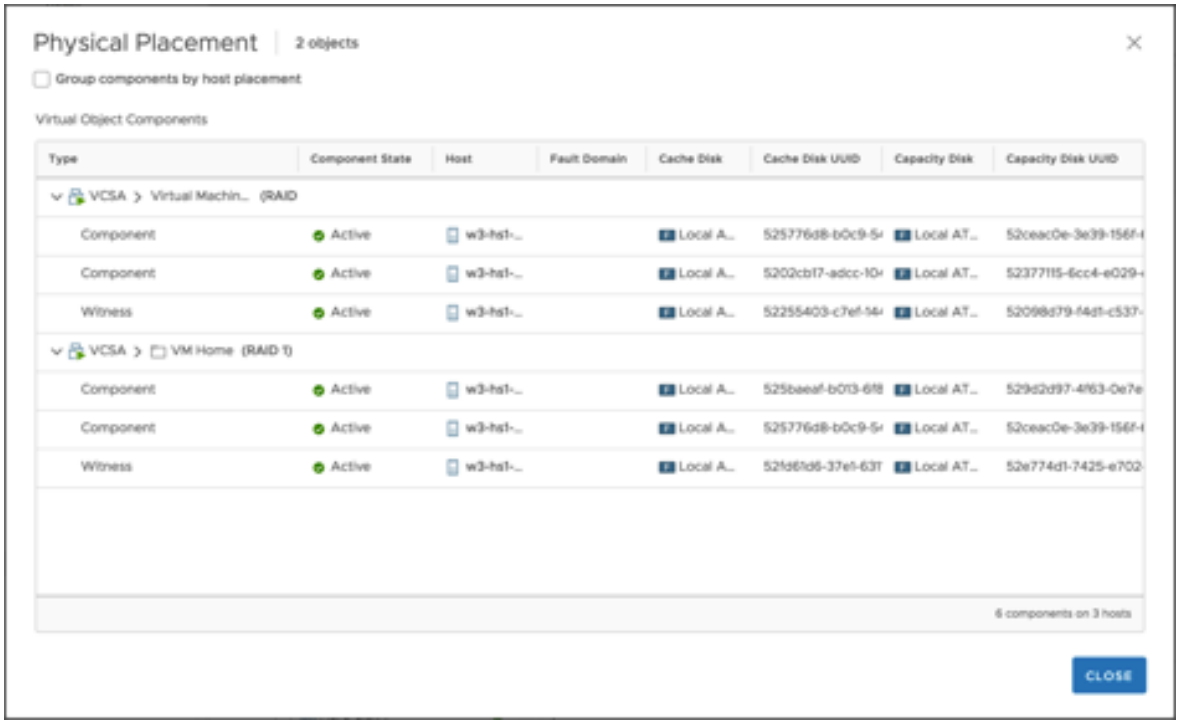
More details on cluster sizing minimums and recommendations can be found in the vSAN Design and Sizing Guide.

VM Swap Object Behavior

VM Swap objects are only created when a virtual machine is powered on and does not have a VM memory reservation that matches the assigned VM memory setting. In vSAN, this has traditionally been hard coded to a storage policy of being Mirrored once with the space reserved to ensure sufficient swap space was available to a VM should the need to swap arise.

When a large amount of swapping occurs, this is advantageous. In deployments with little to no swapping, this hard-coded policy can essentially reserve capacity that could otherwise be utilized. In vSAN 6.2, an advanced setting was added to mitigate this somewhat. When set to "1", the advanced setting **/VSAN/SwapThickProvisionDisabled** changed the default space reservation rule of 100% to 0%. VM swap objects were then sparse and only consumed space used. This setting had to be configured on each vSAN host and has often been overlooked or not known.

In vSAN 6.7, the behavior for VM swap objects now inherits the storage policy assigned to the VM home space object. Different workloads can now be easily assigned policies that are appropriate for their VM swap object.



Object Rebuilding, Resynchronization, & Consolidation

vSAN achieves high availability and extreme performance through the distribution of data across multiple hosts in a cluster. Data is transmitted between hosts using the vSAN network. There are cases where a significant amount of data must be copied across the vSAN network. One example is when you change the fault tolerance method in a storage policy from RAID-1 mirroring to RAID-5 erasure coding. vSAN copies or "resynchronizes" the mirrored components to a new set of striped components.

Another example is repair operations such as when vSAN components are offline due to a host hardware issue. These components are marked "absent" and colored orange in the vSAN user interface. vSAN waits 60 minutes by default before starting the repair operation. vSAN has this delay as many issues are transient. In other words, vSAN expects absent components to be back online in a reasonable amount of time and we want to avoid copying large quantities of data unless it is necessary. An example is a host being temporarily offline due to an unplanned reboot.

vSAN will begin the repair process for absent components after 60 minutes to restore redundancy. For example, an object such as a virtual disk (VMDK file) protected by a RAID-1 mirroring storage policy will create a second mirror copy from the healthy copy. This process can take a considerable amount of time depending on how much data must be copied. The rebuild process continues even if the absent copy comes back online in versions of vSAN prior to 6.6.

Repair Process

The object repair process was improved in vSAN 6.6. If absent components come back online while vSAN is rebuilding another copy, vSAN will determine whether it is more efficient to continue building an entirely new copy or update the existing copy that came back online. vSAN will restore redundancy using the most efficient method and cancel the other action. This enhancement in vSAN 6.6 rebuilds

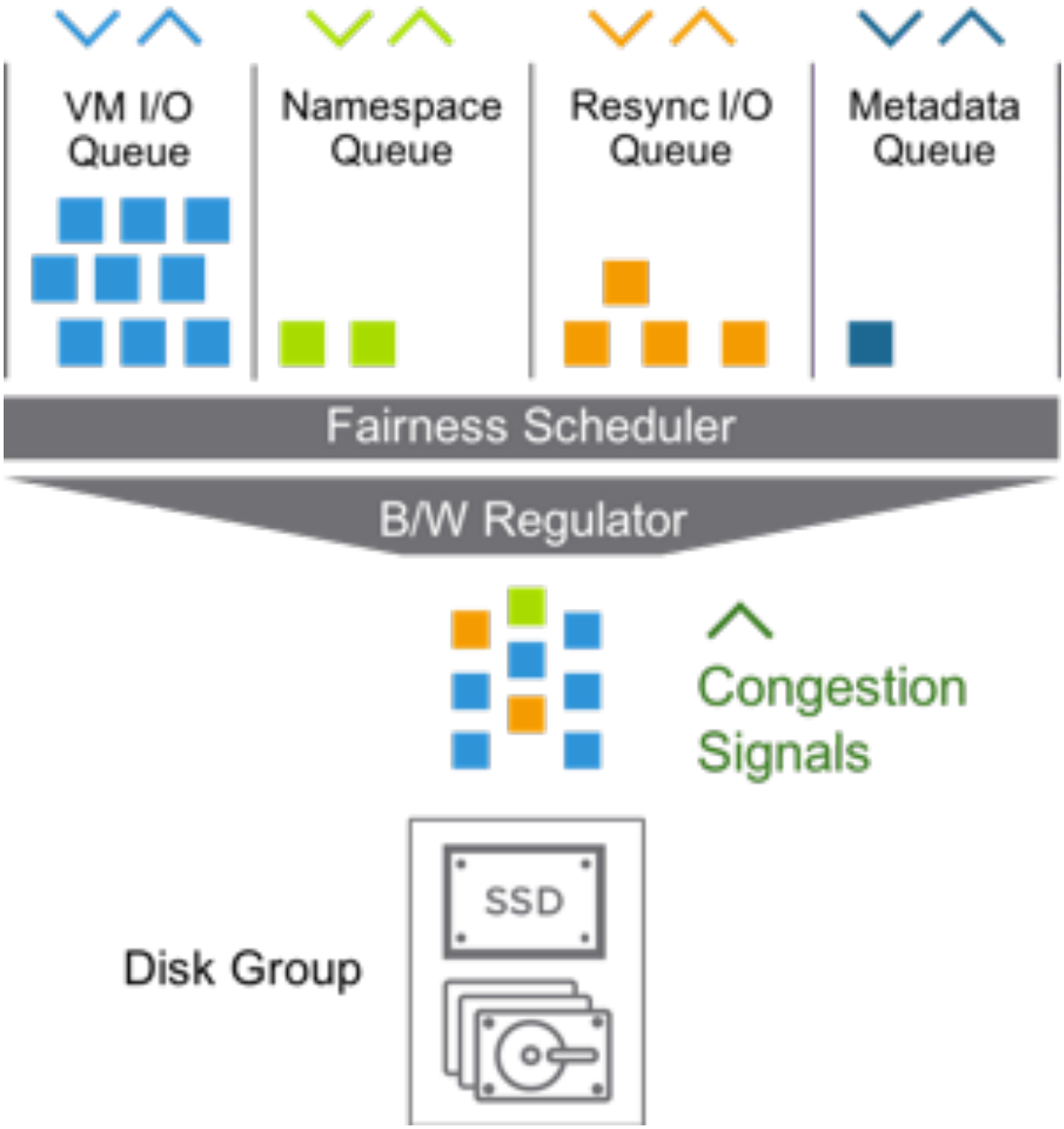
improves the speed and efficiency of object repair operations to reduce risk and minimize resource usage.

In cases where there are not enough resources online to comply with all storage policies, vSAN 6.6 will repair as many objects as possible. This helps ensure the highest possible levels of redundancy in environments affected by unplanned downtime. When additional resources come back online, vSAN will continue the repair process to comply with storage policies.

There are a few other operations that can temporarily increase vSAN “backend” traffic flow. Rebalancing of disk utilization is one of these operations. When a disk has less than 20% free space, vSAN will automatically attempt to balance capacity utilization by moving data from that disk to other disks in the vSAN cluster. Achieving a well-balanced cluster from a disk capacity standpoint can be more challenging if there are many large components. vSAN 6.6 improves efficiency by splitting large components into smaller components to achieve a better balance.

Rysynchronization
Excessive amounts of vSAN backend resynchronization traffic might affect cluster performance. Resynchronization operations in previous versions of vSAN are automated and controlled entirely by vSAN. In other words, administrators are unable to adjust the rate at which resynchronization operations are performed. Throughput of resynchronization operations could be manually adjusted to reduce the impact of excessive resynchronization activity at the cost of an increase in the amount of time needed to rebuild and or resynchronize components.

vSAN 6.7 introduces an Adaptive Resynchronization feature to ensure fair-share resources are available for virtual machine I/O and resynchronization I/O as the I/O changes.

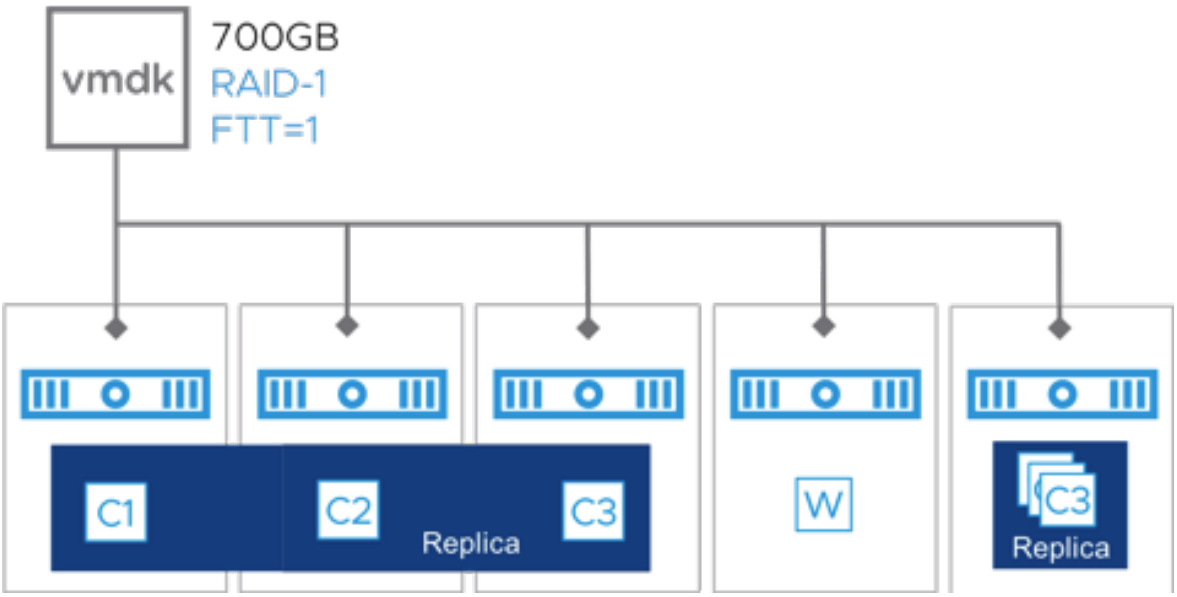


When I/O activity exceeds the sustainable Disk Group bandwidth, Adaptive Resync guarantees bandwidth levels for VM I/O and resynchronization I/O. During times without contention, VM I/O or resynchronization I/O are allowed to use additional bandwidth. If there are no resynchronization operations being performed VM I/O can consume 100% of the available Disk Group bandwidth. During times of contention, resynchronization I/O will be guaranteed 20% of the total bandwidth the Disk Group is capable of. This allows for a more optimal use of resources.

Replica Consolidation
When decommissioning a vSAN Capacity device, Disk Group, or host, data should be evacuated to maintain policy compliance. vSAN 6.6 introduced greater visibility into which components a decommissioning operation would affect, so administrators could make appropriate decommissioning choices.

Over time, due to vSAN datastore rebalancing, partial repairs, or other data placement operations, vSAN components can be split into smaller components. The splitting of these vSAN components is normal and can provide a more optimal utilization of a vSAN datastore.

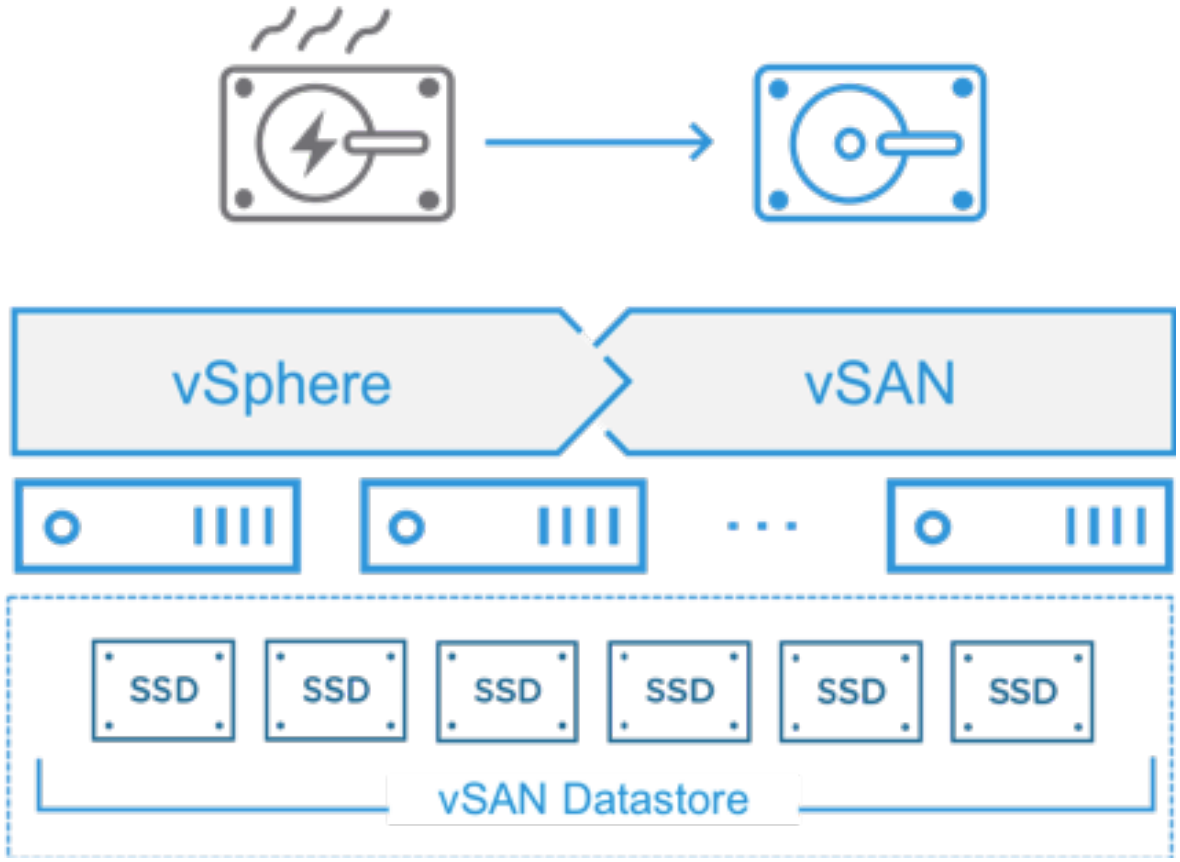
When a decommissioning task is requested, vSAN will attempt to find an open location to move data to that does not violate the anti-affinity data placement rules, so storage policy compliance is still satisfied. But what about cases where there is no available Fault Domain to move the data to?



If a Fault Domain already contains a vSAN component replica, and there is additional capacity for the replica that needs to be evacuated, vSAN now has the ability to consolidate them into a single replica. The smallest replicas are moved first, resulting in less data rebuilt, and less temporary capacity used.

Degraded Device Handling
VMware continues to improve how vSAN handles hardware issues such as a storage device that is showing symptoms of impending failure. In some cases, storage devices issues are easily detected through errors reported by the device, e.g., SCSI sense codes. In other cases, issues are not so obvious.

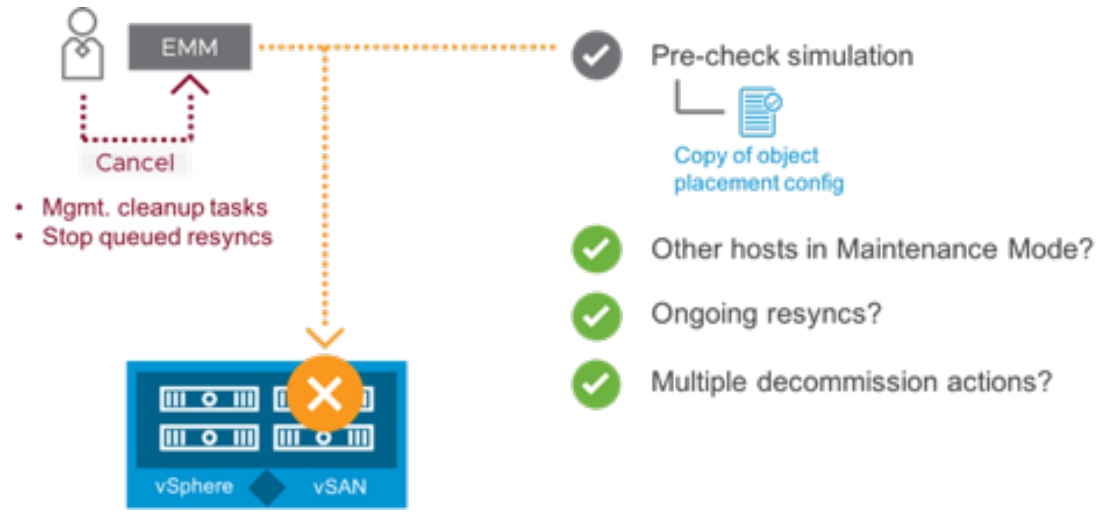
To proactively discover these types of issues, vSAN will track performance characteristics of a device over time. A significant discrepancy in performance is a good indicator of a potential problem with a device. vSAN uses multiple samples of data to help avoid “false positives” where the issue is transient in nature.



When failure of a device is anticipated, vSAN evaluates the data on the device. If there are replicas of the data on other devices in the cluster, vSAN will mark these components as “absent”. “Absent” components are not rebuilt immediately as it is possible the cause of the issue is temporary. vSAN waits for 60 minutes by default before starting the rebuilding process. This does not affect the availability of a virtual machine as the data is still accessible using one or more other replicas in the cluster. If the only replica of data is located on a suspect device, vSAN will immediately start the evacuation of this data to other healthy storage devices. Intelligent, predictive failure handling drives down the cost of operations by minimizing the risk of downtime and data loss.

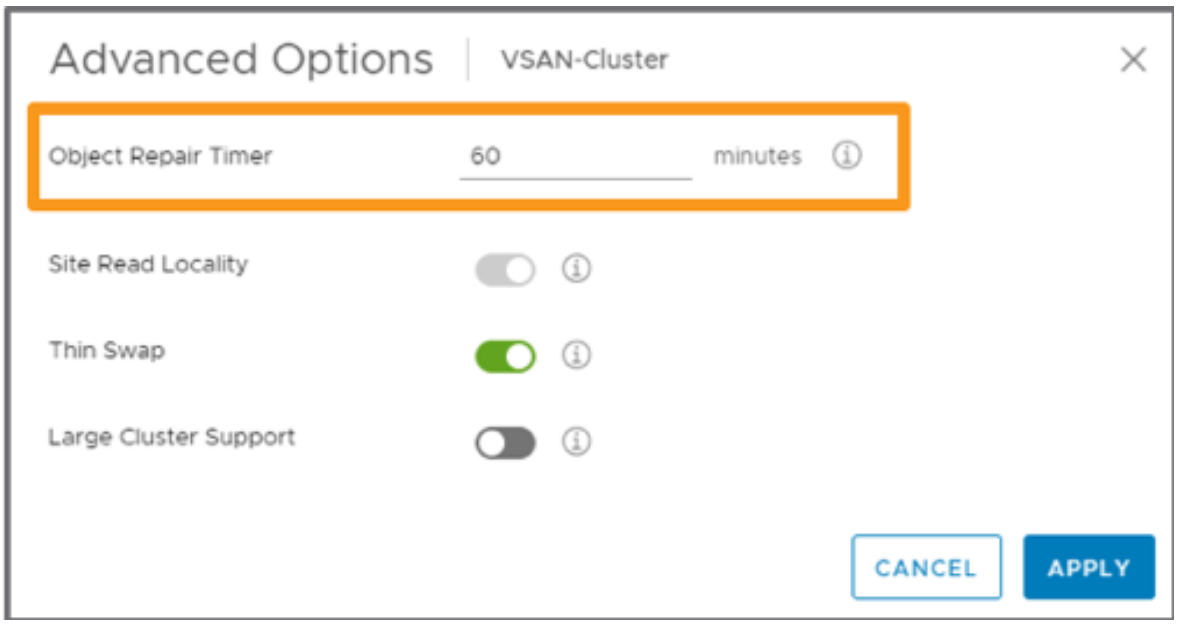
Decommissioning and Maintenance Mode Safeguards in vSAN 6.7 U1
Since each vSAN host in a cluster contributes to the cluster storage capacity, entering a host into maintenance mode takes on an additional set of tasks when compared to a traditional architecture.

vSAN 6.7 U1 has improved the safeguards when performing maintenance and decommissioning activities on vSAN hosts. vSAN will now perform a full simulation of data movement to determine if the process of entering maintenance mode will succeed or fail before it even starts. This will prevent unnecessary data movement and provide a result more quickly to the administrator.



New warnings have been added to entering maintenance mode activities to ensure that there are no other hosts already in maintenance mode or resync activity current performing.

For cases where an administrator needs to adjust the time vSAN waits before it begins to rebuild data to reestablish compliance with storage policies, a new “object repair timer delay” setting is now in the UI.



All of these improvements are added to enhance the overall experience and predictability of host decommissioning activities like entering a host into maintenance mode.

4.2 Fault Domains

"Fault domain" is a term that comes up often in availability discussions. In IT, a fault domain usually refers to a group of servers, storage, and/or networking components that would be impacted collectively by an outage. A common example of this is a server rack. If a top-of-rack switch or the power distribution unit for a server rack would fail, it would take all the servers in that rack offline even though the server hardware is functioning properly. That server rack is considered a fault domain.

Each host in a vSAN cluster is an implicit fault domain. vSAN automatically distributes components of a vSAN object across fault domains in a cluster based on the Number of Failures to Tolerate rule in the assigned storage policy. The following diagram shows a simple example of component distribution across hosts (fault domains). The two larger components are mirrored copies of the object and the smaller component represents the witness component.

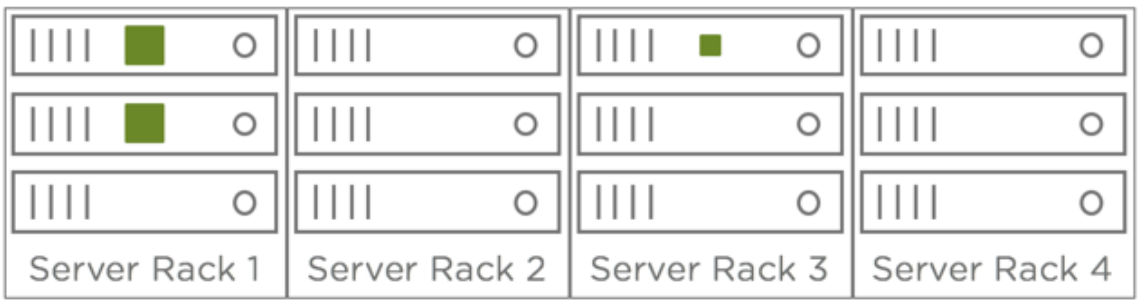
When determining how many hosts or Fault Domains a cluster is comprised of, it is important to remember the following:

- For vSAN objects that will be protected with Mirroring, there must be 2n+1 hosts or Fault Domains for the level of protection chosen.
  - Protecting from 1 Failure would require (2x1+1) or 3 hosts
  - Protecting from 2 Failures would require (2x2+1) or 5 hosts
  - Protecting from 3 Failures would require (2x3+1) or 7 hosts
- For vSAN objects that will be protected with Erasure Coding, there must be 2n+2 hosts or Fault Domains for the level of protection chosen.
  - RAID5 (3+1) requires (2x1+2) or 4 hosts
  - RAID6 (4+2) requires (2x2+2) or 6 hosts

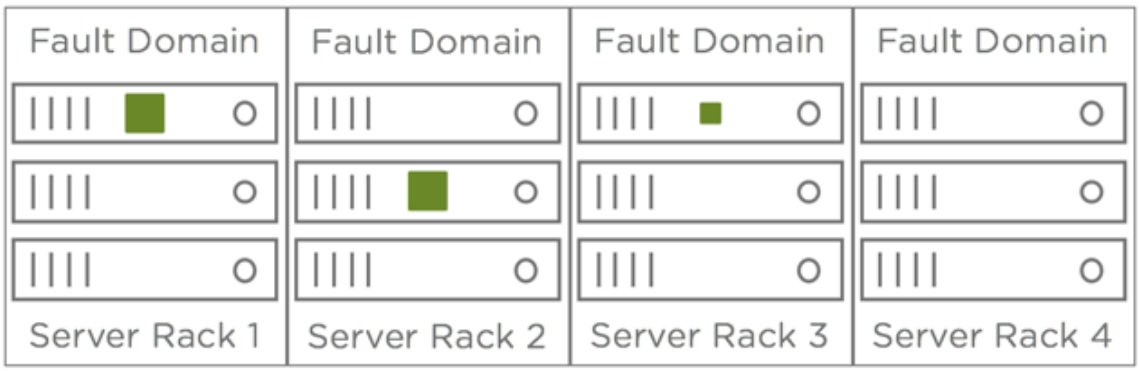
Also consider that the loss of a Fault Domain, or hosts when Fault Domains are not configured, could result in no location to immediately rebuild to. VMware recommends having an additional host or Fault Domain to provide for the ability to rebuild in the event of a failure.

Using Fault Domains for Rack Isolation

The failure of a disk or entire host can be tolerated in the previous example scenario. However, this does not protect against the failure of larger fault domains such as an entire server rack. Consider our next example, which is a 12-node vSAN cluster. It is possible that multiple components that make up an object could reside in the same server rack. If there is a rack failure, the object would be offline.



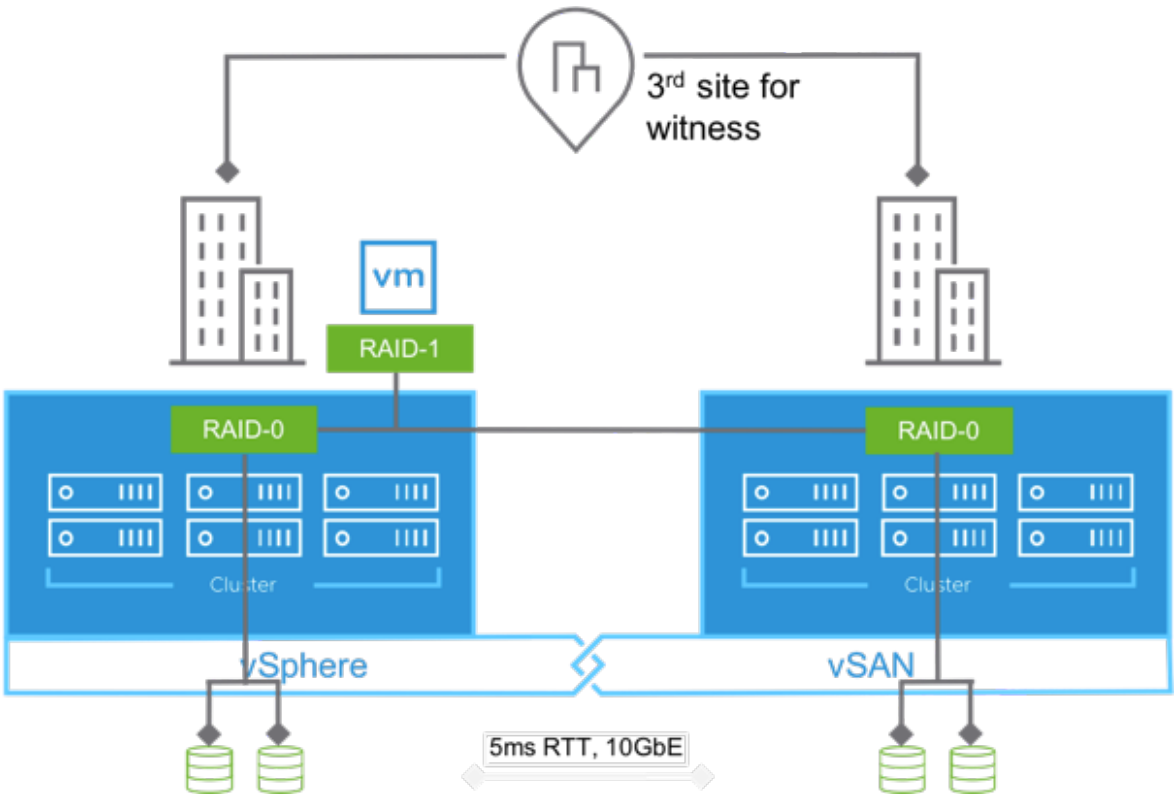
To mitigate this risk, place the servers in a vSAN cluster across server racks and configure a fault domain for each rack in the vSAN UI. This instructs vSAN to distribute components across server racks to eliminate the risk of a rack failure taking multiple objects offline. This feature is commonly referred to as "Rack Awareness". The diagram below shows component placement when three servers in each rack are configured as separate vSAN fault domains.



4.3 Stretched Clusters

vSAN Stretched Clusters provide resiliency against the loss of an entire site. vSAN is integrated tightly with vSphere Availability (HA). If a site goes offline unexpectedly, vSphere HA will automatically restart the virtual machines affected by the outage at the other site with no data loss. The virtual machines will begin the restart process in a matter of seconds, which minimizes downtime. Stretched Clusters are being included in this section, because object placement is handled a bit differently when using Stretched Clusters.

vSAN stretched clusters are also beneficial in planned downtime and disaster avoidance situations. Virtual machines at one site can be migrated to the other site with VMware vMotion. Issues such as an impending storm or rising flood waters typically provide at least some time to prepare before disaster strikes. Virtual machines can easily be migrated out of harm's way in a vSAN Stretched Cluster environment.



The limitations of what is possible centers on network bandwidth and round-trip time (RTT) latency. Nearly all stretched cluster solutions need a RTT latency of 5 ms or less. Writes to both sites must be

committed before the writes are acknowledged. RTT latencies higher than 5 ms introduce performance issues. vSAN is no exception. A 10Gbps network connection with 5 ms RTT latency or less is required between the preferred and secondary sites of a vSAN stretched cluster.

Up to 15 hosts per site are currently supported. In addition to the hosts at each site, a "witness" must be deployed to a third site. The witness is a VM appliance running ESXi that is configured specifically for use with a vSAN stretched cluster. Its purpose is to enable the cluster to achieve quorum when one of the two main data sites is offline. The witness also acts as “tie-breaker” in scenarios where a network partition occurs between the two data sites. This is sometimes referred to as a “split-brain” scenario.

The witness does not store virtual machine data such as virtual disks. Only metadata such as witness components is stored on the witness.

Up to 200ms RTT latency is supported between the witness site and data sites. The bandwidth requirements between the witness site and data sites vary and depend primarily on the number of vSAN objects stored at each site. A minimum bandwidth of 100Mbps is required and the general rule of thumb is at least 2Mbps of available bandwidth for every 1000 vSAN objects. The vSAN Stretched Cluster Bandwidth Sizing guide provides more details on networking requirements.

Stretched Cluster Fault Domains
A vSAN Stretched Cluster consists of three Fault Domains.

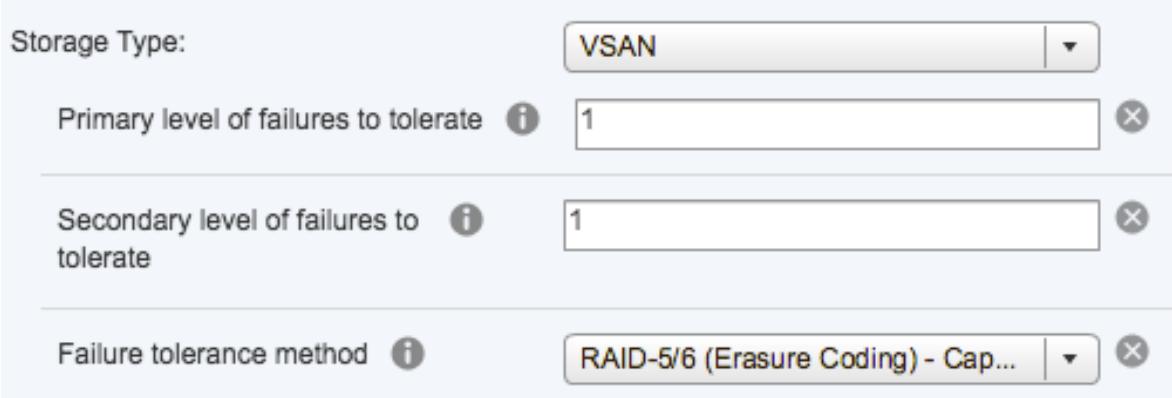
- Physical hosts in the primary or “preferred” location make up one Fault Domain
- Physical hosts in the secondary location are the second Fault Domain
- A vSAN Witness Host is in an implied third Fault Domain placed at a tertiary location. *Because the vSAN Witness Host is not a member of the vSAN Stretched Cluster, the Fault Domain is implied.

When vSAN Stretched Clusters were first introduced, a mirroring storage policy was the only option to protect data across sites. Data was mirrored across sites, with one replica in each site. Metadata (vSAN witness objects) are placed on the vSAN Witness Host at a third site. If any one of the sites goes offline, there are enough surviving components to achieve quorum, so the virtual machine is still accessible.

vSAN 6.6 introduced the ability to configure a secondary level of failures to tolerate. This feature enables resiliency within a site, as well as, across sites. For example, RAID-5 erasure coding protects objects within the same site while RAID-1 mirroring protects these same objects across sites.

Local failure protection within a vSAN stretched cluster further improves the resiliency of the cluster to minimize unplanned downtime. This feature also reduces or eliminates cross-site traffic in cases where components need to be resynchronized or rebuilt. vSAN lowers the total cost of ownership of a stretched cluster solution as there is no need to purchase additional hardware or software to achieve this level of resiliency.

This is configured and managed through a storage policy in the vSphere Web Client. The figure below shows rules in a storage policy that is part of an all-flash stretched cluster configuration. The primary level of failures to tolerate is set to 1, which instructs vSAN to mirror data across the two main sites of the stretched cluster. The secondary level of failures to tolerate specifies how data is protected within the site. In the example storage policy below, RAID-5 erasure coding is used, which can tolerate the loss of a host within the site.



To clarify the selection of cross site and local protection, the new HTML5 based vSphere Client presents this a bit differently.

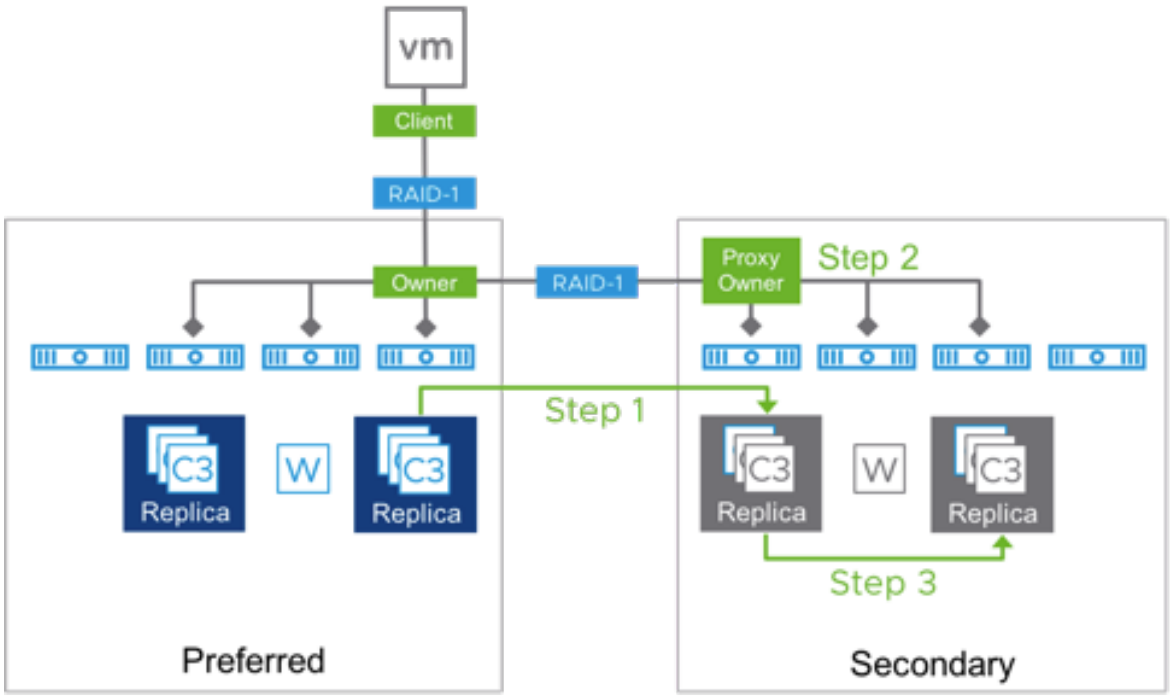
To avoid confusion, the Storage Policy Wizard now explicitly asks for Site disaster tolerance and the number of failures to tolerate.

Stretched Cluster Data Sites
A maximum of thirty (30) hosts may be used in a single vSAN Stretched Cluster across the data sites. In vSAN versions up to 6.5, this was fifteen (15) per site. With the introduction of Site Affinity rules that places data on only one data site or the other, it is possible to have a vSAN Stretched Cluster deployment that does not have an equal number of hosts per site.

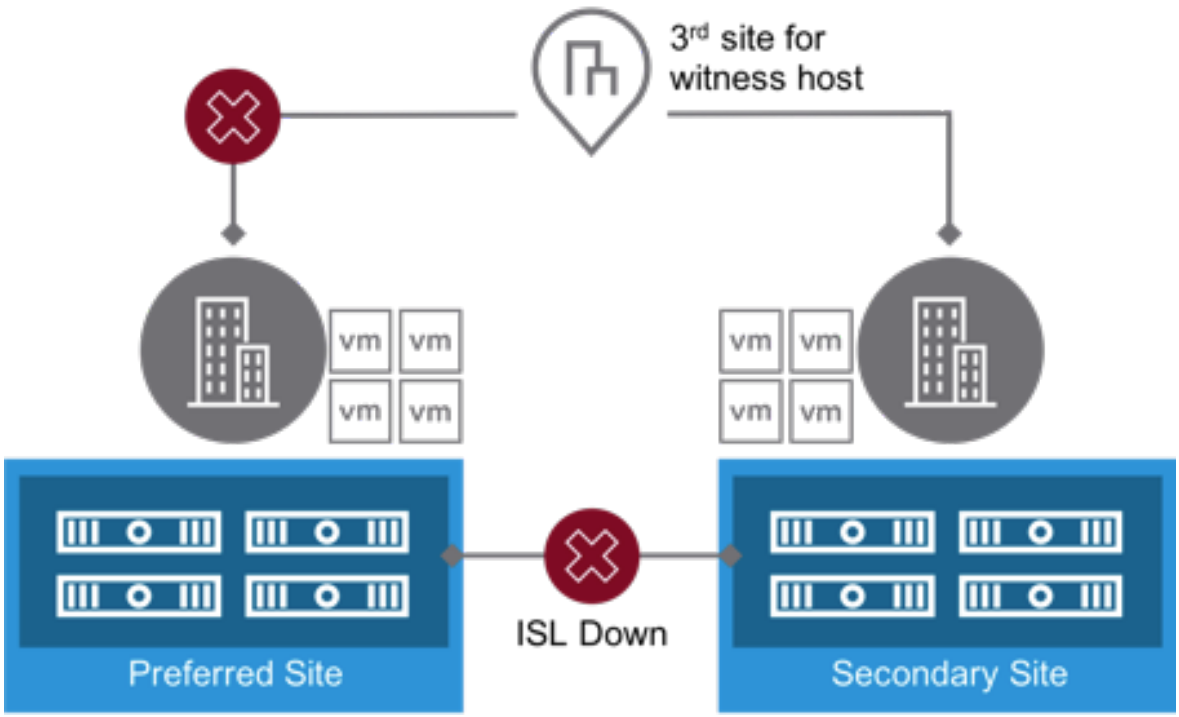
Network bandwidth and round-trip time (RTT) latency are the primary factors that must be considered when deploying a vSAN Stretched Cluster. Because writes are synchronous, they must be committed before they may be acknowledged. When RTT latencies are higher than five milliseconds (5ms) performance issues may result. The maximum supported RTT between the two data sites is 5ms.

A minimum bandwidth of 10Gbps is recommended for the inter-site link connecting the preferred and secondary sites. The actual bandwidth required is entirely dependent on the workload that is running on the Stretched Cluster. Only writes traverse the inter-site link under normal conditions. Read locality was introduced with vSAN Stretched Clusters in version 6.1 to keep reads local to the site a VM is running on and better utilize the inter-site link. The Stretched Cluster Bandwidth Sizing Guide can provide additional guidance to determine the required bandwidth based on the workload profile.

In failure scenarios that result in an incomplete local dataset, it is more efficient to copy only enough pieces necessary to repair the local data set than it is to perform a full copy. vSAN 6.7 performs a partial resynchronization to bring one replica to the degraded site, triggers the local proxy owner to begin the local resync, and the resync/repair process is performed locally.



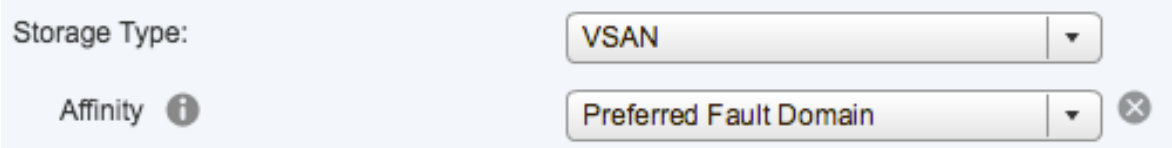
In larger scale failure scenarios, such as when the Preferred site is completely isolated from the vSAN Witness Host and the inter-site link is down, vSAN will fail over to the Secondary site. As connectivity is restored to the Preferred site, it does not become the authoritative (preferred) site until it has regained connectivity with both the vSAN Witness Host and the Secondary site. This prevents the situation where the Preferred site reports as being available and attempts to fail workloads back to stale data.



Stretched Cluster Site Affinity
vSAN 6.6 improved upon the flexibility of storage policy-based management for Stretched Clusters by introducing the Site Affinity rule. You can specify a single site to locate virtual machine objects in cases

where cross-site redundancy is not necessary. Common examples include applications that have built-in replication or redundancy such as Microsoft Active Directory and Oracle Real Application Clusters (RAC). This capability reduces costs by minimizing the storage and network resources used by these workloads.

Affinity is easy to configure and manage using storage policy-based management. A storage policy is created and the Affinity rule is added to specify the site where a virtual machine's objects will be stored.



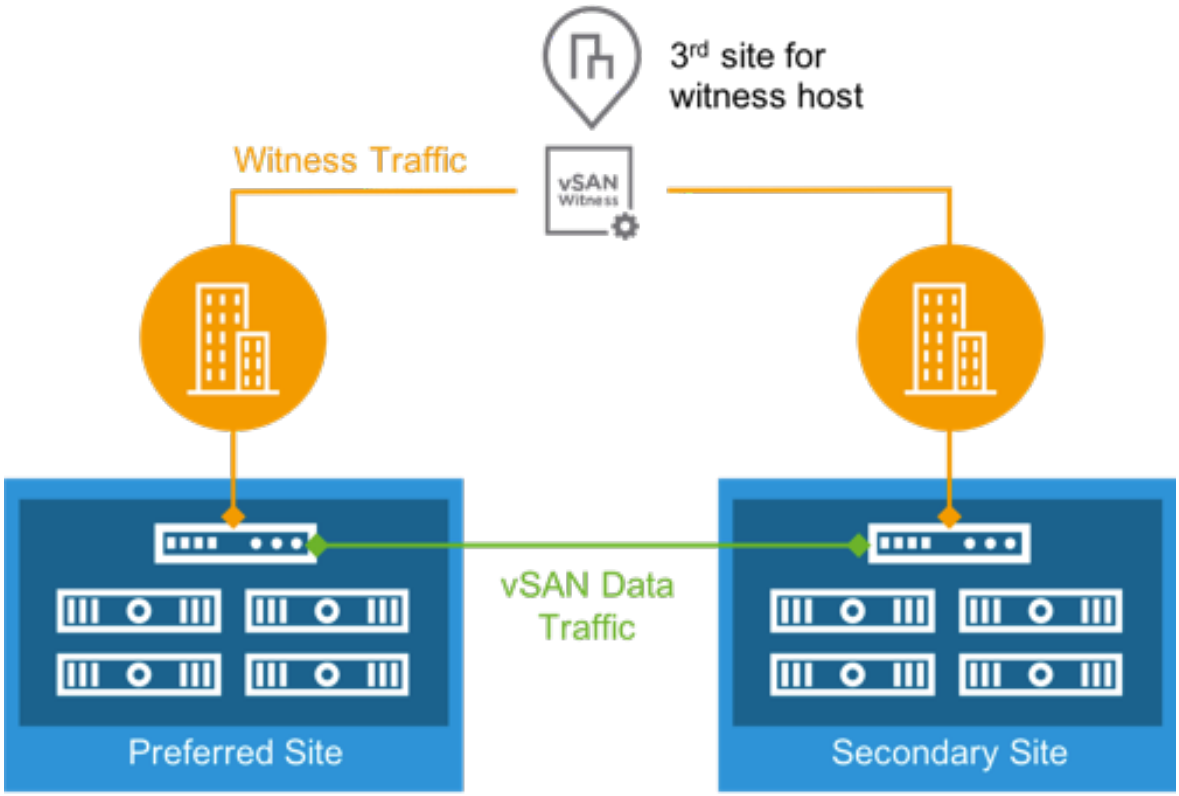
Stretched Clusters Witness Site
A vSAN Witness host must be deployed to a third site. The vSAN Witness Host allows the cluster achieve quorum when one of the two main data sites are offline as well as acts as "tie-breaker" in scenarios where a network partition occurs between the two data sites. This is sometimes referred to as a "split-brain" scenario.

The vSAN Witness Host does not store virtual machine data such as virtual disks. Only metadata is stored on the vSAN Witness Host. This includes witness components for objects residing on the data sites.

Up to 200ms RTT latency is supported between the witness site and data sites. The bandwidth requirements between the witness site and data sites vary and depend primarily on the number of vSAN objects stored at each site. The general rule of thumb is at least 2Mbps of available bandwidth for every 1000 vSAN objects. The vSAN Stretched Cluster Bandwidth Sizing guide provides more details on networking requirements.

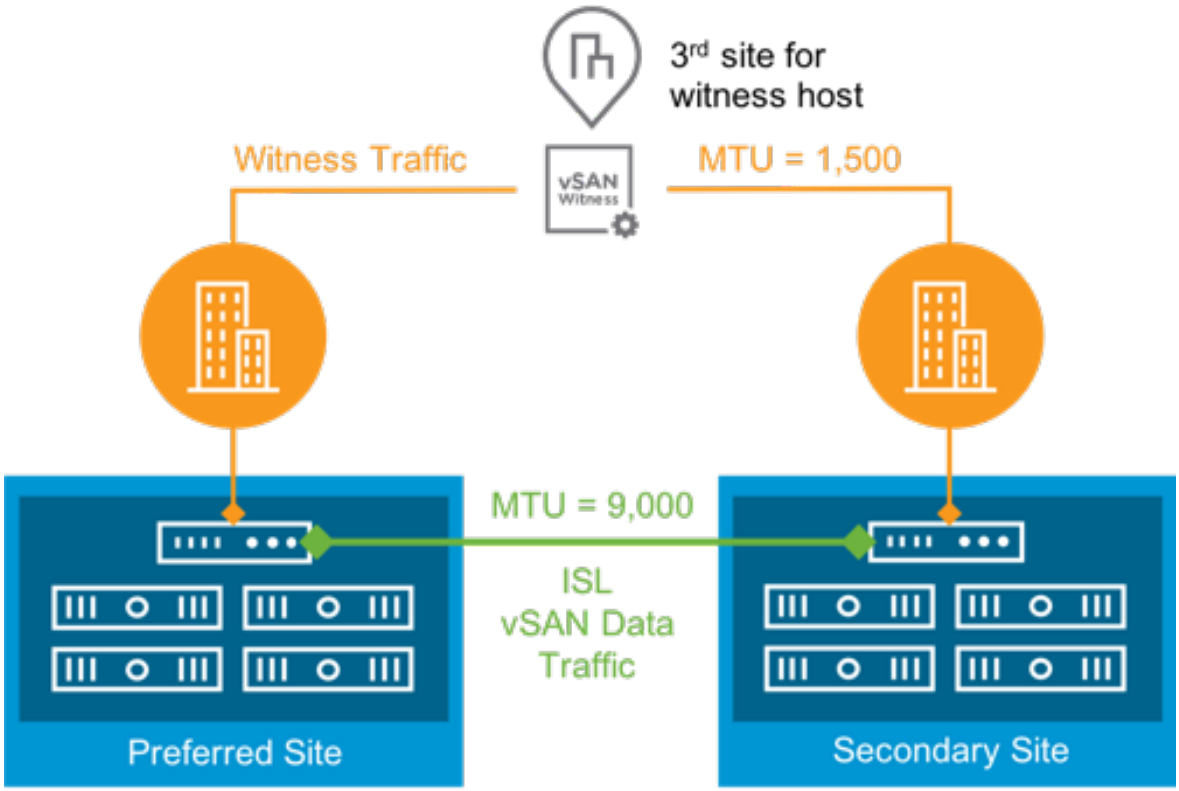
The vSAN data network in Stretched Clusters has been required to have connectivity to the vSAN Witness Host in deployments up until now. Improper configuration could have sent data destined for the alternate site across the lower bandwidth, higher latency network. This improper routing could also interfere with the proper failover of workloads in the event of a failed or isolated site.

Witness Traffic Separation (WTS) was introduced in vSAN 6.5 for 2 Node vSAN clusters and is now supported in vSAN 6.7 with Stretched Clusters.



By allowing vSAN data sites to communicate with the vSAN Witness Host using an alternate interface, there is no opportunity for the vSAN data network to be exposed to the WAN.

Stretched Clusters Mixed MTU Support – New in vSAN 6.7 U1
The ability to isolate witness traffic using dedicated uplinks in a vSAN Stretched Cluster environment was a powerful enhancement made to vSAN 6.7.



vSAN 6.7 U1 introduces additional levels of flexibility to this feature with the support of different MTU sizes configured for witness traffic and the Inter-site link used for vSAN data traffic.

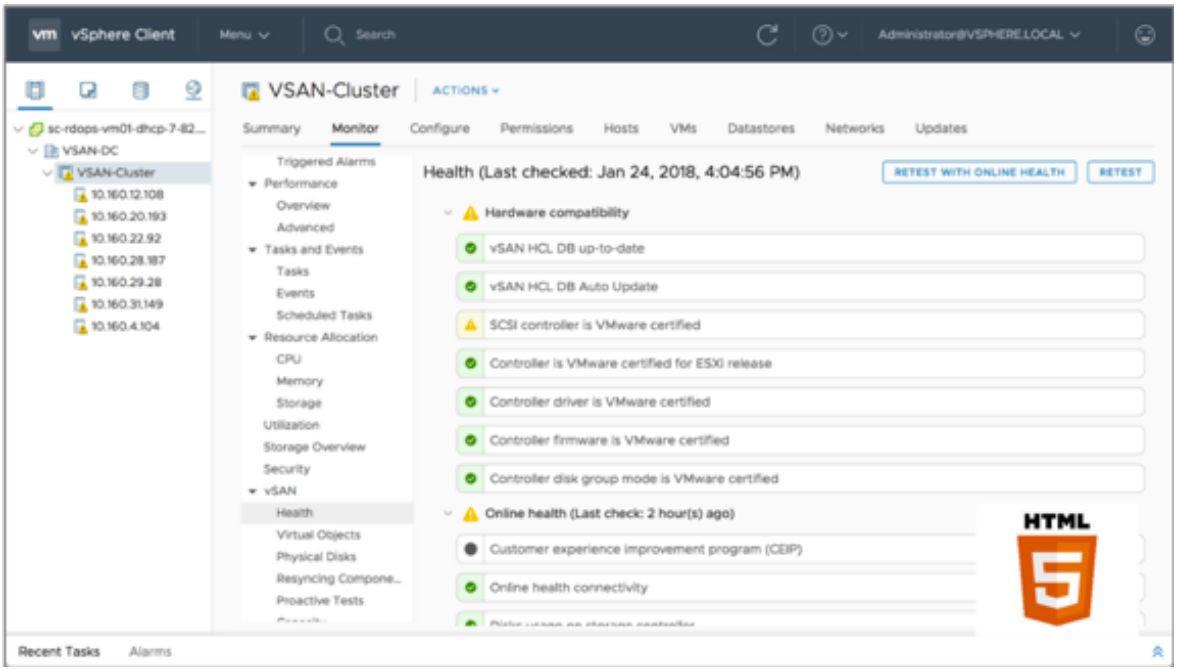
This allows for an administrator to configure the vSAN Inter-site link to use Jumbo Frames, while keeping the witness uplinks going to the more affordable witness site to a more common standard MTU size.

This enhancement gives additional flexibility and choice in allowing for a wider variety of customer topology conditions and reduce potential network issues.

5. Management

5.1 vSphere Client with the Clarity UI

vSAN 6.7 introduces support for the new HTML5 based vSphere Client that is based on the "Clarity" framework seen in other VMware solutions. Some other VMware products that already use the Clarity UI are vRealize Operations and vRealize Log Insight.



The workflows used to manage vSAN in the new vSphere Client for 6.7 have been built from the "ground up" to simplify the overall management experience. Going forward, new features will not be added to the legacy "Flex" based vSphere Web Client. With the release of vSAN 6.7, the vast majority of management tasks are available in the new vSphere Client. In creating these new workflows have been designed to optimize tasks more intuitively and with less clicks.

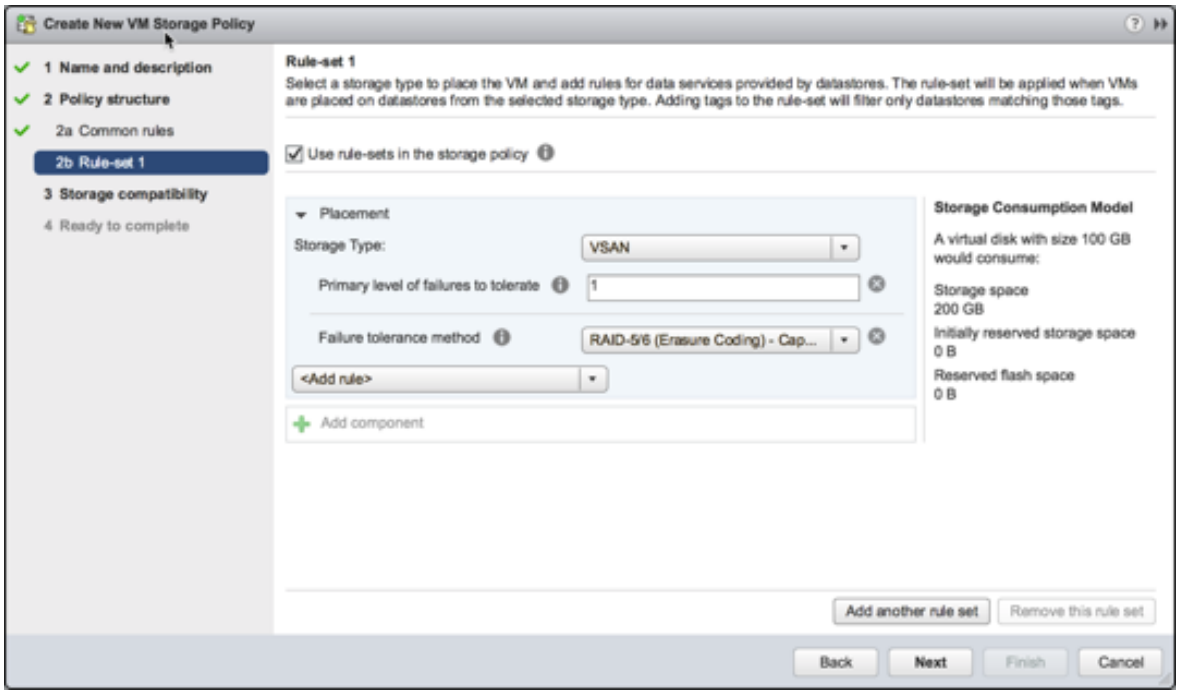
5.2 Storage Policy Based Management

Traditional storage solutions commonly use LUNs or volumes. A LUN or a volume is configured with a specific disk configuration such as RAID to provide a specific level of performance and availability. The challenge with this model is each LUN or volume is confined to providing only one level of service regardless of the workloads that it contains. This leads to provisioning numerous LUNs or volumes to provide the right levels of storage services for various workload requirements. Maintaining many LUNs or volumes increases complexity. Deployment and management of workloads and storage in traditional storage environments are often a manual process that is time-consuming and error prone.

Storage Policy-Based Management (SPBM) from VMware enables precise control of storage services. Like other storage solutions, vSAN provides services such as availability levels, capacity consumption, and stripe widths for performance. A storage policy contains one or more rules that define service levels.

Storage policies can be created and managed using the new vSphere Client, the legacy "Flex" vSphere Web Client, or via PowerCLI/API. Policies can be assigned to virtual machines and individual objects such as a virtual disk. Storage policies are easily changed or reassigned if application requirements change. These modifications are performed with no downtime and without the need to migrate virtual machines from one datastore to another. SPBM makes it possible to assign and modify service levels with precision on a per-virtual machine basis.

The following figure shows storage policy rules created using the vSphere Web Client. This policy contains three rules. The first rule, “Primary level of failures to tolerate," defines how many failures an object can tolerate before it becomes unavailable. The second rule indicates the failure tolerance method that will be used. This policy uses RAID-5/6 erasure coding to minimize capacity consumption. Creating this rule takes roughly 11 steps.



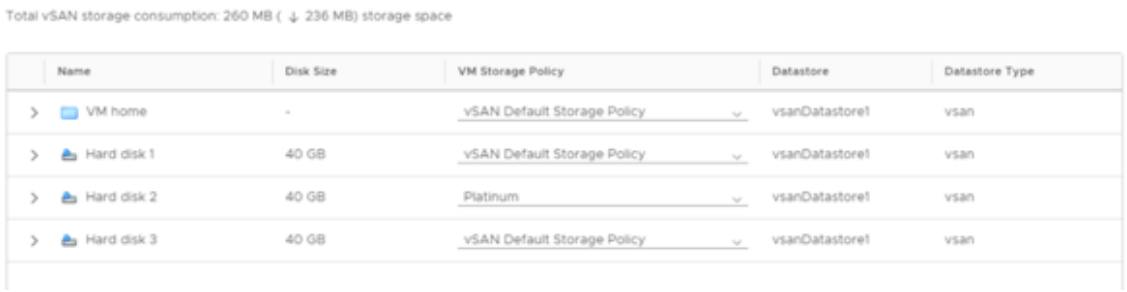
With vSAN 6.7 now supporting the vSphere Client, the same policy can be created with less steps, but it is important to also understand that it is presented a bit differently.

The most common policy choices are protection within a standard cluster and what protection choice. In the vSphere 6.7 client, those would be categorized as Availability rules. The other rules available to

vSAN are categorized as Advanced Policy Rules in the vSphere Client. These other rules such as Object Space Reservation, IOPS Limits, Stripe Width, and so forth are really seldom used for most workloads.

vSAN Storage Policy

Storage policies can be applied to all objects that make up a virtual machine and to individual objects such as a virtual disk. The figure below shows a virtual machine with three virtual disks. One of the virtual disks has the "Platinum" storage policy assigned while the rest of the objects have the default storage policy assigned.



Here is a Storage Policy Based Management demo

Click to see topic media

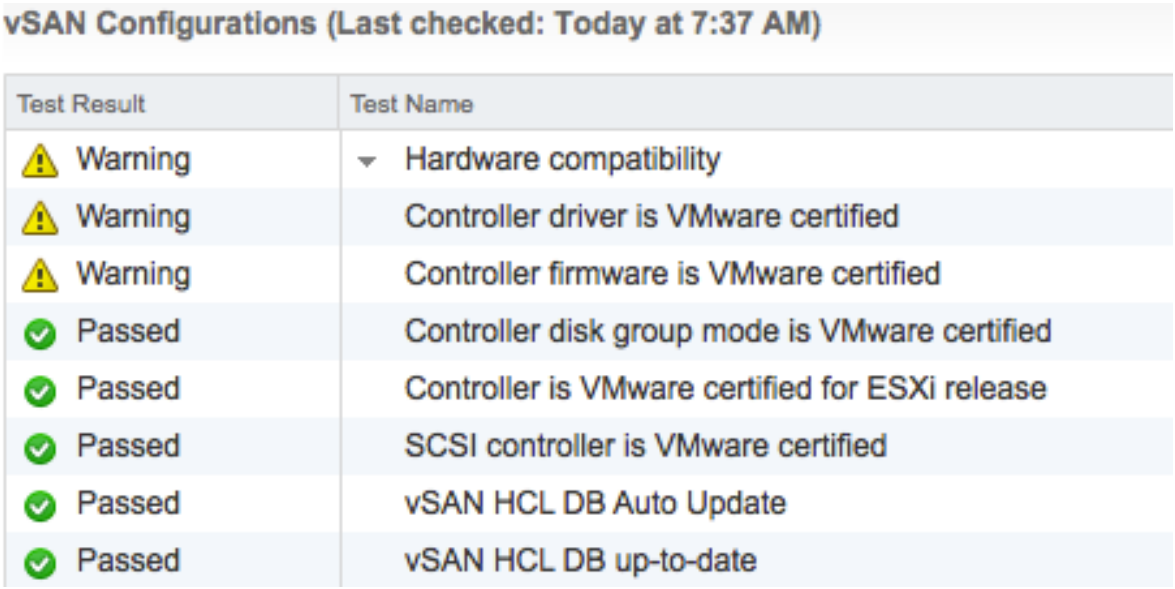
5.3 vSAN Configuration Assist

An important aspect of a healthy vSAN environment is ensuring correct configurations, device firmware, and device drivers. vSAN 6.6 introduced vSAN Configuration Assist to check hardware compatibility, burn-in testing, network configuration, vSAN configuration, and adherence to VMware cluster recommendations.

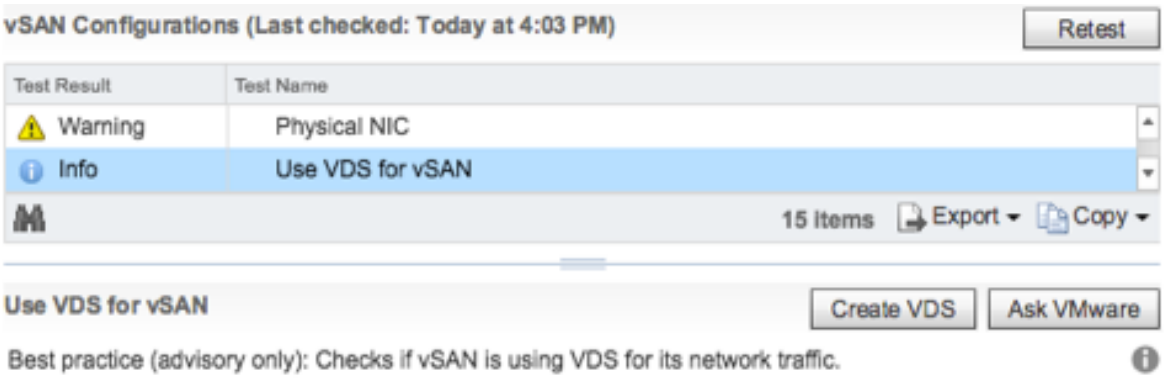
Outdated controller firmware and driver versions are identified and the option to download and install the latest supported software is provided. Driver and firmware upgrades to controller hardware can

now be initiated by a single click and orchestrated across the entire cluster to streamline hardware lifecycle management. This feature eliminates the need for vendor-specific tools. Automatic downloads and notifications reduce management overhead and lower the risk associated with manual processes.

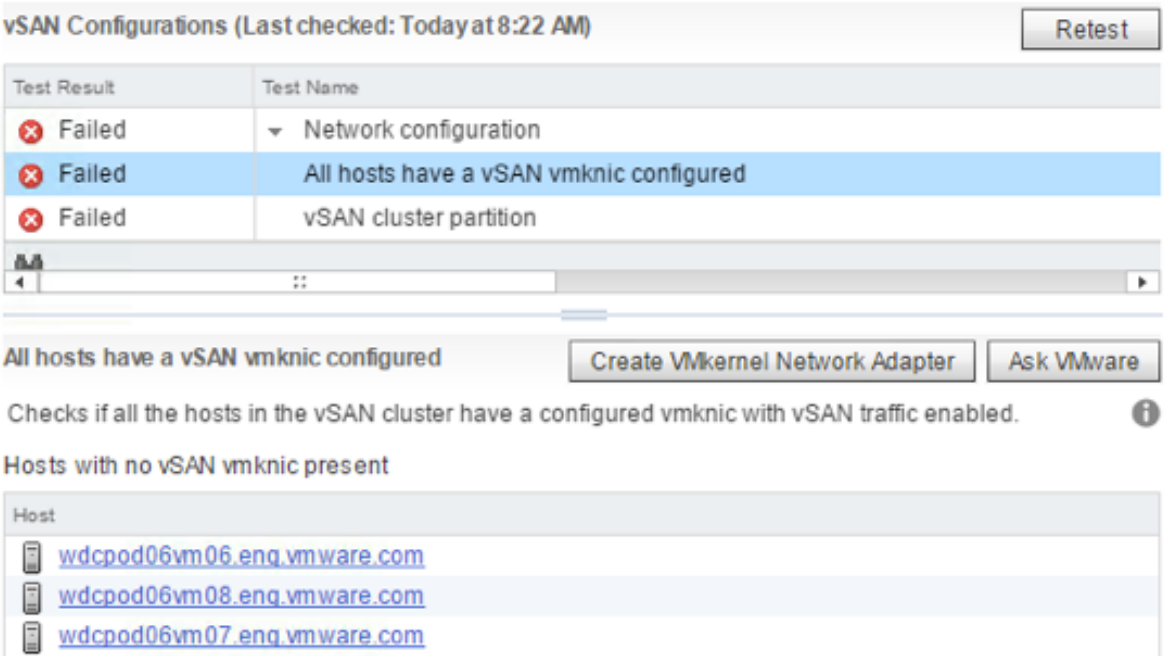
Note: A subset of the OEMs and server models that run vSAN are currently supported. Many do not yet support automated remediation. In these cases, the vSAN Health UI shows items that need attention, but remediation is a manual process.



As example of a recommendation or "best practice" is each vSAN vmknic should be backed by two physical network interface cards (NICs). Configuration Assist will check the configuration of the vSAN vmknic and recommend two NICs if only one is configured. Another example is the recommendation to use a **vSphere Distributed Switch** (VDS) for vSAN. In the figure below, the recommendation to use a VDS is visible and a "Create VDS" button is provided to start the process.



vSAN Configuration Assist simplifies other configuration aspects, as well. For example, the creation of a VMkernel network adapter for vSAN on each host in a vSAN cluster. The figure below shows a 3-Node cluster where the hosts do not have a vSAN vmknic configured. Configuration Assist identifies the issue configuration issue and includes the "Create VMkernel Network Adapter" button. Clicking this button initiates the process of configuring the vSAN vmknics on the vSphere Distributed Switch. This helps ensure proper configuration and consistency across the cluster. If more information is needed, the Ask VMware button opens a relevant VMware knowledge base (KB) article in a new Web browser window.

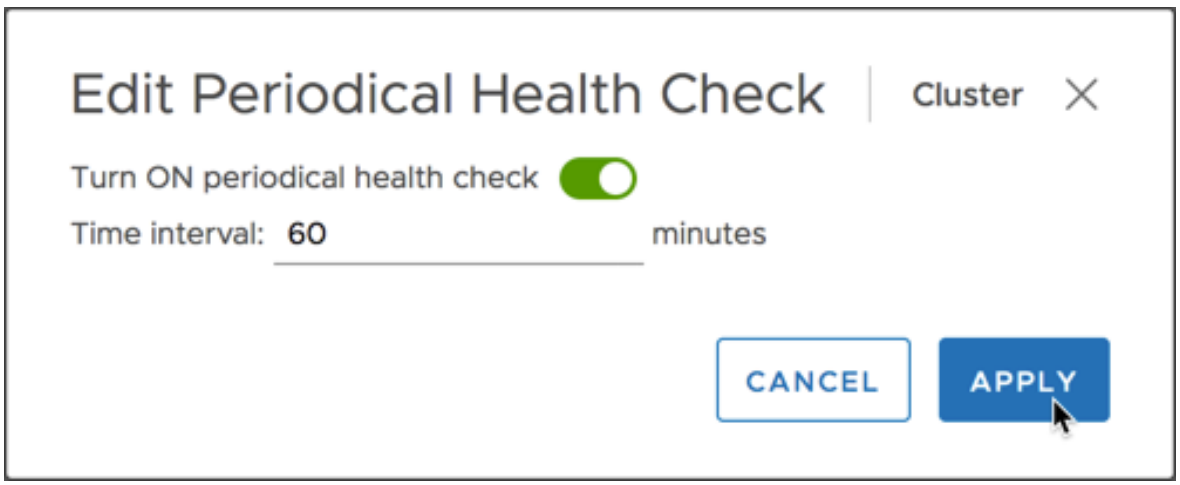


The vSAN Configuration Assist utility is currently only available in the vSphere Web Client (Flex) and may be ported to the new vSphere Client at a later time.

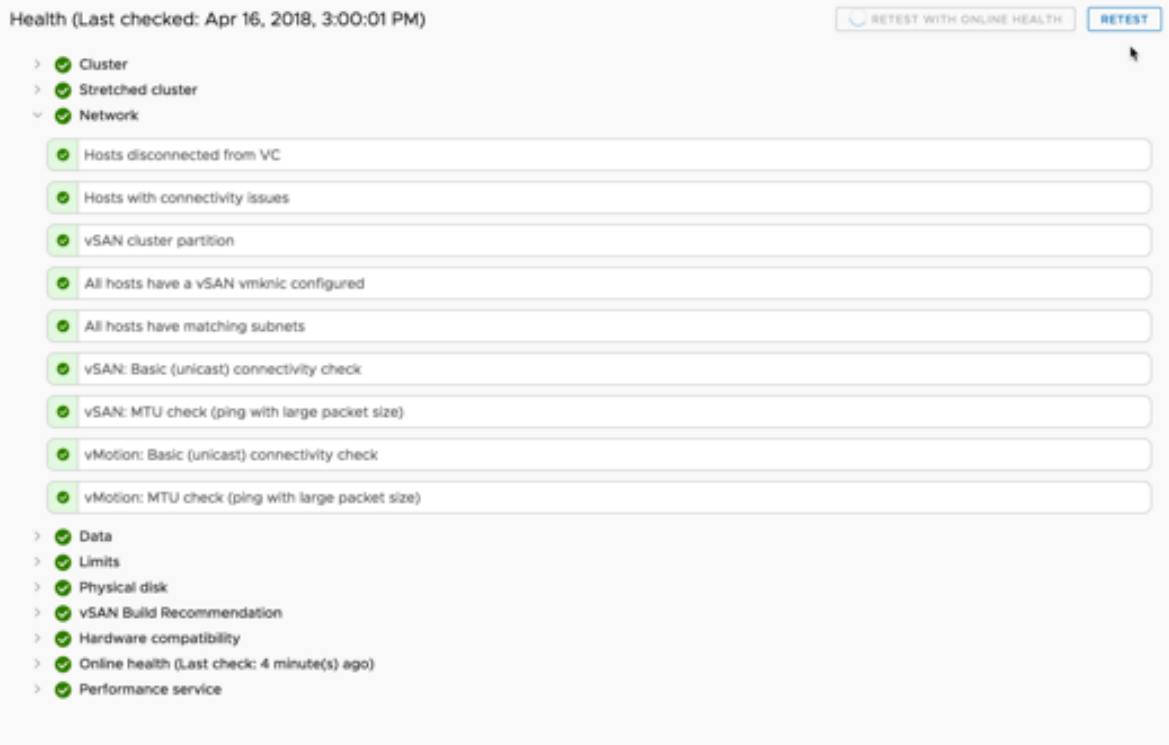
5.4 vSAN Health Check

vSAN features a comprehensive health service appropriately called vSAN Health that actively tests and monitors many items such as hardware compatibility, verification of storage device controllers, controller queue depth, and environmental checks for all-flash and hybrid vSAN configurations. vSAN Health examines network connectivity and throughput, disk and cluster health, and capacity consumption. Proactive monitoring and alerting in vSAN Health helps ensure the environment is optimally configured and functioning properly for the highest levels of performance and availability.

vSAN Health is enabled by default and configured to check the health of the vSAN environment every 60 minutes. The 60-minute time interval is the recommended setting, but this setting can be changed.



vSAN Health is thorough in the number of tests it performs. As an example, there are 10 tests just in the Network section of the vSAN Health UI.

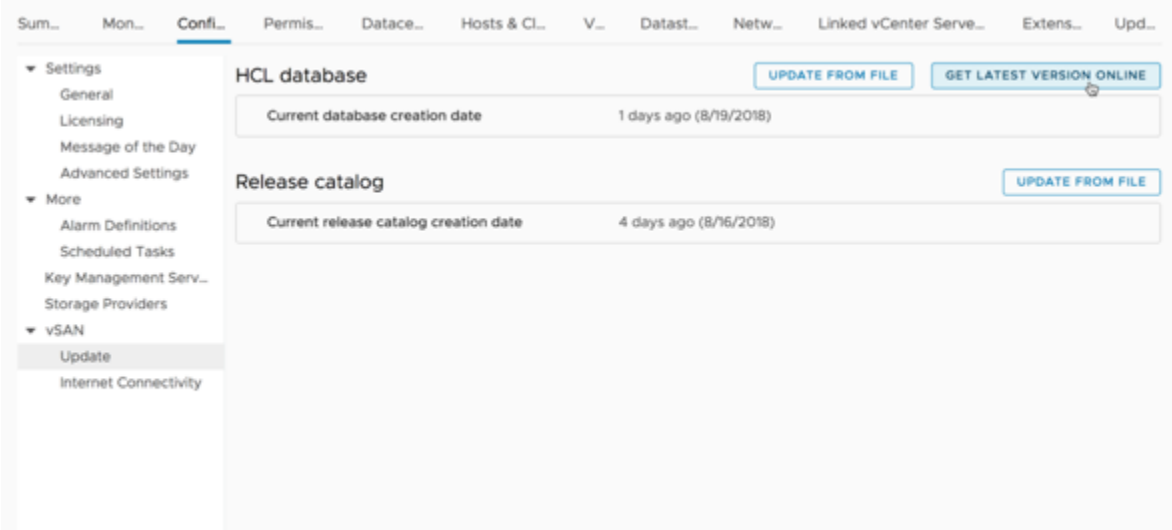


If an issue is detected, a warning is immediately visible in the vSAN UI. Clicking the warning provides more details about the issue. In addition to providing details about the warning, vSAN Health also has an Ask VMware button, which brings up the relevant VMware Knowledge Base article. This simplifies and streamlines remediation efforts.

Hardware Compatibility

vSphere and vSAN support a wide variety of hardware configurations. The list of hardware components and corresponding drivers that are supported with vSAN can be found in the VMware Compatibility Guide. It is very important to use only hardware, firmware, and drivers found in this guide to ensure stability and performance.

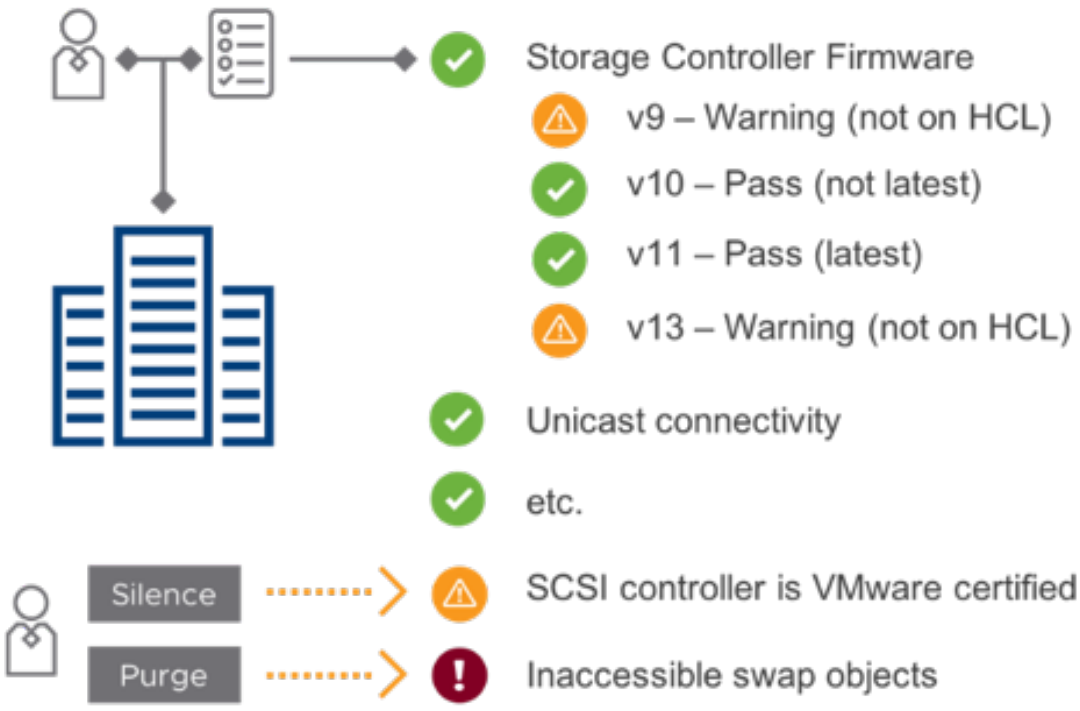
The list of certified hardware, firmware, and drivers is contained in a hardware compatibility list (HCL) database. vSAN makes it easy to update this information. If the environment has Internet connectivity, updates are obtained directly from VMware. Otherwise, HCL database updates can be downloaded for offline use.



If an issue does arise that requires the assistance of VMware Global Support Services, it is easy to upload support bundles to help expedite the troubleshooting process. Clicking the Upload Support Bundles to Service Request button enables you to enter an existing support request (SR) number and easily upload the necessary logs.

Improved Health Check Guidance & Recommendations in vSAN 6.7 U1

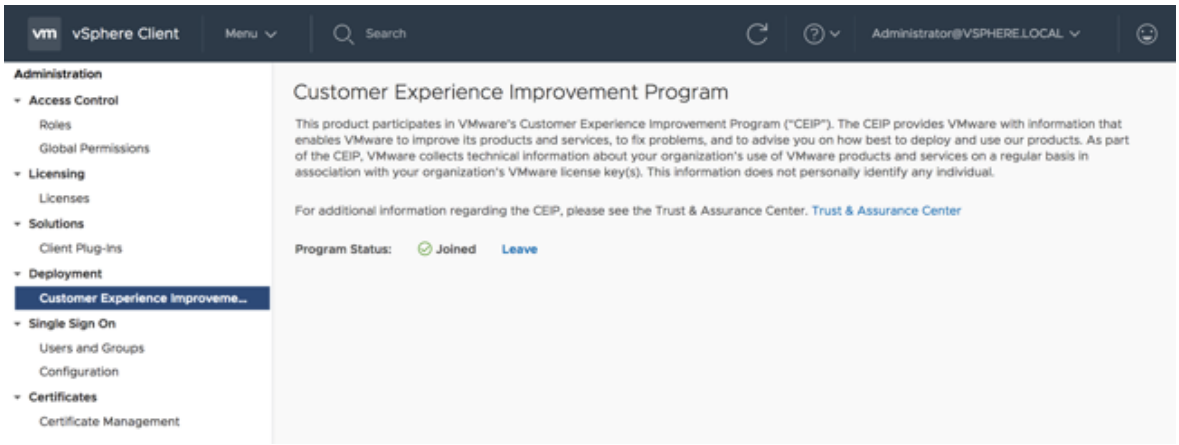
vSAN 6.7 U1 extends this feature even more, with a more robust way of handling multiple approved firmware levels for storage controllers.



A new Unicast network performance health check and test ensures that proper continuity is achieved between

vSAN hosts and will report network bandwidth results for the tests. vSAN 6.7 U1 also introduces functionality that is now accessible in the UI. Health checks can be silenced granularly, directly in the UI, as well as being able to purge inaccessible swap objects that are no longer needed. These improvements improve the effectiveness of vSAN's ability to not only recognize issues but remediate them.

Customer Experience Improvement Program
Participating in the Customer Experience Improvement Program (CEIP) enables VMware to provide higher levels of proactive and reactive customer assistance. Many Benefits of participating in this program include streamlined troubleshooting, real-time notifications and recommendations for your environment, diagnostics, and the potential to remedy issues before they become problems.



More specific information on participation in the VMware Customer Experience Improvement Program (CEIP) can be found here:
https://www.vmware.com/solutions/trustvmware/ceip.html

Proactive Cloud Health Checks
Participating in the Customer Experience Improvement Program (CEIP) enables VMware to provide higher levels of proactive and reactive customer assistance. Benefits of participating in this program include streamlined troubleshooting, real-time notifications and recommendations for your environment, diagnostics, and the potential to remedy issues before they become problems.

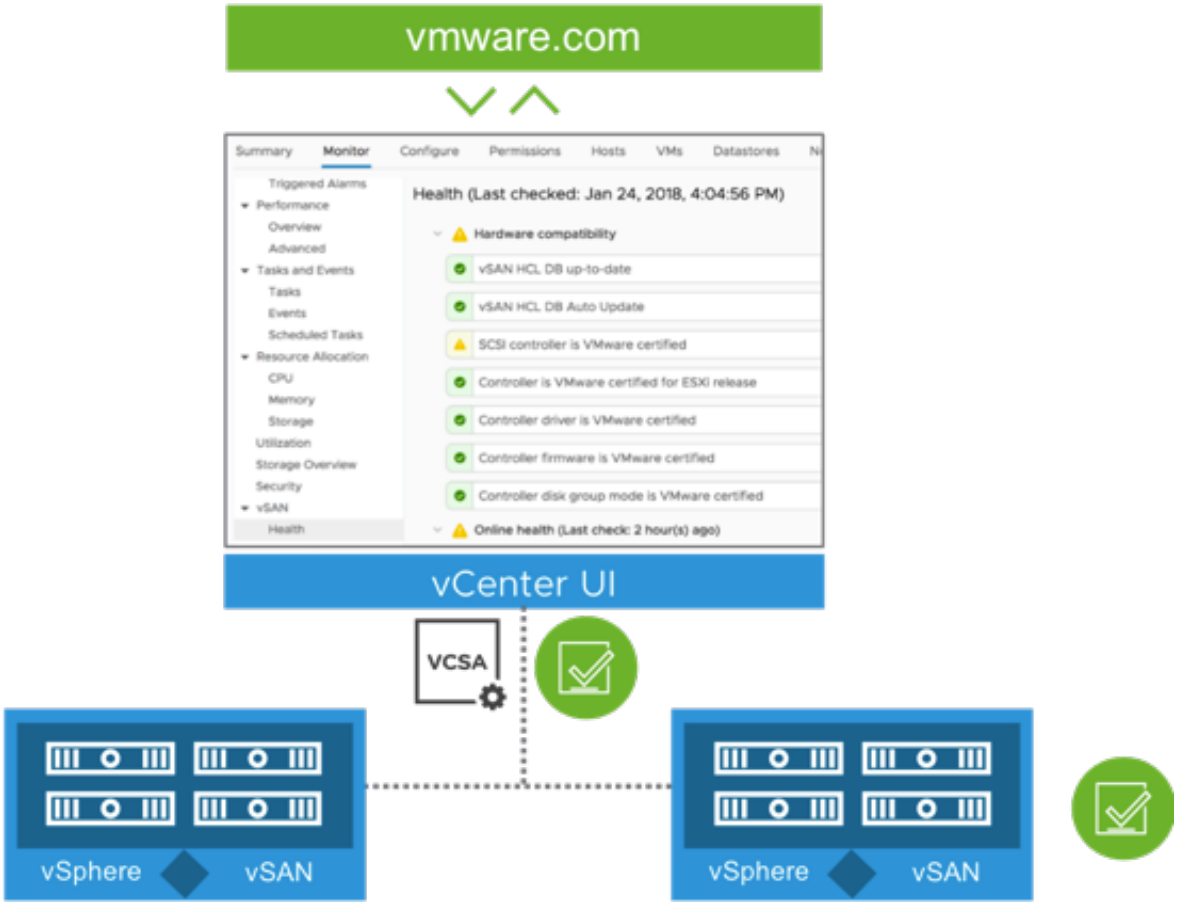
vSAN Health is cloud-connected. New health checks appear as new VMware Knowledge Base (KB) articles are created and published. An "Ask VMware" button is supplied in the user interface, which guides administrators to the relevant VMware knowledge base article.

Online Health Checks are added between product updates. This is beneficial because they are delivered without the need to upgrade vSphere and vSAN. This enhancement consistently provides administrators with latest checks and remedies for optimal reliability and performance.

Some hardware specific health checks will appear only in the event a cluster has particular hardware, while other health checks previously only in Online Health Checks will available locally when CEIP is disabled or vCenter has no Internet accessibility.

A couple tests have been retired, such as the Multicast test due to vSAN 6.6 and higher do not utilize Multicast, as well as the Storage Load test. HCIBench is recommended for performing vSAN benchmarking tests.

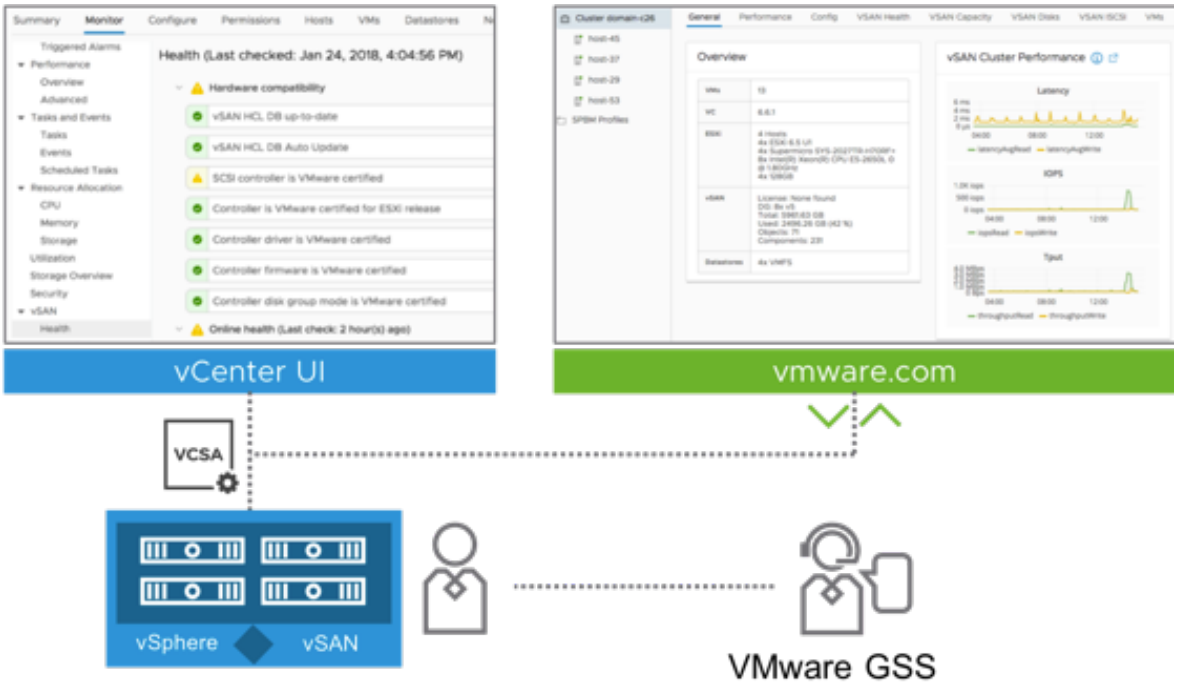
- New health check tests include:
- Host maintenance mode verification
- Host consistency settings for advanced settings
- Improved vSAN and vMotion network connectivity checks
- Improved vSAN Health Service installation check
- Physical Disk Health checks combine multiple checks into a single health check
- Improved HCL check. Firmware checks are now independent of driver checks.



Note: CEIP provides VMware with information that enables VMware to improve its products and services, to fix problems, and to advise you on how best to deploy and use our products. As part of the CEIP, VMware collects technical information about your organization's use of VMware products and services on a regular basis in association with your organization's VMware license key(s). This information does not personally identify any individual. For additional information regarding the CEIP, please see the Trust & Assurance Center.

vSAN Support Insight
Customers that participate in the Customer Experience Improvement Program (CEIP) also get added benefit by proactively providing VMware Global Support Services (GSS) with some data about their environment.

Those that are reluctant to enable CEIP and provide data to GSS can rest assured their data is safe. Cluster data is anonymized and uploaded to VMware's analytics cloud and converted to useful information such as configuration specifics, performance metrics, and health status. GSS has no visibility to information such as host names, virtual machine names, datastore names, IP addresses, SPBM profiles, and more as their actual values are not distinguishable without an obfuscation map that is only available to the cluster owner.



The anonymized data can be associated with ongoing support cases to better equip GSS in the troubleshooting process. Together, the ability to protect sensitive customer information and reducing the amount of time associated with the troubleshooting process, can provide more secure and faster problem resolution.

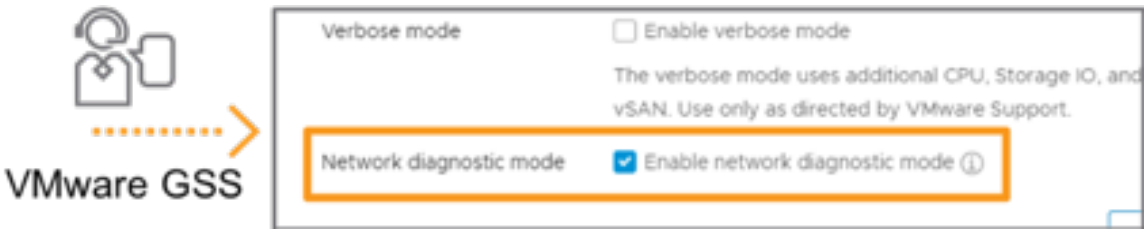
Performance for Support Diagnostics in vSAN 6.7 U1

In addition to vSAN Support Insight, vSAN 6.7 U1 adds new “Performance for Support” dashboards natively within the UI to reduce the need for support bundles and allow VMware Global Support Services for a faster time to resolution.

Data previously available in the vSAN Observer utility is now immediately available to GSS personnel to better support the environment. GSS Support personnel now have better visibility to the current state of vSAN, and in many cases can reduce the need for customer submission of support bundles.

The vSAN Performance Service and vRealize Operations are still the primary locations for administrators to view cluster, host, and virtual machine performance metrics.

Additionally, an on-demand network diagnostics mode can temporarily collect granular network intelligence, that can be submitted with log bundles for deeper visibility at both the host and network layers.



Performance for Support Diagnostics, along with the ability to temporarily collect enhanced network metrics are two great additions to the vSAN toolset to better speed problem resolution.

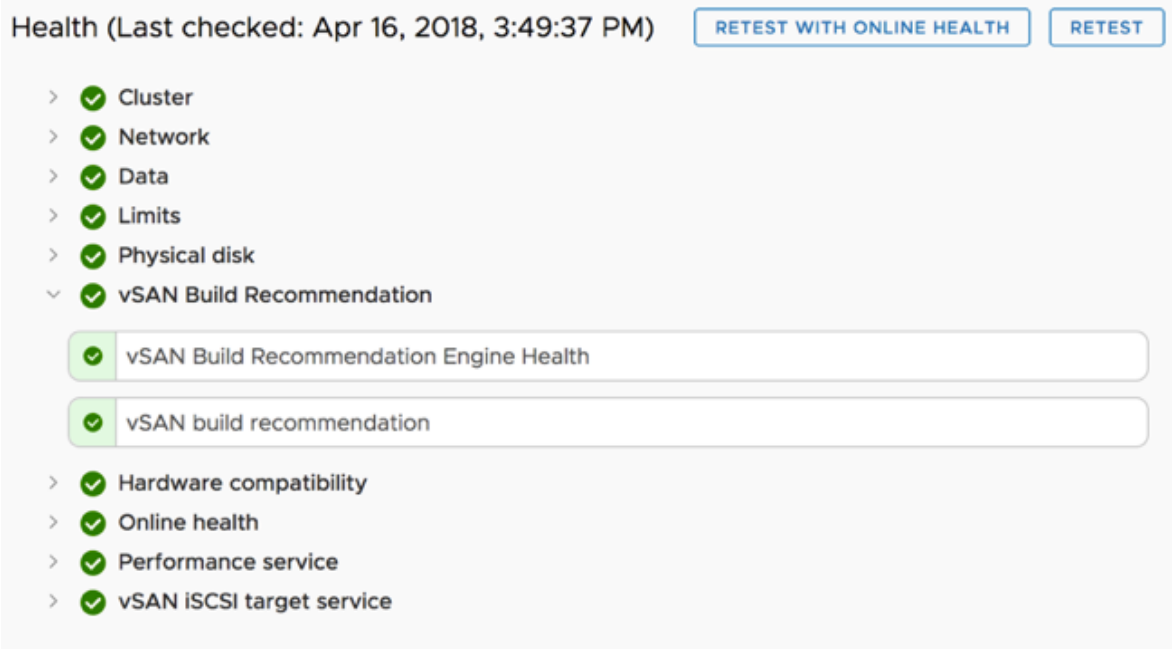
5.5 vSphere Update Manager Integration

vSAN 6.7

vSphere administrators have turned to VMware vSphere Update Manager™ for quite some time now to simplify and automate the patching and upgrading of vSphere clusters. In previous versions with vSAN, admins had to do a bit of research and perform manual steps before upgrading a vSAN cluster.

The primary concern was verifying hardware compatibility (SAS and SATA controllers, NVMe devices, etc.) with the new version of vSphere and vSAN. This was a manual process of checking the VMware Compatibility Guide to ensure the upgraded version was supported with the hardware deployed in that vSAN cluster.

Those concerns and manual steps have been eliminated and automated. vSphere Update Manager generates automated build recommendations for vSAN clusters. Information in the VMware Compatibility Guide and vSAN Release Catalog is combined with information about the currently installed ESXi release. The vSAN Release Catalog maintains information about available releases, preference order for releases, and critical patches needed for each release. It is hosted on the VMware Cloud. When a new, compatible update becomes available, a notification is proactively displayed in vSAN Health. This eliminates the manual effort of researching and correlating information from various sources to determine the best release for an environment.

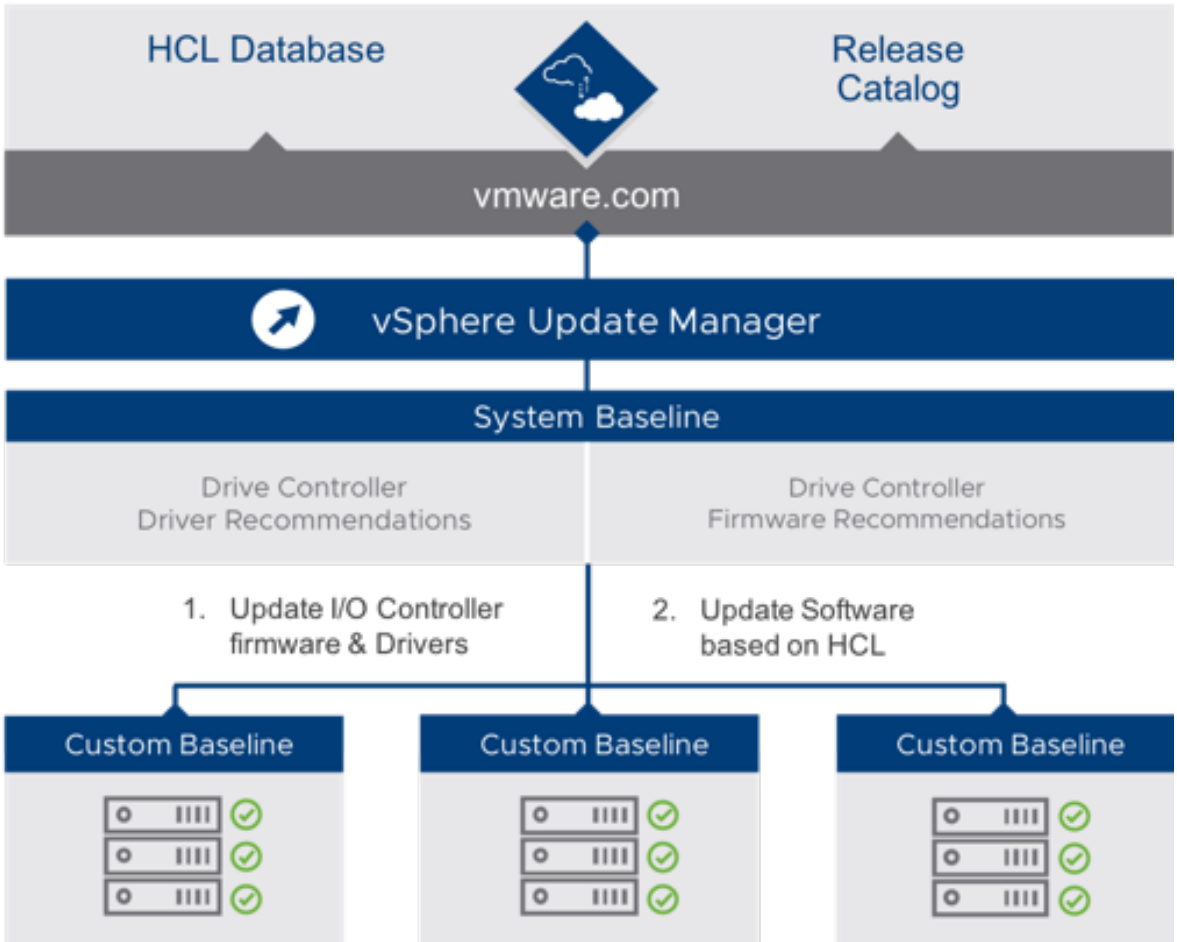


The first step to enable this functionality is entering valid My VMware Portal credentials. vSAN Health produces a warning if credentials have not been entered. A vSAN Cluster baseline group is created using information from the VMware Compatibility Guide and underlying hardware configuration. It is automatically attached to the vSAN cluster.

The functionality that vSphere Update Manager provides compliments vSAN Configuration Assist in vSAN 6.7. Configuration Assist evaluates the environment and verifies the hardware is certified and compatible with the software currently in use. vSphere Update Manager provides notification of newer software updates certified for use with that same hardware.

VUM Driver & Firmware Updating in vSAN 6.7 U1

Updated in vSAN 6.7 U1, all ESXi, driver, and firmware update functions previously handled by the Configuration Assist workflow have been moved to vSphere Update Manager.

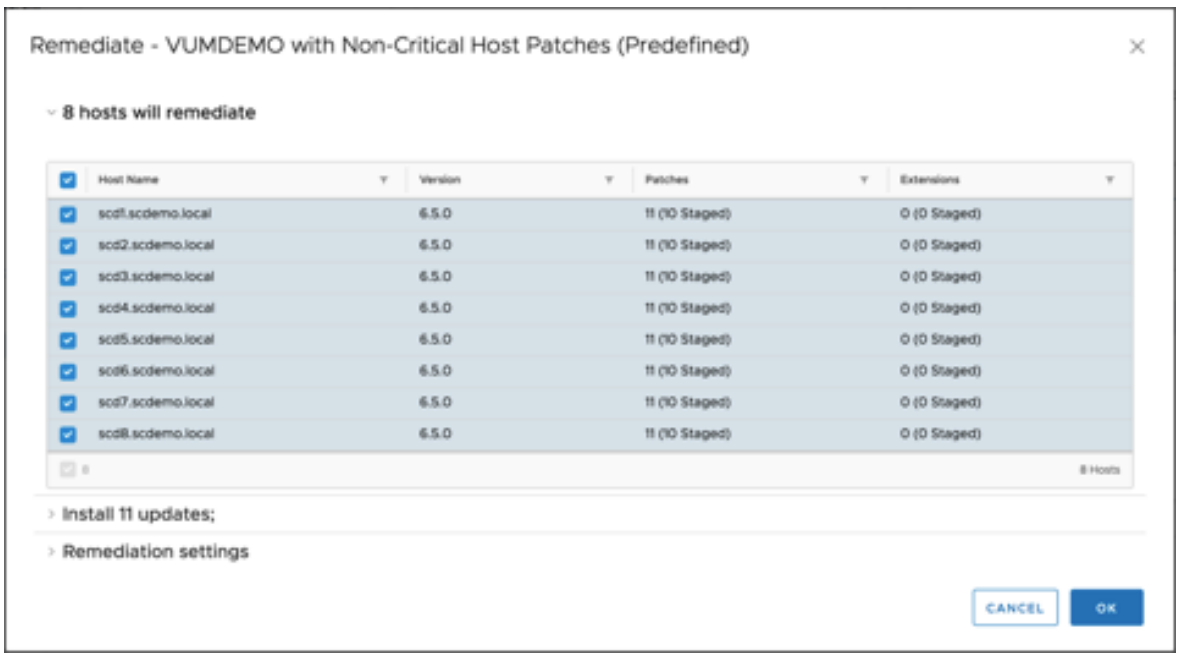


Specific OEM builds can be supported in vSphere Update Manager for 6.7 U1 because it will support the use of OEM vendor ISOs. For those needing to update environments that do not have Internet connectivity, new workflows provide guidance for adding updates to the network isolate environment.

A list of IO Controllers that Support Firmware updating can be found in KB Article 60382 (https://kb.vmware.com/kb/60382)

Updating with vSphere Update Manager

Updates are easily accomplished using baselines that are attached to hosts or clusters to maintain compliance. The upgrade process is automated for hosts that do not comply with the baseline. ISO images and drivers are downloaded automatically. Simply use the Remediate option in vSphere Update Manager to perform a rolling upgrade of the cluster.



vSphere Update Manager migrates virtual machines from the host being upgraded to other hosts in the cluster with no downtime. The host is then placed into maintenance mode and upgraded. The host is restarted if needed. Upon completion of the upgrade, the host exits maintenance mode and vSphere Update Manager starts the same process on the next host. This cycle continues until all hosts in the cluster are compliant with the baseline.

Per host updates in vSAN 6.7 U1 add intelligence to the update process. Updates are validated to ensure deployment and several aspects of host health before allowing a host to exit maintenance mode, giving administrators the opportunity to address the situation before allowing the upgrade process to continue to the next host.

vSphere Update Manager makes it much easier to determine the latest supported versions of vSphere and vSAN for an environment; reduces risk by helping to ensure every host in a cluster is running the same build; and eliminates the time-consuming process of manually upgrading each host.

5.6 Support and Troubleshooting

vSAN ReadyCare

Organizations rely on vendor support more than ever to maintain data center performance through knowledgeable and accessible teams, robust professional services, and advanced analytics and artificial intelligence (AI) technologies. With the launch of vSAN 6.7, VMware is introducing vSAN ReadyCare to describe the focused investments and innovations being made into vSAN support. vSAN ReadyCare represents the combined investments the vSAN team is making in people, professional services, and technology to provide a superior vSAN support experience. This is similar to vSAN ReadyLabs, which focus on certifications and new technologies.

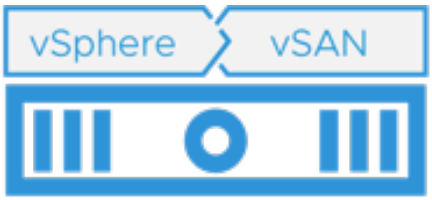
On the people and services side, VMware more than doubled the number of support staff dedicated to vSAN over the past year and has rolled out a host of new professional services, including remote services to provide simple, efficient help in getting started for customers that want a little extra guidance and assurance.

On the technology front, vSAN has invested significantly in delivering new capabilities that can identify potential issues and recommend solutions before a support request is required. vSAN 6.7 introduces over a dozen new health checks, raising the total to over fifty. vSAN also enhances the predictive modelling and other advanced analytical tools to identify common issues across thousands

of vSAN deployments. Utilizing this information, VMware can push proactive alerts to customers and arm the VMware support teams with real-time information to accelerate troubleshooting and further enhance a customer's support experience.

Here is a look at some of the innovations and capabilities available today with vSAN helping us deliver the enhanced support experience of vSAN ReadyCare:

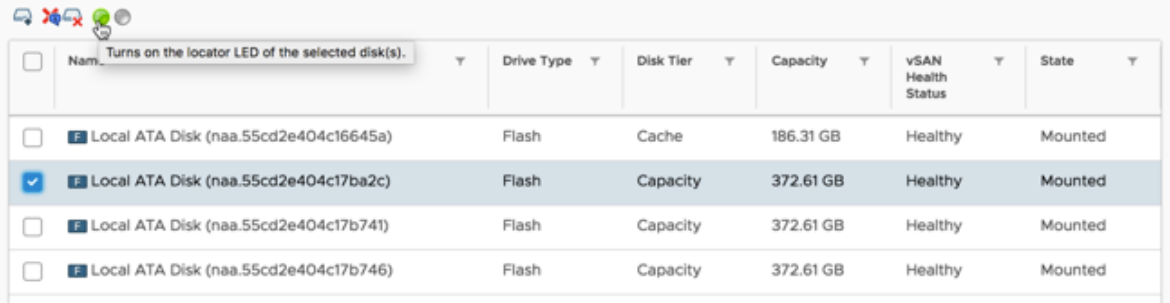
Enhanced Diagnostics Partition Support
The information stored on diagnostic partitions are can be a critical factor in being able to properly diagnose an issue. The ESXi diagnostics partition, known as the coredump is created by default when using traditional storage devices. While the default ESXi coredump size is sufficient for traditional vSphere installations, there is added benefit with a larger coredump when using vSAN.



Extending the coredump has previously been accomplished using the steps in KB article 2147881 using scripting and some interaction from an administrator.

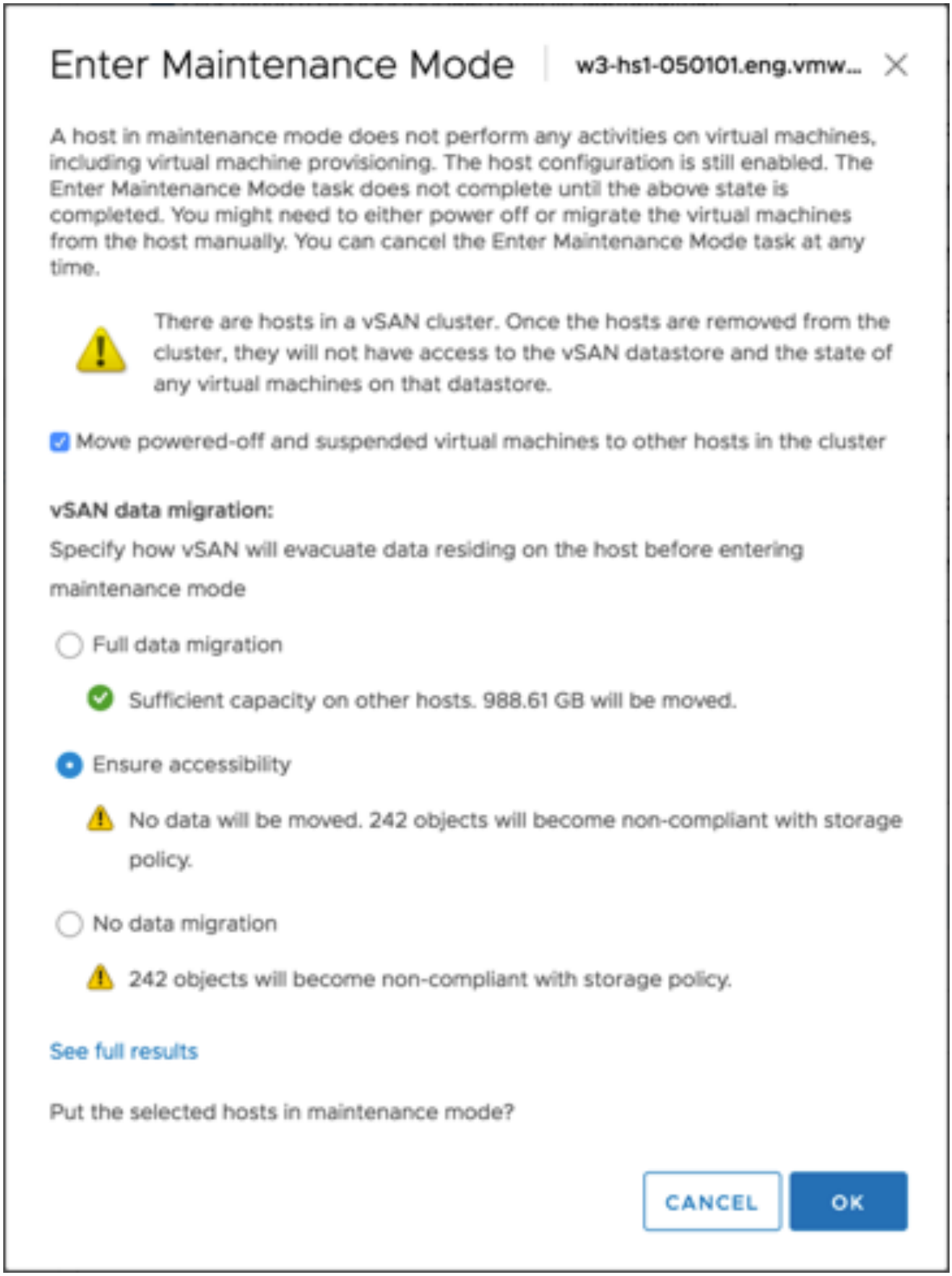
Administrators no longer have to manually perform this process with vSAN 6.7. This is because vSAN 6.7 now brings more intelligence to this matter, by automatically resizing the coredump partitions on USB or SSD media if there is free space on the device.

Storage Device Serviceability
To help simplify operational tasks like hardware maintenance, vSphere includes plugins that provide extended management and information support for drives connected to local storage controllers. This is especially beneficial in vSAN environments. Information such as drive location in the server chassis is provided through these plugins to ease administration and maintenance. Storage Device Serviceability also provides a method for turning on storage device LEDs to assist with locating drives. This was previously available only for controllers in RAID mode. It is now supported for controllers in HBA or "pass-through" mode.



Note: This capability is currently supported in HPE DL and ML series servers with Gen9 storage controllers.

Host Maintenance Mode



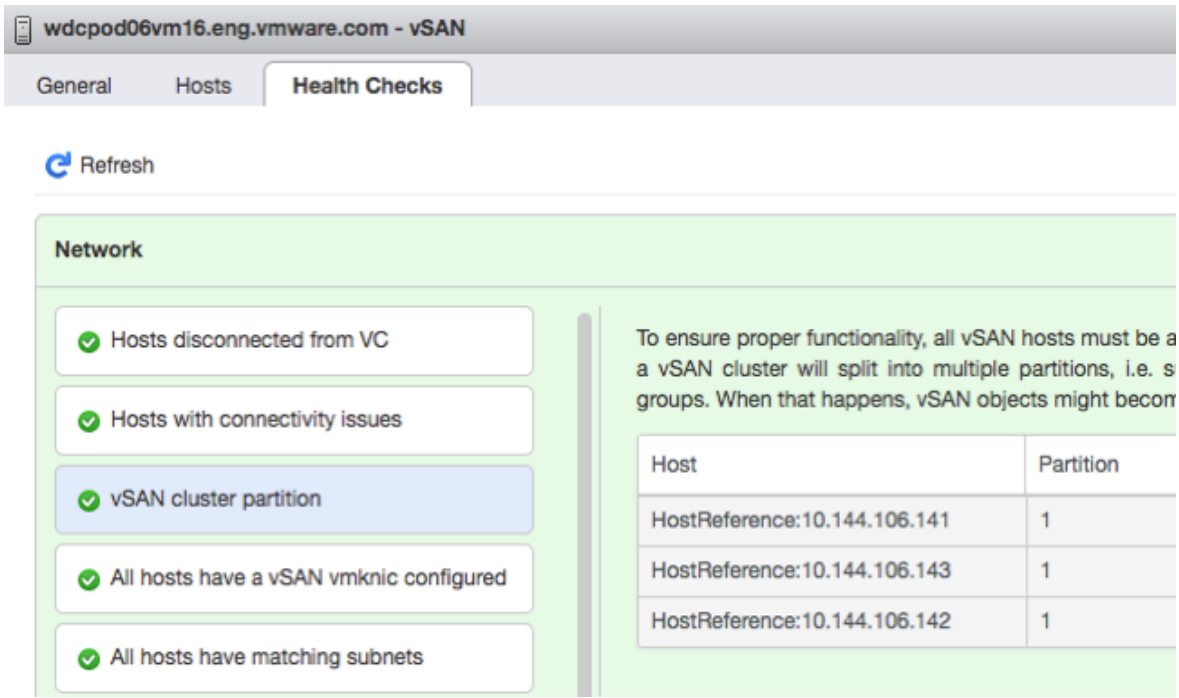
A maintenance mode pre-check is included in vSAN to help ensure adequate capacity remains after a host is evacuated. This function is also used when removing a disk or disk group.

Conducting this pre-check prior to evacuating a host provides a better understanding of how the cluster will be impacted from a capacity standpoint.

Changes in storage policy compliance are also indicated. This is one more example of how vSAN reduces complexity, minimizes risk, and lowers operational overhead.

Highly Available Control Plane for Health Checks
Previous versions of vSAN required vCenter Server and the vSphere Web Client server (part of vCenter Server) to be online to check the health of the cluster. vSAN 6.6 included the ability to perform vSAN health checks using the VMware Host Client in the rare event vCenter Server is offline.

Hosts in a vSAN cluster cooperate in a distributed fashion to check the health of the entire cluster. Any host in the cluster can be used to view vSAN Health. This provides redundancy for the vSAN Health data to help ensure administrators always have this information available. The figure below shows the Network – vSAN Cluster Partition health.

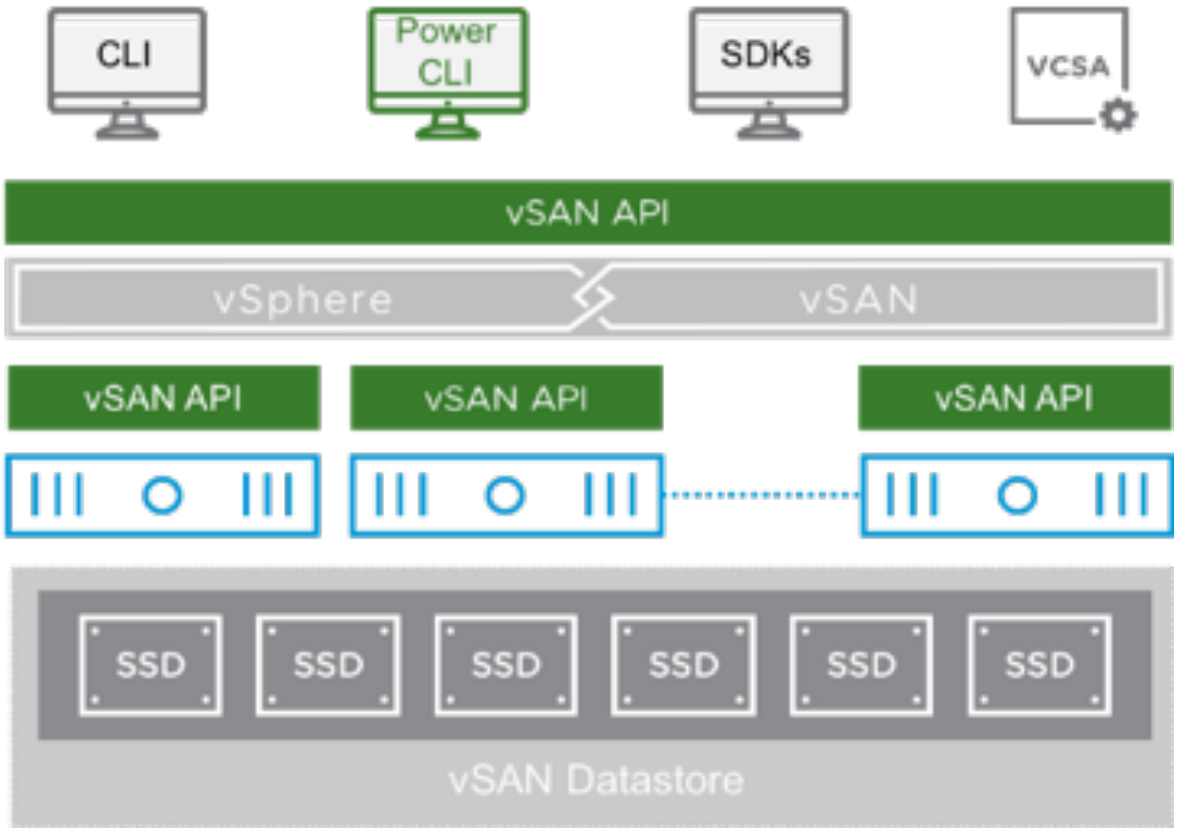


Command line functionality is keeping pace with information available in the vSphere Web Client and VMware Host Client graphical user interfaces. Administrators can use esxcli vsan commands to check vSAN health, perform debugging, and manage configurations for items such as fault domains, storage policies, and iSCSI targets.

```
Usage: esxcli vsan {cmd} [cmd options]

Available Namespaces:
  cluster               Commands for vSAN host cluster configuration
  datastore             Commands for vSAN datastore configuration
  debug                 Commands for vSAN debugging
  health                Commands for vSAN Health
  iscsi                 Commands for vSAN iSCSI target configuration
  network               Commands for vSAN host network configuration
  resync                Commands for vSAN resync configuration
  storage               Commands for vSAN physical storage configuration
  faultdomain           Commands for vSAN fault domain configuration
  maintenancemode       Commands for vSAN maintenance mode operation
  policy                Commands for vSAN storage policy configuration
  trace                 Commands for vSAN trace configuration
```

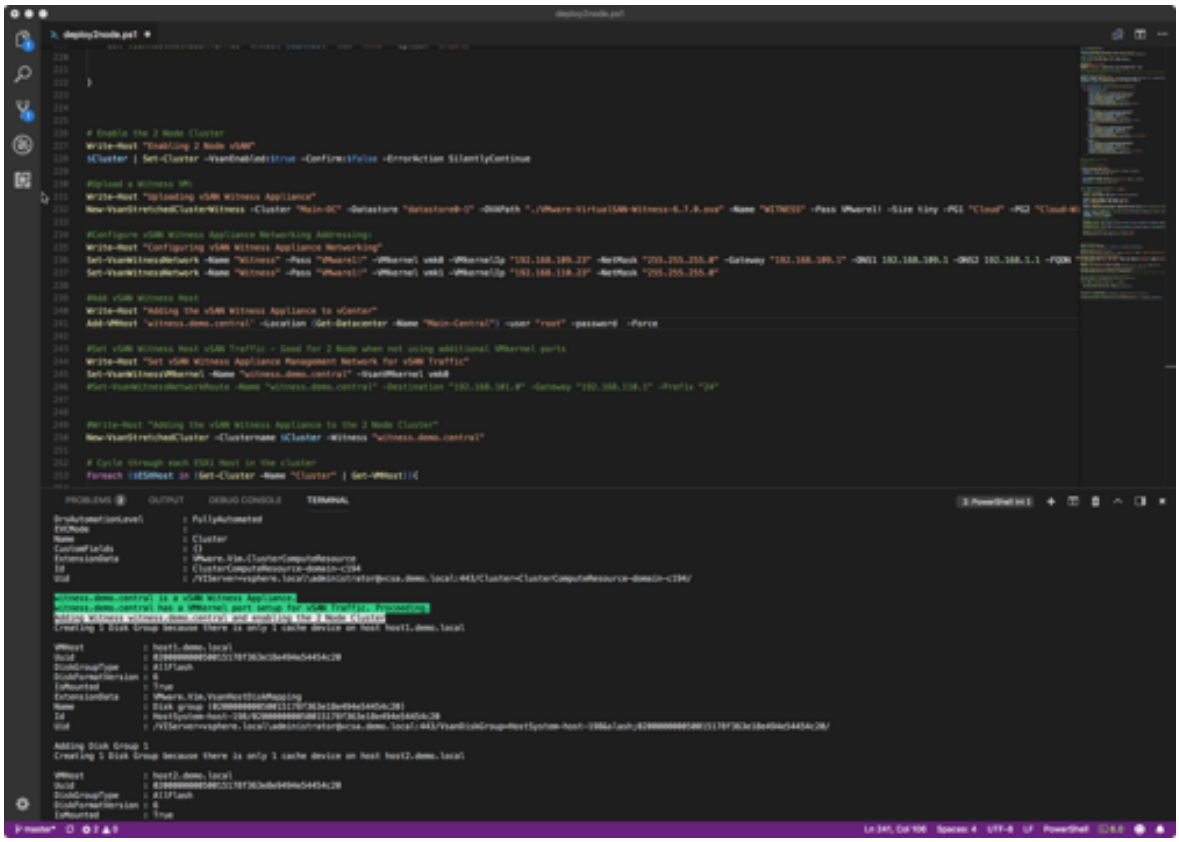
Automation
vSAN features an extensive management API and multiple software development kits (SDKs) to provide IT organizations options for rapid provisioning and automation. Administrators and developers can orchestrate all aspects of installation, configuration, lifecycle management, monitoring, and troubleshooting in vSAN environments.



VMware PowerCLI is one of the most widely adopted extensions to the PowerShell framework. VMware PowerCLI includes a very comprehensive set of functions that abstract the vSphere API down to simple and powerful cmdlets including many for vSAN. This makes it easy to automate several

actions from enabling vSAN to deployment and configuration of a vSAN stretched cluster. Here are a few simple examples of what can be accomplished with vSAN and PowerCLI:

- Assigning a Storage Policy to Multiple VMs
- Set Sparse Virtual Swap Files
- vSAN Encryption Rekey
- Setup vSAN Stretched Cluster DRS Rules
- Backup or Restore SPBM Profiles in VCSA



PowerCLI includes cmdlets for performance monitoring, some cluster configuration information tasks such as upgrades, and vSAN iSCSI operations. A Host-level API can query cluster-level information. S.M.A.R.T. device data can also be obtained through the vSAN API.

SDKs are available for several programming languages including .NET, Perl, and Python. The SDKs are available for download from VMware Developer Center and include libraries, documentation, and code samples.

vSphere administrators and DevOps shops can utilize these SDKs and PowerCLI cmdlets to lower costs by enforcing standards, streamlining operations, and enabling automation for vSphere and vSAN environments.

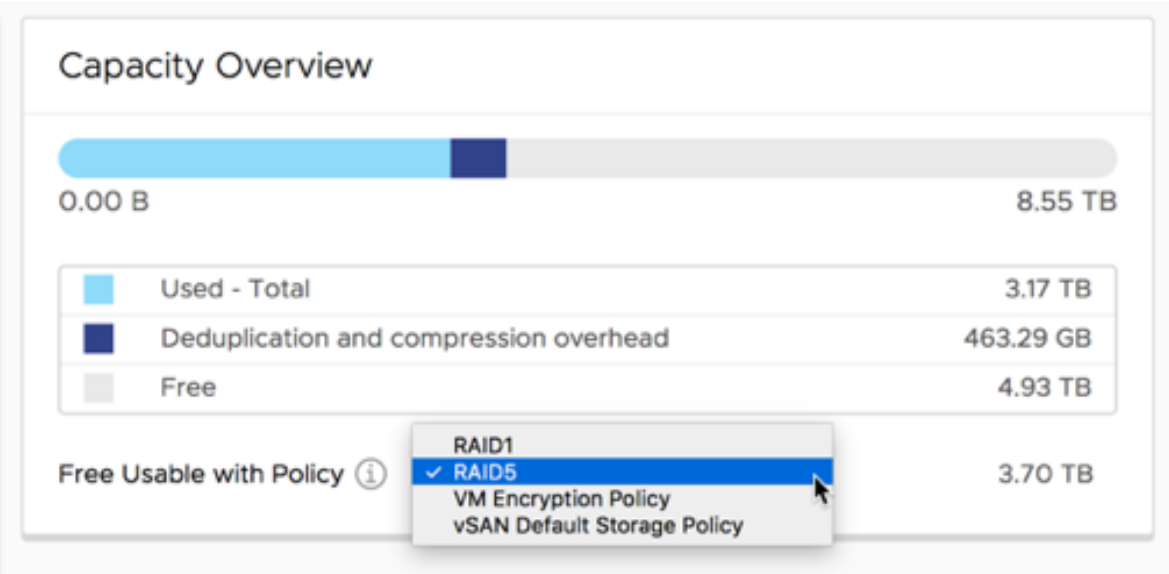
6. Monitoring

6.1 Capacity Reporting

Capacity utilization is a top of mind concern for administrators as part of day-to-day administration of any storage infrastructure.

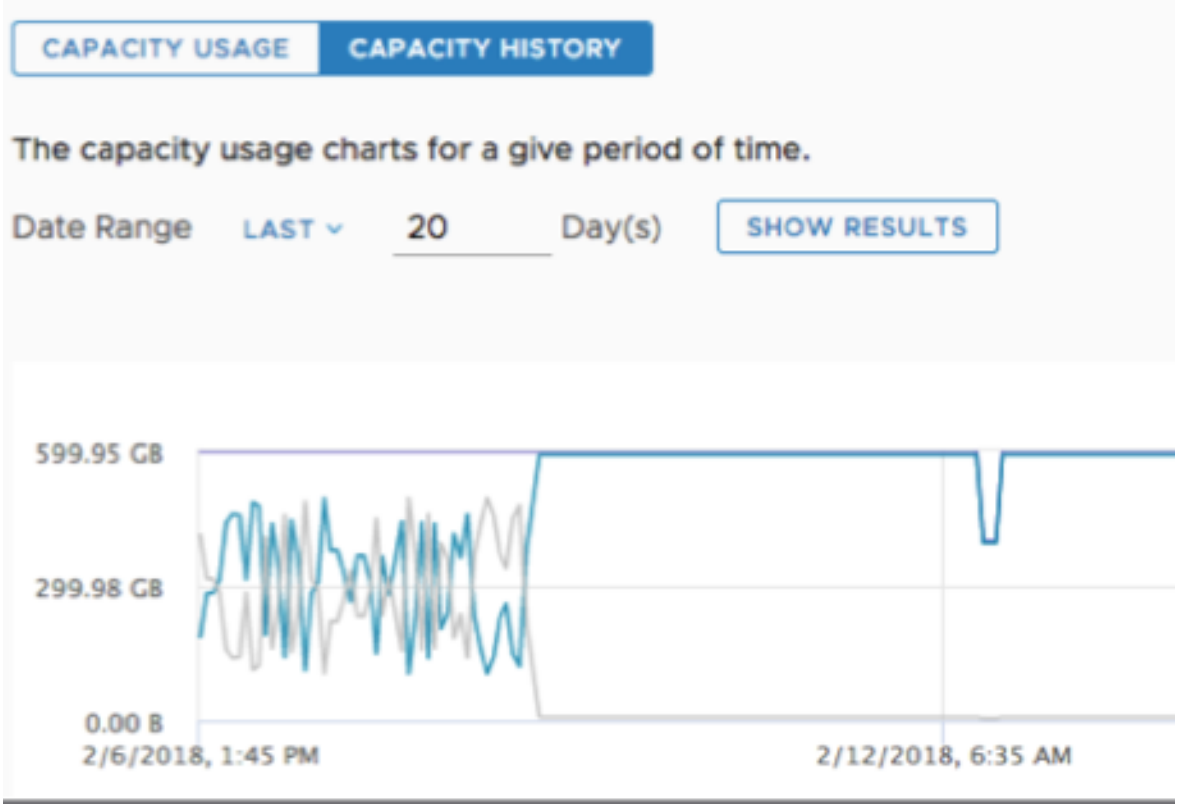
Tools that provide easy access to how much capacity is being consumed and by what type of data, the available usable capacity going forward with different policy choices, as well as the historical change in storage consumption are key to successful storage management.

The vSAN Capacity Overview quickly and easily shows the amount of capacity being consumed on a vSAN datastore.

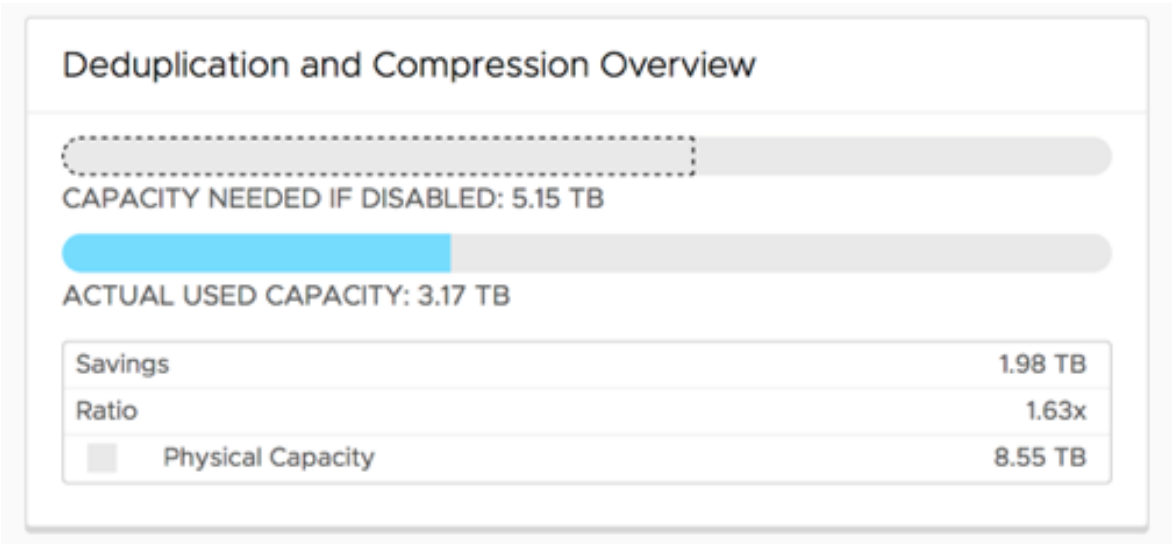


New in vSAN 6.7 U1, administrators have the ability to estimate the amount of usable capacity based on a desired storage policy.

Also new in vSAN 6.7 U1, administrators can see the Capacity History of the vSAN Cluster over time. Historical metrics such as changes in capacity used/free and deduplication ratios, can provide administrators better understanding of how changes in the environment impact storage consumption. and smarter decision making.



For deployments that have Deduplication and Compression enabled, the Deduplication and Compression Overview provides detailed information about the space consolidation ratio and the amount of capacity that has been saved.



New in vSAN 6.7 U1, the Deduplication and Compression Overview shows the amount of space that would be required if Deduplication and Compression were to be disabled.

The Used Capacity Breakdown provides more details on the distribution of object types.

Note: Percentages are of used capacity, not of total capacity.

This following list provides more details on the object types in the Used Capacity Breakdown chart.

- Virtual disks: Virtual disk consumption before deduplication and compression
- VM home objects: VM home object consumption before deduplication and compression
- Swap objects: Capacity utilized by virtual machine swap files

Performance management objects: When the vSAN performance service is enabled, this is the amount of capacity used to store the performance data

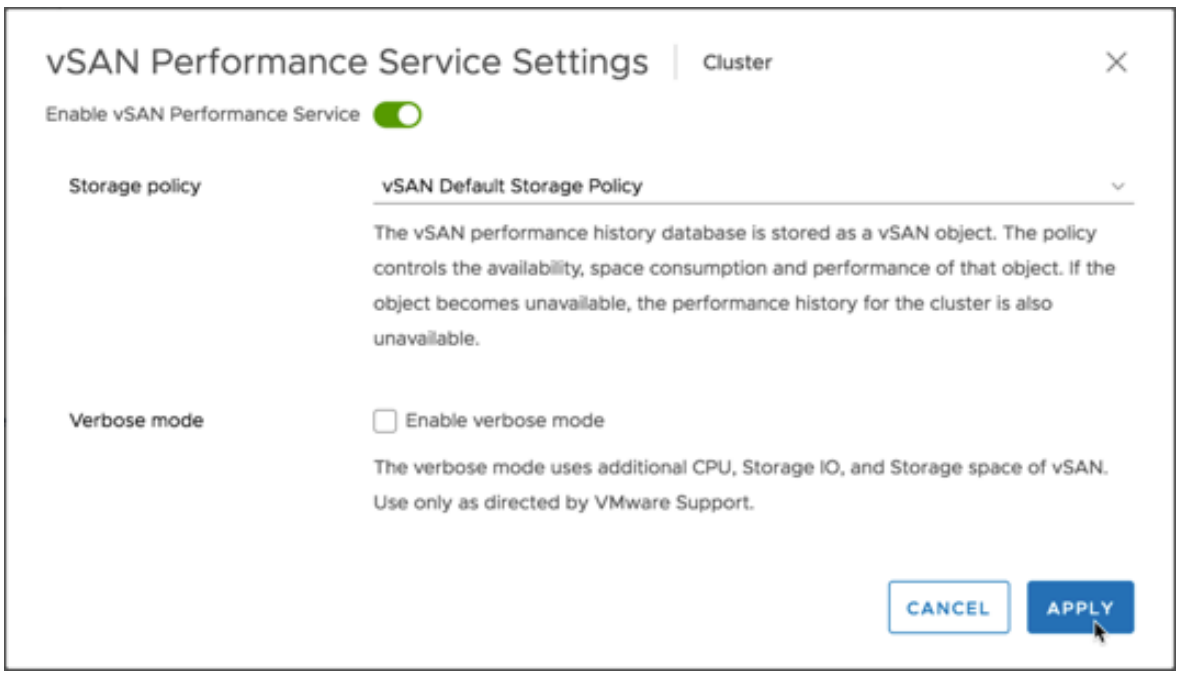
- File system overhead: Capacity required by the vSAN file system metadata
- Deduplication and compression overhead: deduplication and compression metadata, such as hash, translation, and allocation maps
- Checksum overhead: Capacity used to store checksum information
- Other: Virtual machine templates, unregistered virtual machines, ISO files, and so on that are consuming vSAN capacity

6.2 Performance Service

A healthy vSAN environment is one that is performing well. vSAN includes many graphs that provide performance information at the cluster, host, network adapter, virtual machine, and virtual disk levels.

There are many data points that can be viewed such as IOPS, throughput, latency, packet loss rate, write buffer free percentage, cache de-stage rate, and congestion. Time range can be modified to show information from the last 1-24 hours or a custom date and time range. It is also possible to save performance data for later viewing.

The performance service is enabled at the cluster level. The performance service database is stored as a vSAN object independent of vCenter Server. A storage policy is assigned to the object to control space consumption and availability of that object. If it becomes unavailable, performance history for the cluster cannot be viewed until access to the object is restored.



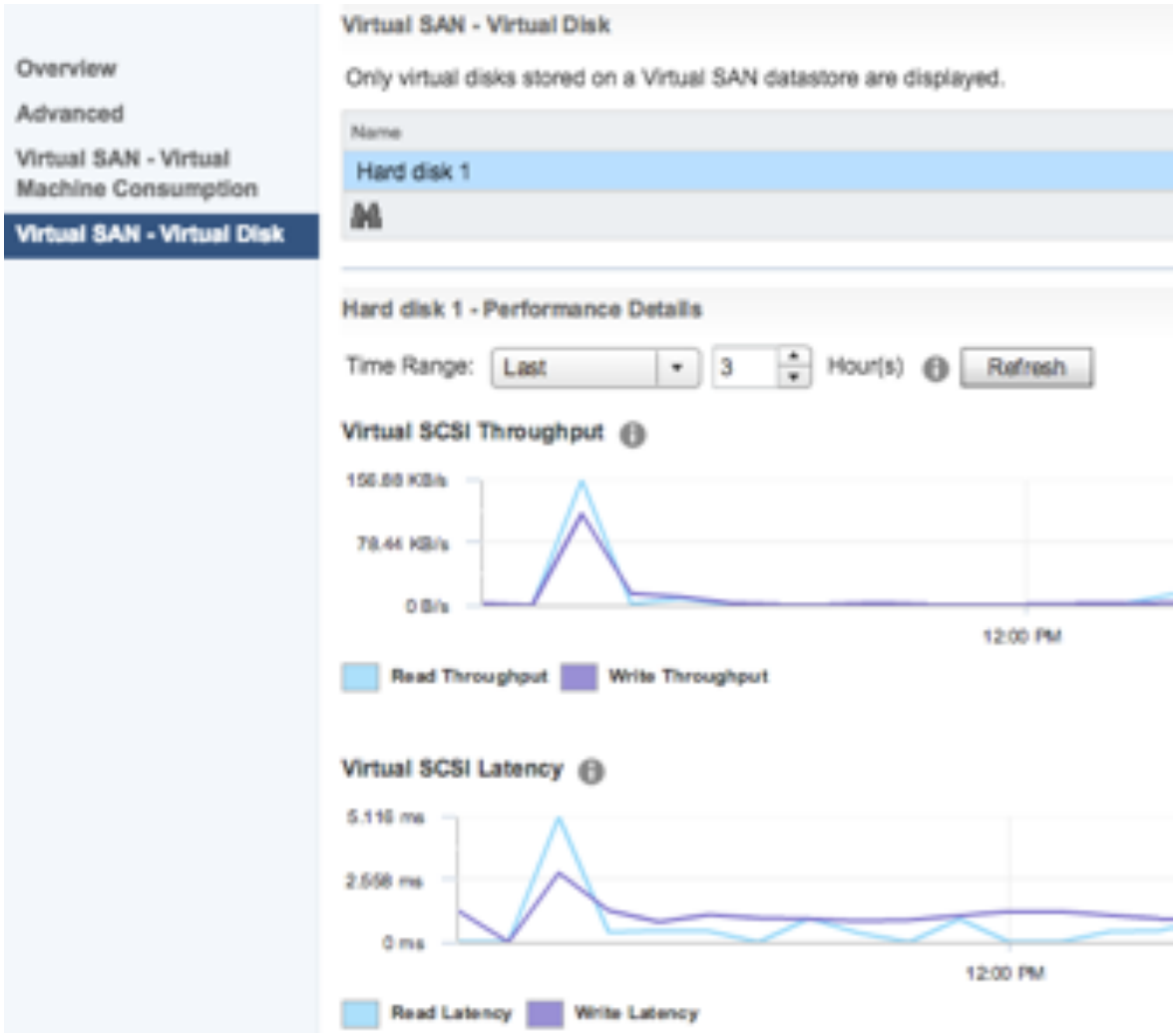
Note: The vSAN Performance Service is turned off by default in versions prior to vSAN 6.7.

6.3 Performance Metrics

Cluster Metrics
These metrics provide visibility to the entire vSAN cluster. Graphs show IOPs, throughput, latency, congestion, and outstanding I/O. "vSAN – Virtual Machine Consumption" graphs show metrics generated by virtual machines across the cluster. In addition to normal virtual machines reads and writes, "vSAN – Backend" consumption adds traffic such as metadata updates, component rebuilds, and data migrations.

Host Metrics
In addition to virtual machine and backend metrics, disk group, individual disk, physical network adapter, and VMkernel adapter performance information is provided at the host level. Seeing metrics for individual disks eases the process of troubleshooting issues such as failed storage devices. Throughput, packets per second, and packet loss rate statistics for network interfaces help identify potential networking issues.

Virtual Machine Metrics
Virtual machine metrics include IOPS, throughput, and latency. It is also possible to view information at the virtual disk level. The figure below shows virtual disk-level Virtual SCSI throughput and latencies for reads and writes.



6.4 Performance Diagnostics

vSAN Performance Diagnostics analyzes previously executed benchmarks. Administrators can select the desired benchmark goal such as maximum throughput or minimum latency and a time range during which the benchmark ran. Performance data is collected and analyzed. If an issue is discovered, it is reported in the vSphere Web Client along with recommendations for resolving the issue. The Ask

VMware button and links provide access to relevant VMware Knowledge Base articles.

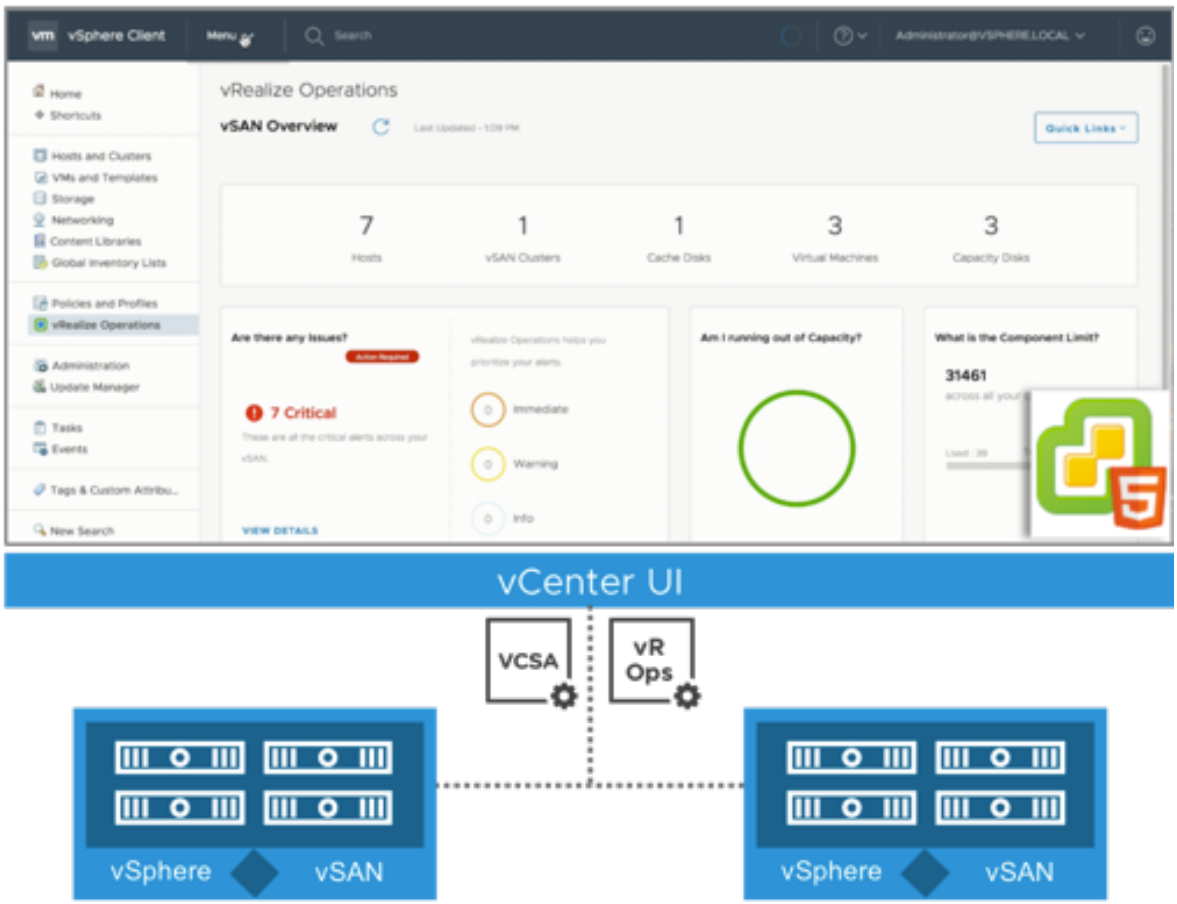
HCIBench has API-level integration with Performance Diagnostics. Administrators run a benchmark test and HCIBench will save the time range in vCenter Server. Detailed results of the test including supporting graphs are easily retrieved from the Performance Diagnostics section in the vSphere Web Client. This integration simplifies activities such and conducting a proof of concept and verifying a new vSAN cluster deployment.

Note: Joining CEIP and enabling the vSAN Performance Service is required to use the Performance Diagnostics feature.

6.5 Native vRealize Operations Integration

VMware vRealize Operations streamlines and automates IT operations management. Intelligent operations management from applications to infrastructure across physical, virtual, and cloud environments can be achieved with vRealize Operations.

vRealize Operations Management Packs could be added to extend the capabilities of vRealize Operations by including prebuilt dashboards that focus on design, management, and operations for a variety of solutions and use cases. Until the release of vSAN 6.7, using vRealize Operations with vSAN meant that an administrator needed to use two separate interfaces.

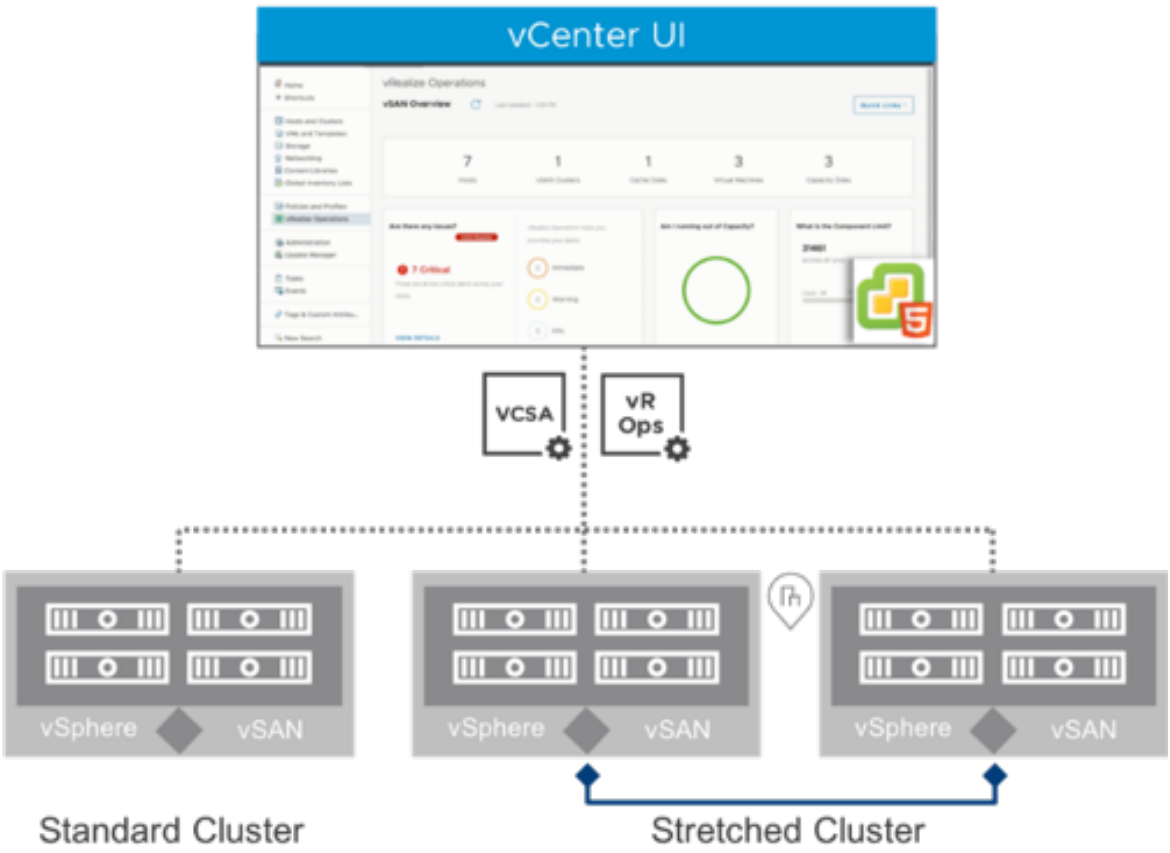


In vSAN 6.7, “vRealize Operations within vCenter” provided an easy way for customers to see basic vRealize intelligence with vCenter. vCenter now includes an embedded plugin that allows for users to easily configure and integrate a new vRealize Operations instance or to utilize an existing vRealize Operations instance.

The deployment process of vRealize Operations through the vSphere Client in vSphere 6.7 has been enhanced with better error handling and support for vSphere Distributed Switches to make it even easier to get started.

The vRealize Operations instance is now visible natively in the vSphere Client. The integrated nature of “vRealize Operations within vCenter” also allows the user to easily launch the full vRealize Operations user interface to see the full collection of vRealize Operations dashboards. These provide additional visibility into and analytics for vSAN environments.

New in vSAN 6.7 U1, vRealize Operations dashboards have the ability to differentiate between normal and stretched vSAN clusters, displaying appropriate intelligence for each.



An incredible number of metrics are exposed to assist with monitoring and issue remediation. vRealize Operations makes it easier to correlate data from multiple sources to speed troubleshooting and root cause analysis.

7. Data Services

7.1 Deduplication and Compression

Space efficiency features such as deduplication, compression, and erasure coding reduce the total cost of ownership (TCO) of storage. Even though flash capacity is currently more expensive than magnetic disk capacity, using space efficiency features makes the cost-per-usable-gigabyte of flash devices the same as or lower than magnetic drives. Add in the benefits of higher flash performance and it is easy to see why all-flash configurations are more popular.

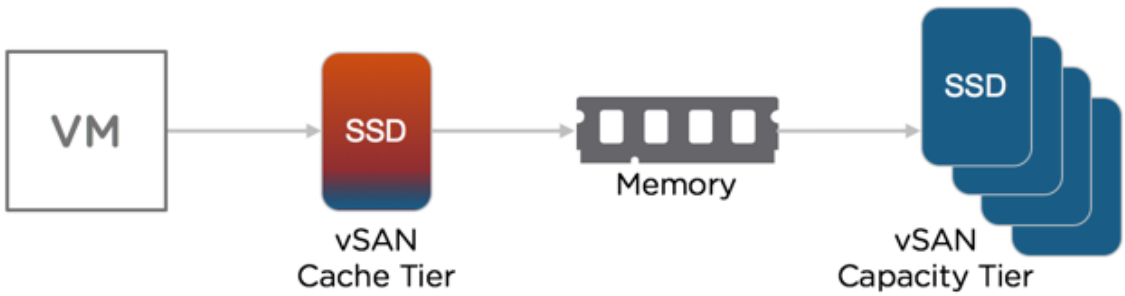
Enabling deduplication and compression can reduce the amount of physical storage consumed by as much as 7x. Environments with redundant data such as similar operating systems typically benefit the most. Likewise, compression offers more favorable results with data that compresses well like text, bitmap, and program files. Data that is already compressed such as certain graphics formats and video files, as well as files that are encrypted, will yield little or no reduction in storage consumption from compression. Deduplication and compression results will vary based on the types of data stored in an all flash vSAN environment.

Deduplication and compression is a single cluster-wide setting that is disabled by default and can be enabled using a simple drop-down menu.

Note: A rolling format of all disks in the vSAN cluster is required when deduplication and compression are enabled on an existing cluster. This can take a considerable amount of time. However, this process does not incur virtual machine downtime.

Deduplication and compression are implemented after writes are acknowledged in the vSAN cache tier to minimize impact to performance. The deduplication algorithm utilizes a 4K fixed block size and is performed within each disk group. In other words, redundant copies of a block within the same disk group are reduced to one copy, but redundant blocks across multiple disk groups are not deduplicated.

“Cold” data in the cache tier that is ready to be de-staged is moved to memory where it is deduplicated and compressed. It is then written to the capacity tier.



The compression algorithm is applied after deduplication has occurred just before the data is written to the capacity tier. Considering the additional compute resource and allocation map overhead of compression, vSAN will only store compressed data if a 4K block can be reduced to 2K. Otherwise, the block is written uncompressed to avoid the use of additional resources.

The processes of deduplication and compression on any storage platform incur overhead and potentially impact performance in terms of latency and maximum IOPS. vSAN is no exception.

However, considering deduplication and compression are only supported in all flash vSAN configurations, these effects are predictable in most use cases. The extreme performance and low latency of flash devices easily outweighs the additional resource requirements of deduplication and compression. The space efficiency generated by deduplication and compression lowers the cost-per usable-GB of all flash.

7.2 Erasure Coding (RAID-5/6)

RAID-5/6 erasure coding is a space efficiency feature optimized for all flash configurations. Erasure coding provides the same levels of redundancy as mirroring, but with a reduced capacity requirement. In general, erasure coding is a method of taking data, breaking it into multiple pieces and spreading it across multiple devices, while adding parity data so it may be recreated in the event one of the pieces is corrupted or lost.

Unlike deduplication and compression, which offer variable levels of space efficiency, erasure coding guarantees capacity reduction over a mirroring data protection method at the same failure tolerance level. As an example, let's consider a 100GB virtual disk. Surviving one disk or host failure requires 2 copies of data at 2x the capacity, i.e., 200GB. If RAID-5 erasure coding is used to protect the object, the 100GB virtual disk will consume 133GB of raw capacity—a 33% reduction in consumed capacity versus RAID-1 mirroring.

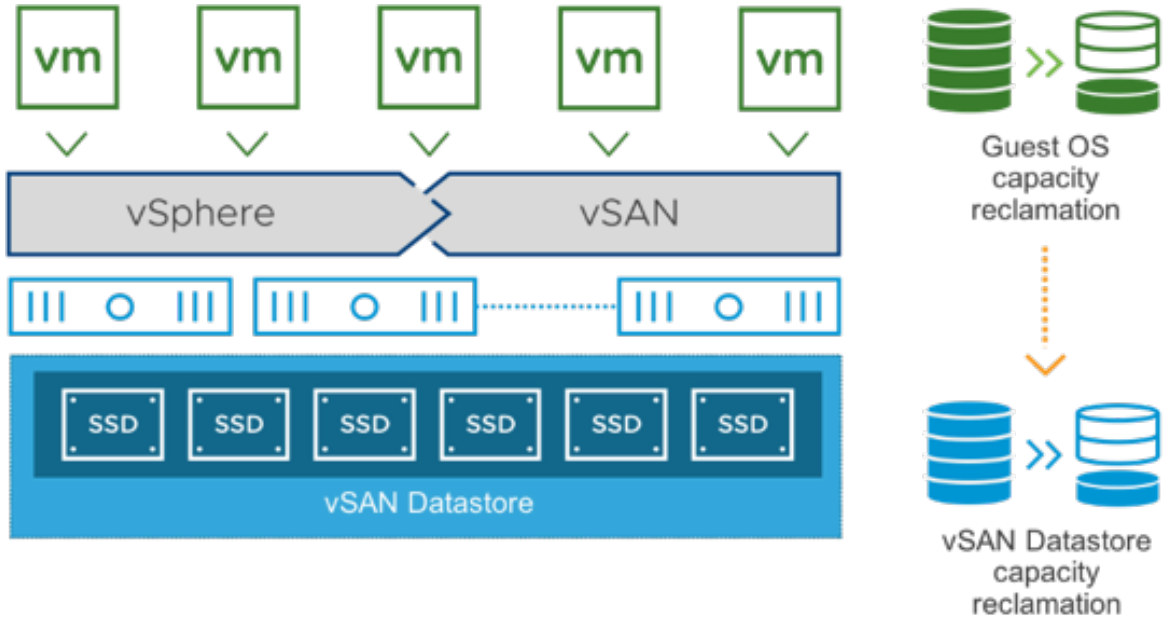
RAID-5 erasure coding requires a minimum of four hosts. Let's look at a simple example of a 100GB virtual disk. When a policy containing a RAID-5 erasure coding rule is assigned to this object, three data components and one parity component are created. To survive the loss of a disk or host (FTT=1), these components are distributed across four hosts in the cluster.

RAID-6 erasure coding requires a minimum of six hosts. Using our previous example of a 100GB virtual disk, the RAID-6 erasure coding rule creates four data components and two parity components. This configuration can survive the loss of two disks or hosts simultaneously (FTT=2).

While erasure coding provides significant capacity savings over mirroring, understand that erasure coding requires additional processing overhead. This is common with any storage platform. Erasure coding is only supported in all flash vSAN configurations. Therefore, the performance impact is negligible in most cases due to the inherent performance of flash devices.

7.3 TRIM/UNMAP

Modern guest operating systems have the ability to reclaim no longer used space once data is deleted inside of a guest operating system. Using commands known as TRIM/UNMAP for the respective ATA and SCSI protocols, this helps the guest operating systems be more efficient with storage space usage.



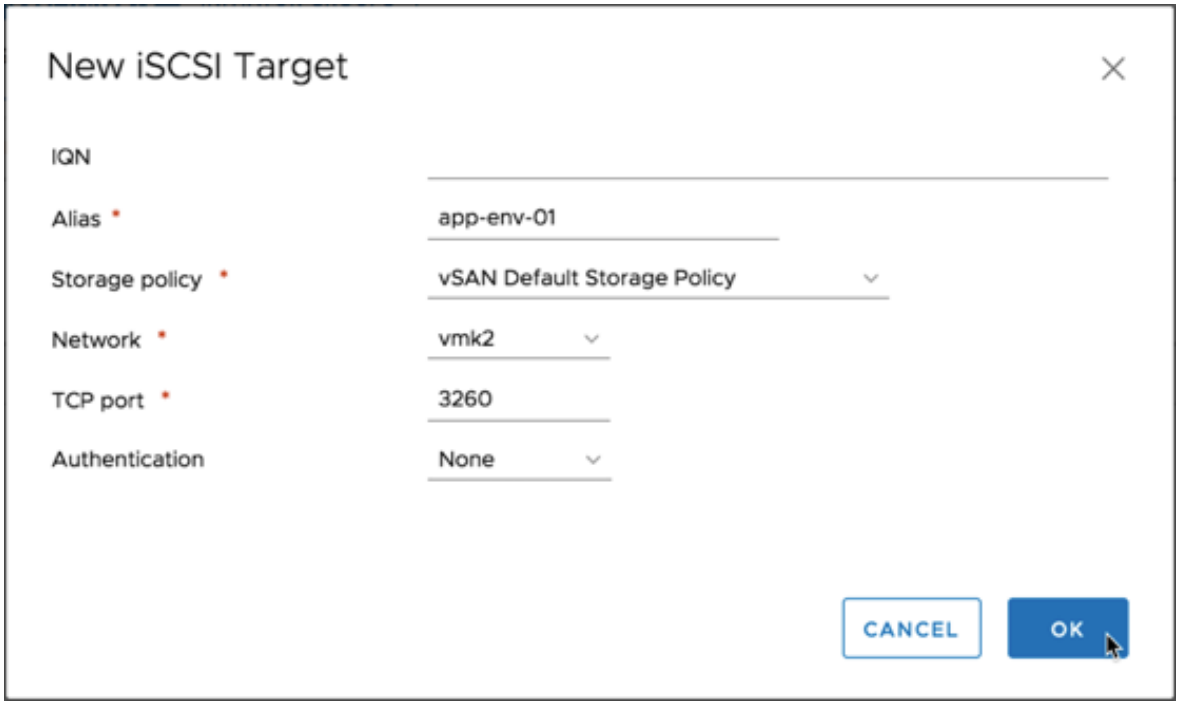
vSAN 6.7 U1 now has full awareness of TRIM/UNMAP commands sent from the guest OS and can reclaim the previously allocated storage as free space.

This is an opportunistic space efficiency feature that can deliver much better storage capacity utilization in vSAN environments. Administrators may find in some cases, dramatic space savings in their production vSAN environments.

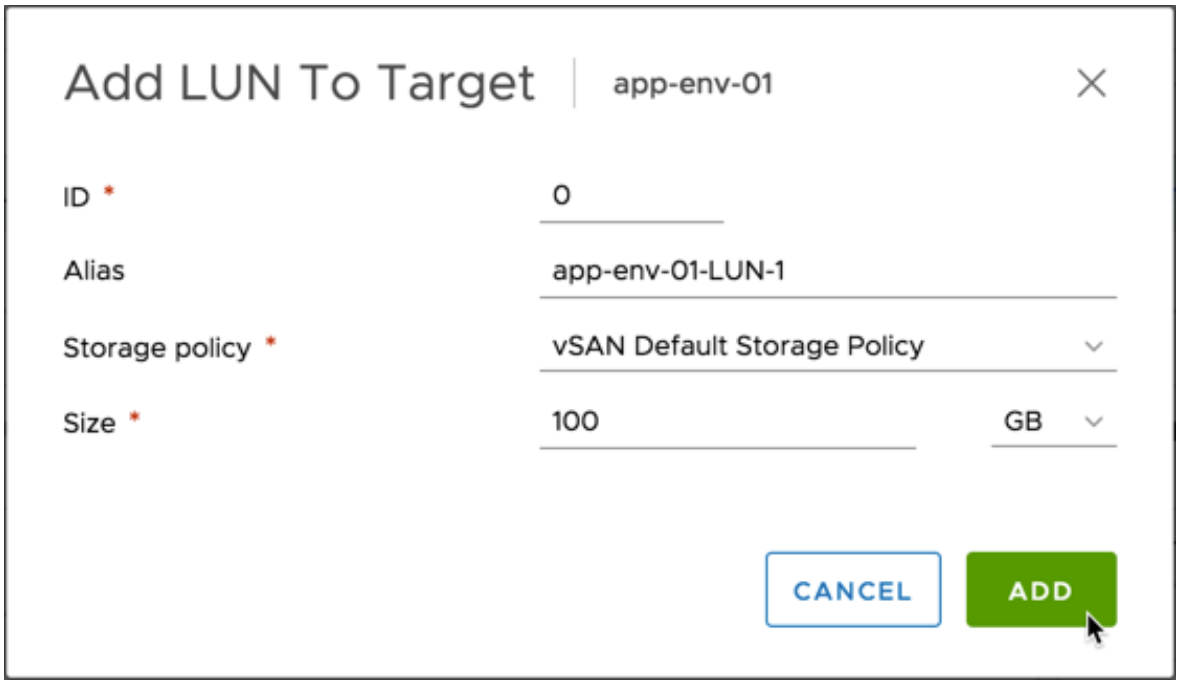
7.4 iSCSI Target Service

Block storage can be provided to physical workloads using the iSCSI protocol. The vSAN iSCSI target service provides flexibility and potentially avoids expenditure on purpose-built, external storage arrays. In addition to capital cost savings, the simplicity of vSAN lowers operational costs.

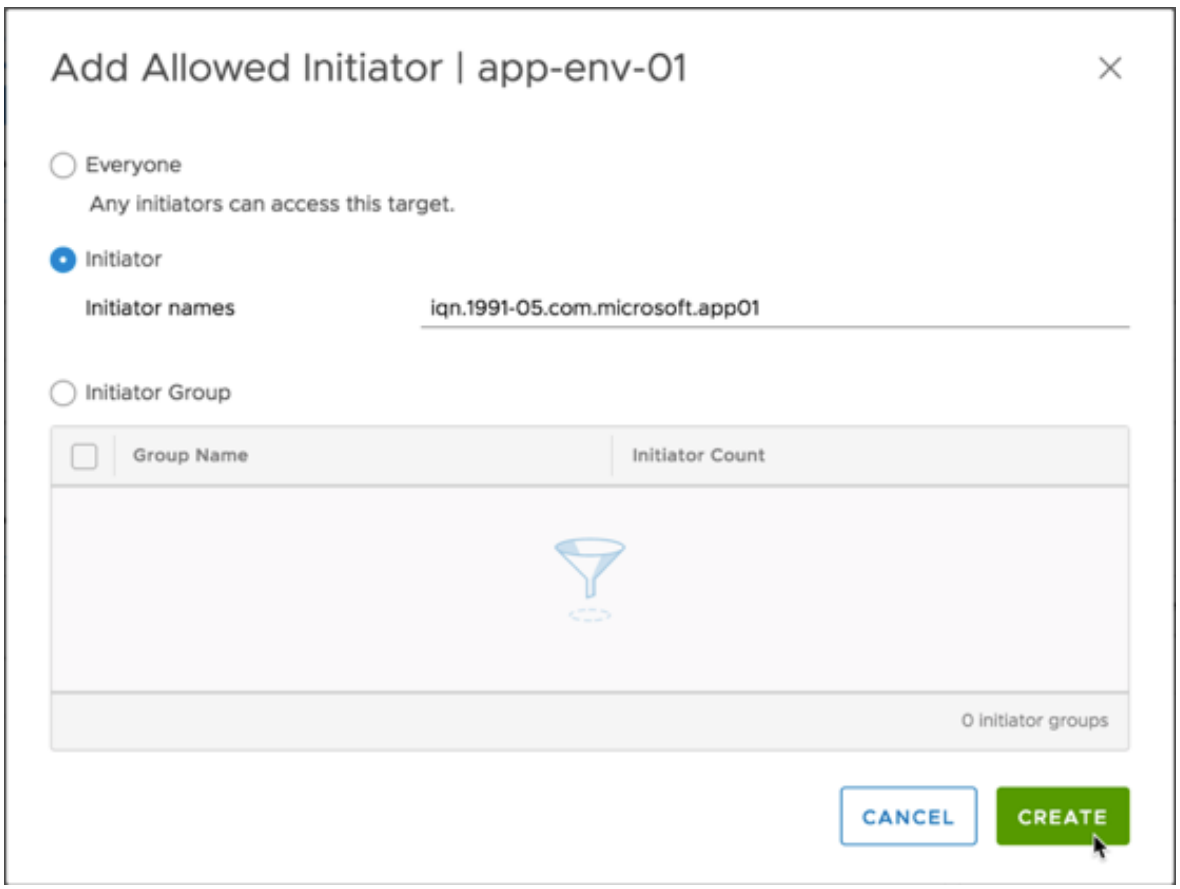
The vSAN iSCSI target service is enabled with just a few mouse clicks. CHAP and Mutual CHAP authentication are supported. vSAN objects that serve as iSCSI targets are managed with storage policies just like virtual machine objects. After the iSCSI target service is enabled, iSCSI targets and LUNs can be created. The figure below shows vSAN iSCSI target configuration.



When a target has been added, it is easy to add a LUN.



The last step is adding initiator names or an initiator group, which controls access to the target. It is possible to add individual names or create groups for ease of management.



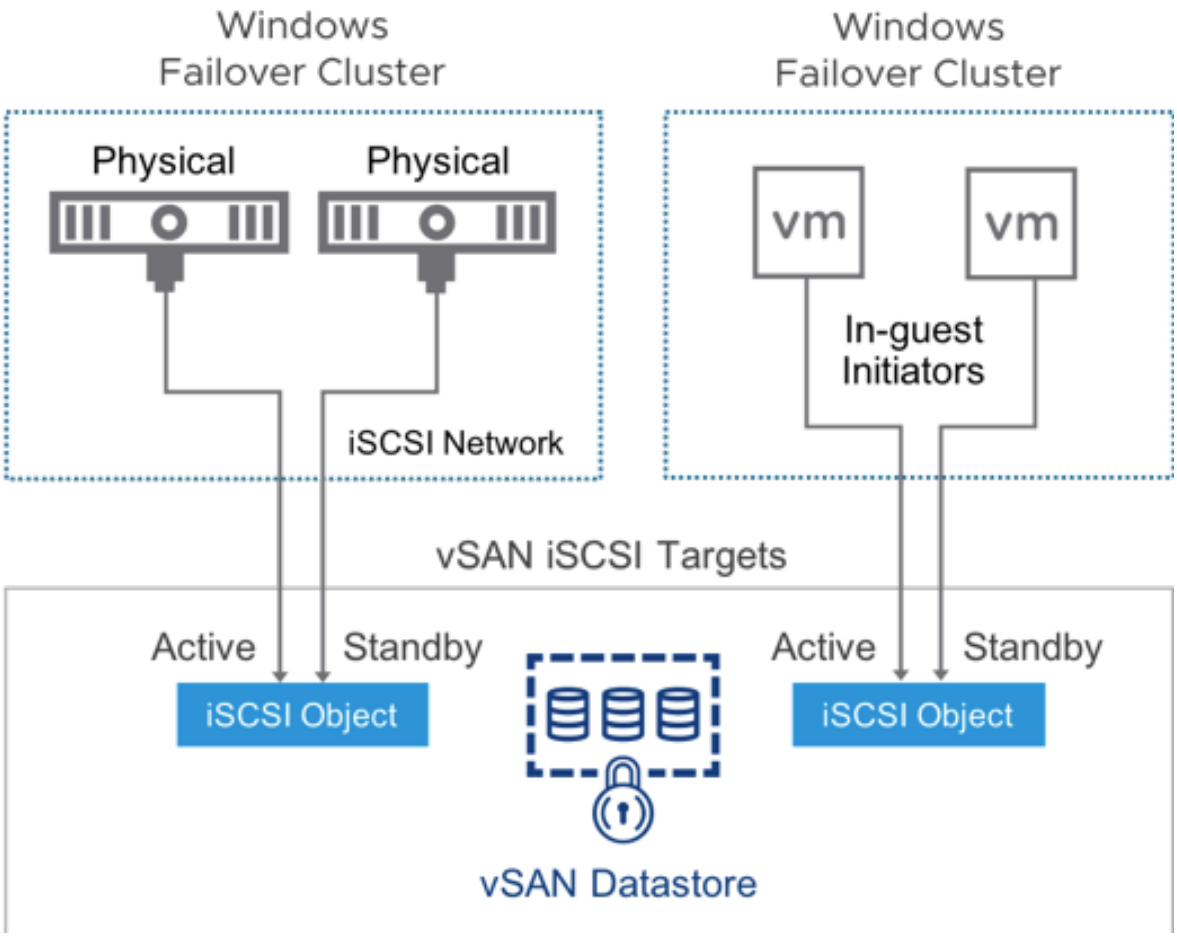
In nearly all cases, it is best to run workloads in virtual machines to take full advantage of vSAN's simplicity, performance, and reliability. However, for those use cases that truly need block storage, it is now possible to utilize vSAN iSCSI Target Service.

Windows Server Failover Clustering with vSAN iSCSI Service

vSAN 6.7 now expands the flexibility of the vSAN iSCSI service to support Windows Server Failover Clusters, or WSFC. vSAN already natively supports modern Windows application layer clustering

technologies such as Microsoft SQL Always-on availability Groups (AAG), and Microsoft Exchange Database Availability Groups (DAG).

For scenarios in which an organization has WSFC servers in a physical or virtual configuration, vSAN can support legacy shared target storage when the storage target is exposed using the vSAN iSCSI service. This allows data center administrators to run workloads using legacy clustering technologies on vSAN, as they eventually transition to more modern implementations of application layer clustering like AAGs and DAGs.



In-guest support to WFSC VMs is on vSphere only; moreover, WFSC VMs using an in-guest iSCSI initiator may reside on a vSAN datastore or other vSphere supported storage.

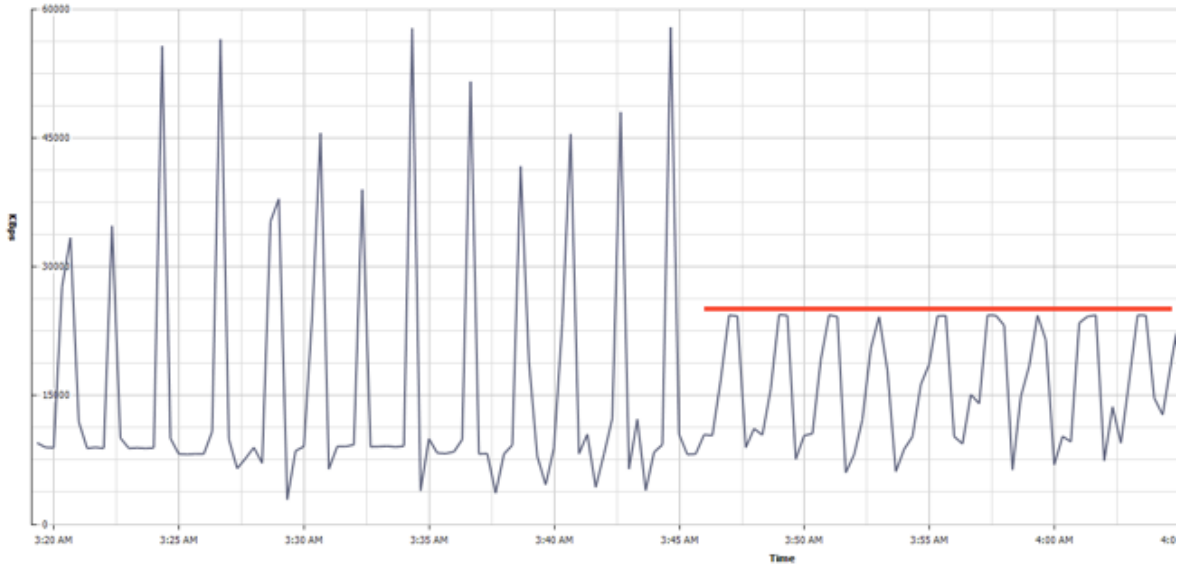
7.5 Certified File Services Solutions

Certified solutions for file services are available through the VMware Ready for vSAN™ program. Customers can deploy these solutions with confidence to extend HCI environments with proven, industry-leading solutions. Using these solutions with vSAN provides benefits such as simplified setup and management, documented recommendations, and robust support. The diagram below shows a virtual storage appliance providing standard file services to physical workloads.

7.6 IOPS Limits

vSAN can limit the number of IOPS a virtual machine or virtual disk generates. There are situations where it is advantageous to limit the IOPS of one or more virtual machines. The term noisy neighbor is often used to describe when a workload monopolizes available IO or other resources, which negatively impact other workloads or tenants in the same environment.

An example of a possible noisy neighbor scenario is month-end reporting. Management requests delivery of these reports on the second day of each month so the reports are generated on the first day of each month. The virtual machines that run the reporting application and database are dormant most of the time. Running the reports take just a few hours, but this generates very high levels of storage I/O. The performance of other workloads in the environment is sometimes impacted while the reports are running. To remedy this issue, an administrator creates a storage policy with an IOPS limit rule and assigns the policy to the virtual machines running the reporting application and database.



The IOPS limit eliminates possible performance impact to the other virtual machines. The reports take longer, but they are still finished in plenty of time for delivery the next day.

Keep in mind storage policies can be dynamically created, modified, and assigned to virtual machines. If an IOPS limit is proving to be too restrictive, simply modify the existing policy or create a new policy with a different IOPS limit and assign it to the virtual machines. The new or updated policy will take effect just moments after the change is made.

7.7 New Application Technologies

New application architecture and development methods have emerged that are designed to run in today's mobile-cloud era. Container technologies such as Docker and Kubernetes are a couple of the many solutions that have emerged as options for deploying and orchestrating these applications. VMware is embracing these new application types with a number of products and solutions. Here are a few examples:

- Photon OS – a minimal Linux container host optimized to run on VMware platforms
- Photon Controller – a distributed, multi-tenant ESXi host controller for containers
- Lightwave – an enterprise-grade identity and access management services
- Admiral – a highly scalable and very lightweight container management platform

Cloud native applications naturally require persistent storage just the same as traditional applications. Deploying vSAN for Photon Platform enables the use of a vSAN cluster in cloud native application environments managed by Photon Controller.

VMware vSphere Integrated Containers™ provides enterprise container infrastructure to help IT Ops run both traditional and containerized applications side-by-side on a common platform. Supporting containers in a virtualized environment provides a number of benefits: IT teams get the security, isolation, and management of VMs, while developers enjoy the speed and agility of containers—all within the familiar vSphere platform. Availability and performance features in vSphere and vSAN can be utilized by vSphere Integrated Containers just the same as traditional storage environments.

VMware vSphere Docker Volume Service enables vSphere users to create and manage Docker container data volumes on vSphere storage technologies such as VMFS, NFS, and vSAN. This driver makes it very simple to use containers with vSphere storage and provides the following key benefits:

- DevOps-friendly API for provisioning and policy configuration.
- Seamless movement of containers between vSphere hosts without moving data.
- Single platform to manage—run virtual machines and containers side-by-side on the same vSphere infrastructure.

vSAN along with the solutions above provides an ideal storage platform for developing, deploying, and managing cloud native applications.

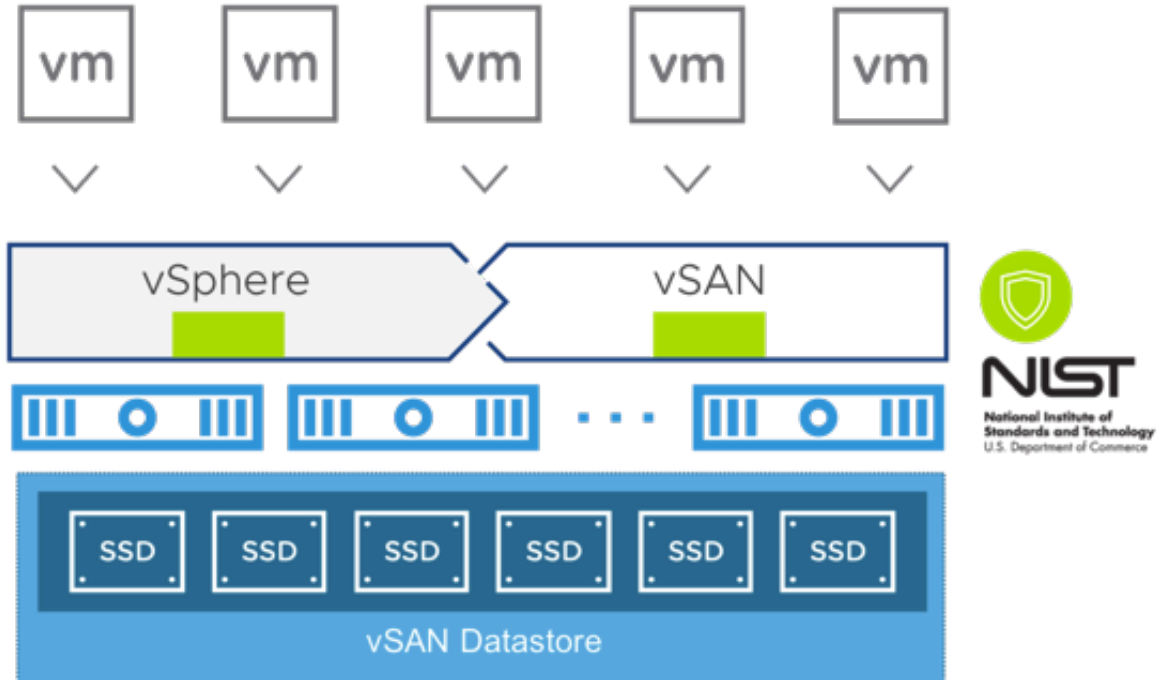
7.8 Data at Rest Encryption

vSAN Data at Rest Encryption will be covered in the VMware HCI Security Section.

8. VMware HCI Security

8.1 Native VMkernel Cryptographic Module

In December of 2017, VMware achieved FIPS 140-2 validation under the Cryptographic Module Validation Program (CMVP). The CMVP is a joint program between NIST and the Communications Security Establishment (CSE). FIPS 140-2 is a Cryptographic Modules Standards that governs security requirements in 11 areas relating to the design and implementation of a cryptographic module.



The VMware VMkernel Cryptographic Module has successfully satisfied all requirements of these 11 areas and has gone through required algorithms and operational testing, rigorous review by CMVP and third-party laboratory before being awarded certificate number 3073 by the CMVP. The details of this validation, along with the tested configurations are available at: https://csrc.nist.gov/Projects/Cryptographic-Module-Validation-Program/Certificate/3073

The implementation and validation of the VMkernel Cryptographic Module shows VMware's commitment to providing industry leading virtualization, cloud, and mobile software that embrace commercial requirements, industry standards, and government certification programs.

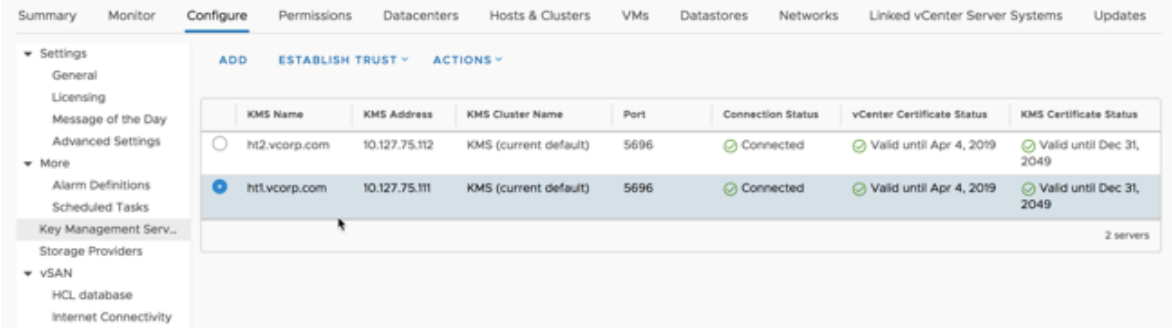
Because the VMware VMkernel Cryptographic Module is part of the ESXi kernel, it can easily provide FIPS 140-2 approved cryptographic services to various VMware products and services.

Virtual machines encrypted with VM Encryption or vSAN Encryption work with all vSphere supported Guest Operating Systems and Virtual Hardware versions, and do not allow access to encryption keys by the Guest OS.

8.2 Key Management

Key management is a core requirement for being able to use vSAN Encryption and VM Encryption. A Key Management Solution using Key Management Interoperability Protocol (KMIP) version 1.1 is required. Any Key Management Server using KMIP 1.1 is supported, but VMware has worked with several industry vendors to validate their solutions with these features.

Initial KMS configuration is done in the vCenter Server UI. A KMS solution profile is added to vCenter, and a trust relationship is established. Different KMS vendors have different processes to establish this trust but is relatively simple.

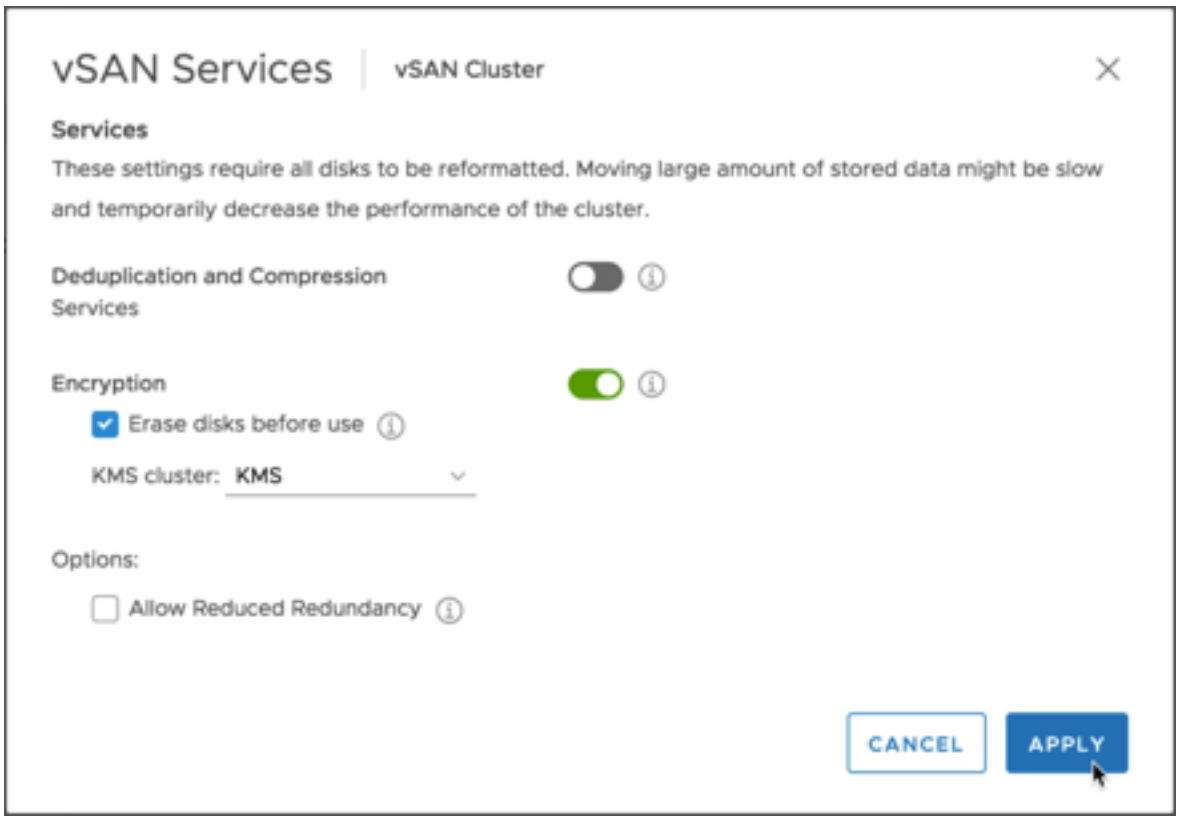


For a complete list of validated KMS vendors, there solutions, and links to their documentation, reference the VMware Hardware Compatibility List.

8.3 vSAN Encryption

The VMware VMkernel Cryptographic Module is used by vSAN to provide Data at Rest Encryption to protect data on the disk groups and their devices in a vSAN cluster.

Because vSAN Encryption is a cluster level data service, when encryption is enabled, all vSAN objects are encrypted in the Cache and Capacity tiers of the vSAN datastore. With encryption configured at the cluster level, only a single Key Management Server may be used to provide Key Management Services.

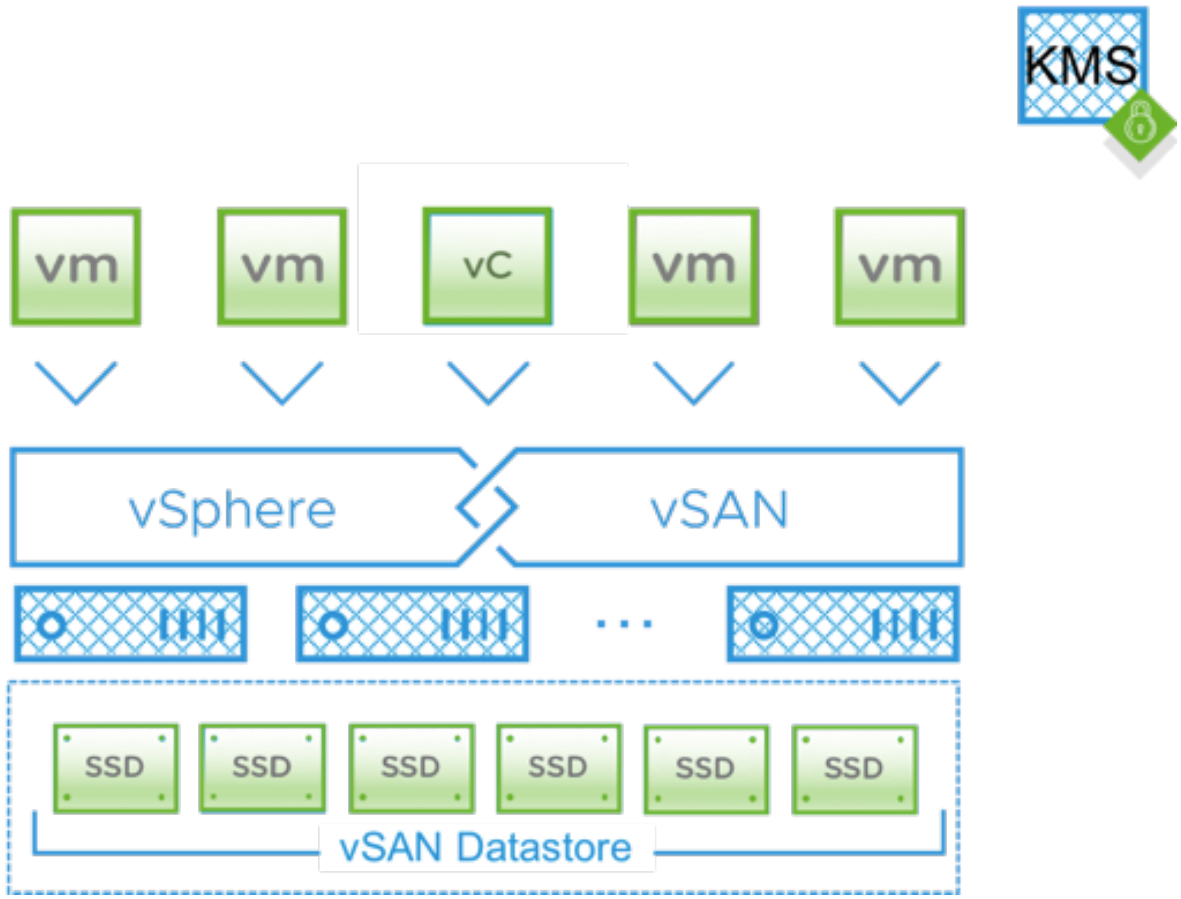


Encryption occurs just above the device driver layer of the vSphere storage stack, which means it is compatible with all vSAN features such as deduplication and compression, RAID-5/6 erasure coding, stretched cluster configurations. All vSphere features including vSphere vMotion, vSphere Distributed Resource Scheduler (DRS), vSphere Availability (HA), and vSphere Replication are supported. Data is

only encrypted in the Cache and Capacity tiers and not encrypted “on the wire” when being written to vSAN.

Because data is encrypted on physical vSAN devices, vSAN Encryption does not follow a virtual machine if it is migrated to another datastore. The target location must provide its own encryption mechanism to allow the data to remain encrypted at rest.

When enabling or disabling vSAN Encryption, a rolling reformat of Disk Groups will be required as the vSAN cluster encrypts or decrypts individual storage devices.

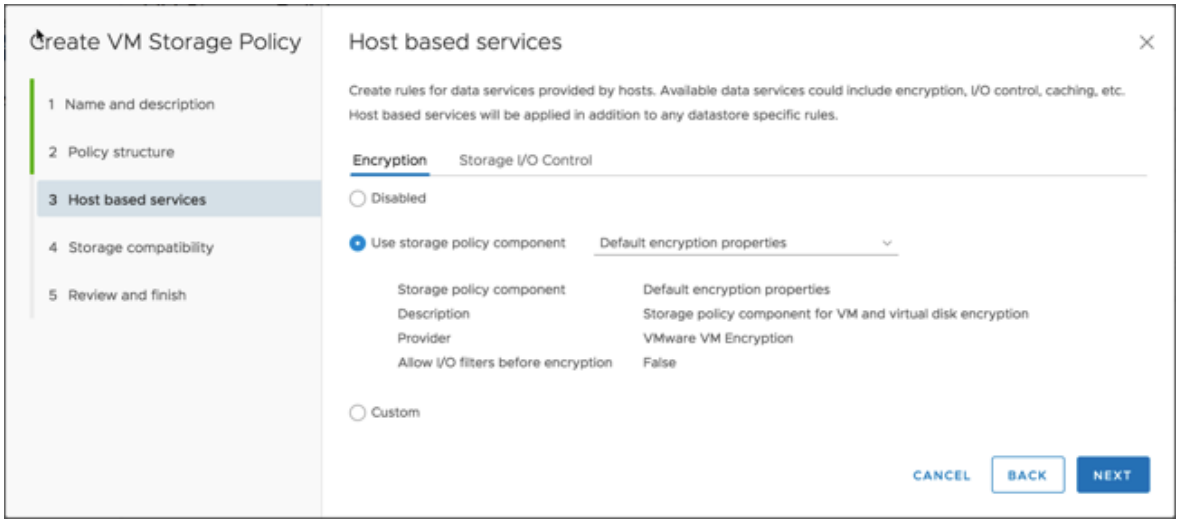


vCenter and PSC services can be run on an encrypted vSAN cluster, because each host will directly contact the Key Management Server upon reboot. There is no dependency on vCenter being available during a host boot up, key retrieval, and mounting of encrypted vSAN disks.

8.4 VM Encryption

The VMware VMkernel Cryptographic Module is also used by vSphere Virtual Machine (VM) Encryption. VM Encryption can be used create new encrypted virtual machines or to encrypt existing virtual machines. VM Encryption is applied to virtual machines and their disks individually, through the use of Storage Policies.

Because of this, each virtual machine must be independently assigned a Storage Policy with a common rule that includes encryption. Unlike vSAN Encryption, where the whole cluster uses a single Key Encryption Key (KEK), each VM has its own KEK. This makes VM Encryption very advantageous where there is a requirement for multiple Key Management Servers, such as scenarios where different departments or organizations could be managing their own KMS.



Once encrypted with VM Encryption, workloads that were once similar, and could be easily deduplicated, are no longer similar. As a result, VM Encryption may not be best suited for use cases where high storage consolidation savings are required.

VM Encryption performs encryption of virtual machine data as it is being written to the virtual machine's disks. Because of this, the virtual machine itself is encrypted. If a virtual machine is encrypted with VM Encryption and it is migrated to an alternate vSphere datastore, or offline device, the encryption remains with the virtual machine. VM Encryption works with any supported vSphere datastore.

Encrypted vMotion

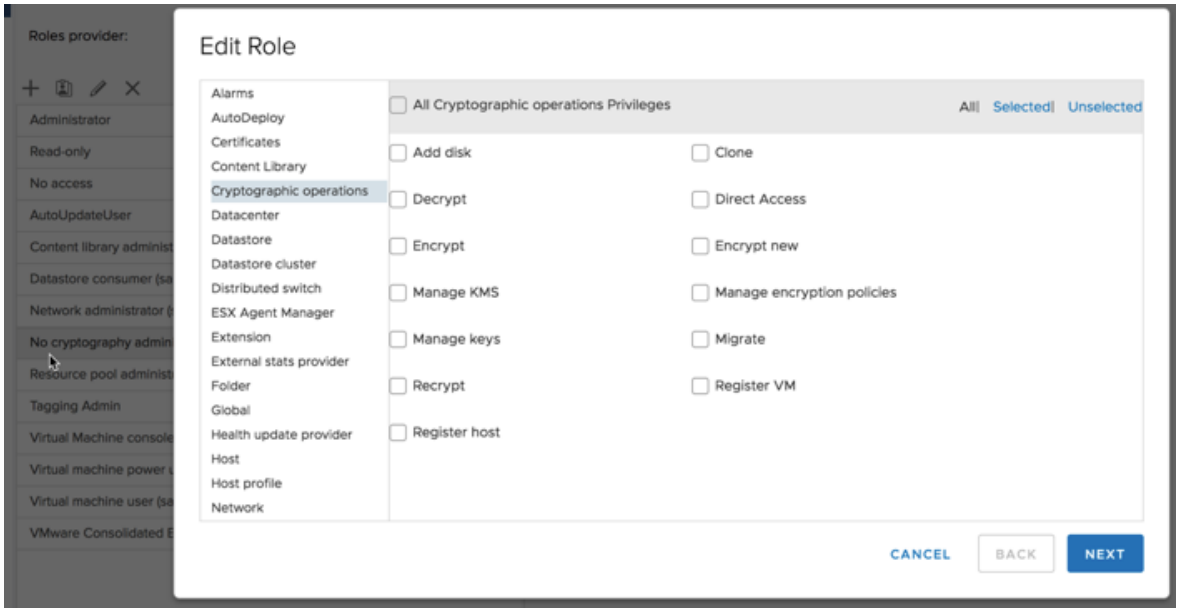
The Encrypted vMotion feature available in vSphere enables a secure way of protecting critical VM data that traverses across clouds and over long distances. The VMware VMkernel Cryptographic Module implemented in vSphere is utilized to encrypt all the vMotion data inside the VMkernel by using the most widely used and secured AES-GCM algorithm, thereby providing data confidentiality, integrity, and authenticity even when vMotion traffic traverses over untrusted network links.

8.5 Role Based Access Control

Securing workloads does not end with the use of encryption technologies. Access granted to data and its management must also be properly secured. Effective access to these workloads must align with the responsibilities associated with their management, configuration, reporting, and use requirements.

No Cryptography Administrator Role

vSphere 6.5 introduced the No Cryptography Administrator role along with the introduction of VM Encryption. This role is very similar to the normal administrator with many of the same privileges. Operations such as power on or off a virtual machine, boot, shutdown, vMotion, as well as normal vSAN management may be performed. However, this role is not allowed to perform any cryptographic operations.



The permissions in the illustration, show that users assigned the No Cryptography Administrator role do not have any permissions to perform any operations that require any cryptographic operations.

No Cryptography Administrator and VM Encryption

Users assigned to the No Cryptography Administrator role are not granted the following privileges:

- Ability to encrypt or decrypt virtual machines with VM Encryption
- Direct console access to virtual machines that are encrypted with VM Encryption
- The ability to download virtual machines that are encrypted with VM Encryption. This will prevent the user from downloading a virtual machine to a USB or other offline media.
- The ability to add hosts to vCenter. This limitation exists, because the process of adding a host to vCenter grants the host access to the cryptographic keystore.

No Cryptography Administrator and vSAN Encryption

Users assigned to the No Cryptography Administrator role are **not granted** the following privileges:

- The ability to enable or disable vSAN Encryption
- The ability to generate new encryption keys (Shallow or Deep Rekey)
- The ability to add hosts to vCenter.

Users assigned to the No Cryptography Administrator role **are granted** the following privileges:

- Direct console access to virtual machines that reside on a vSAN Cluster with vSAN Encryption enabled
- The ability to download virtual machines that reside on a vSAN Cluster with vSAN Encryption enabled.
- The ability to add hosts to a vSAN Cluster*.

* In a situation where a host needs to be added to a vSAN Cluster, a user with Cryptographic rights would have to add the host to vCenter. Once added to vCenter a Non-Cryptographic Administrator could then add the host to an encrypted vSAN Cluster.

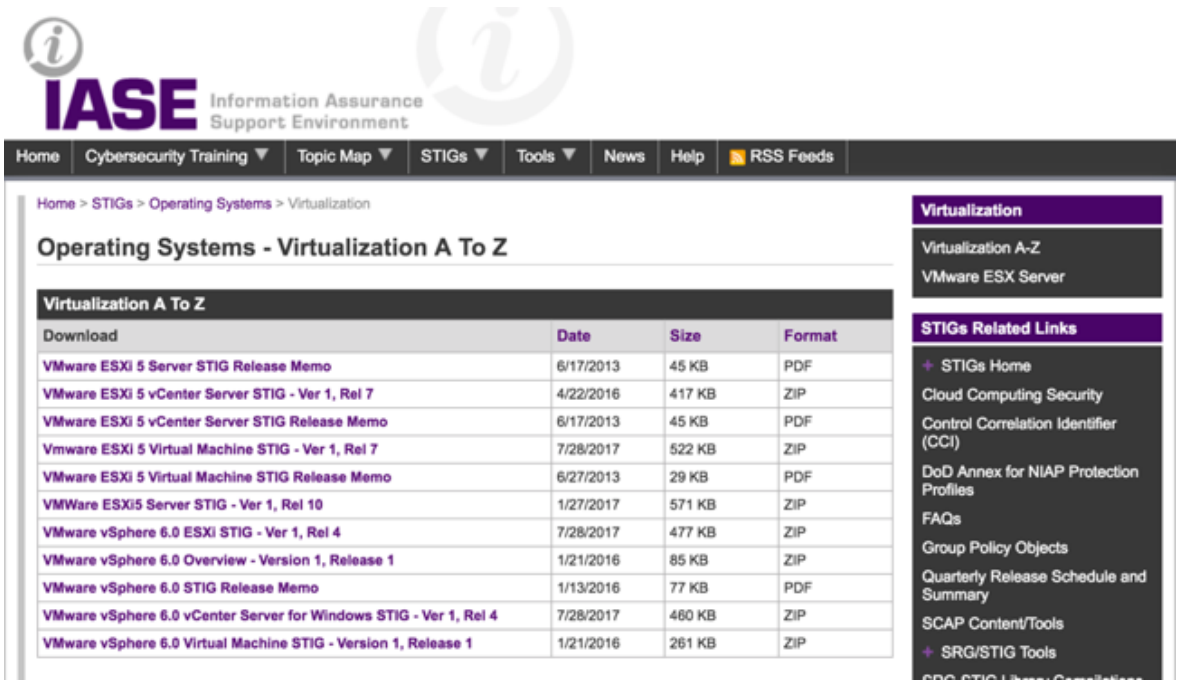
8.6 Compliance

vSAN is native to the vSphere hypervisor and, because of that tight integration, shares the robust security and compliance benefits realized by the vSphere platform. 2-factor authentication methods, such as RSA SecurID and Common Access Card (CAC), are supported by vCenter Server, vSphere, and vSAN.

In April of 2017, VMware announced that VMware vSAN has been added to the VMware vSphere STIG Framework. The updated Defense Information Systems Agency (DISA) Security Technical Implementation Guide (STIG) defines secure installation requirements for deploying vSphere and vSAN on U.S. Department of Defense (DoD) networks. VMware worked closely with DISA to include vSAN in this update to the existing VMware vSphere STIG Framework. With this update, VMware HCI, comprised of vSphere and vSAN, is the first and only HCI solution that has DISA published and approved STIG documentation.

The purpose of a DISA STIG, is to reduce information infrastructure vulnerabilities. These guides are primarily created as guidance for deployment and operations of U.S. Government infrastructure, though it is not uncommon for other organizations use them as well.

This is because information security is not exclusive to the U.S. Government. Security is important to organizations across all verticals, such financial, health care, and retail to name a few. Any organization that is interested in operating with a more security aware posture, can use these publicly available STIGs to better secure their environment. DISA STIGs can be found on the Information Assurance Support Environment (IASE) website.



The acronym STIG is not a copyrighted term but is ***uniquely associated*** with DISA.

DISA is mandated to develop STIGs against a very specific set of standards in collaboration with the NSA and other organizations. This is a formal process that is very time consuming, requiring close collaboration among all involved. When the Risk Management Executive signs and approves the STIG, it validates that the product in the STIG meets the risk acceptance level for use in the DoD. If important requirements are not met, DISA can and will refuse to sign/approve a proposed STIG.

It is not uncommon to hear the term "STIG Compliant," but this does not indicate being included in a certified, approved, and published DISA STIG. Achieving the inclusion in a DISA STIG is no small feat. Only through the coordination with and approval by DISA can security guidelines be part of a DISA STIG.

At VMware, we are excited to have VMware HCI included in the VMware vSphere STIG Framework to be able to provide this level of security to customers who need complete certainty about their security profile.

To view the official STIG approval, visit the IASE website.

9. Summary

9.1 HCI Powered by vSAN

HyperConverged Infrastructure, or HCI, consolidates traditional IT infrastructure silos onto industry standard servers. The physical infrastructure is virtualized to help evolve data centers without risk, reduce total cost of ownership (TCO), and scale to tomorrow with timely support for new hardware, applications, and cloud strategies.

HCI solutions powered by VMware consist of a single, integrated platform for storage, compute and networking that build on the foundation of VMware vSphere, the market-leading hypervisor, and VMware vSAN, the software-defined enterprise storage solution natively integrated with vSphere. vCenter Server provides a familiar unified, extensible management solution.

Seamless integration with vSphere and the VMware ecosystem makes it the ideal storage platform for business-critical applications, disaster recovery sites, remote office and branch office (ROBO) implementation, test and development environments, management clusters, security zones, and virtual desktop infrastructure (VDI). Today, customers of all industries and sizes trust vSAN to run their most important applications.

VMware provides the broadest choice of consumption options for HCI:

VMware Cloud Foundation™, a unified platform that brings together VMware's vSphere, vSAN and VMware NSX™ into an integrated stack for private and public clouds.

Dell EMC VxRail™, a turn-key HCI appliance tailored for ease of use and deployment

vSAN ReadyNodes™ with hundreds of systems from all major server vendors catering to flexibility of deployment needs and vendor preferences

vSAN focuses on enabling customers to modernize their infrastructure by enhancing three key areas of today's IT need: higher security, lower cost, and faster performance.

The industry's first native encryption solution for HCI is delivered through vSphere and vSAN. A highly available control plane is built in to help organizations minimize risk without sacrificing flash storage efficiencies.

vSAN lowers TCO by providing highly available, powerful, and economical stretched clusters that are 60% less than leading legacy storage solutions. Operational costs are also reduced with new, intelligent operations that introduce 1-click hardware updates for predictable hardware experience and pro-active cloud health checks for custom, real-time support.

vSAN is designed to scale to tomorrow's IT needs by optimizing flash performance for traditional and next-generation workloads. This enables organizations to realize the benefits of vSAN for all workloads.