DOCUMENT – SEPTEMBER 2020

PRINTED 17 MAY 2021

VSAN 7 TECHNOLOGY OVERVIEW

VMwareStorage

vmware®

Table of Contents

vSAN 7 Technology Overview

Introduction

VMware vSAN continues to be the Hyperconverged Infrastructure (HCI) market leader. vSAN has proven to be an excellent fit for all types of workloads. Traditional applications like Microsoft SQL Server and SAP HANA; next-generation applications like Cassandra, Splunk, and MongoDB; and even container-based services orchestrated through Kubernetes are run on vSAN by customers today. The success of vSAN can be attributed to many factors such as performance, flexibility, ease of use, robustness, and pace of innovation.

Paradigms associated with traditional infrastructure deployment, operations, and maintenance include various disaggregated tools and often specialized skill sets. The hyperconverged approach of vSphere and vSAN simplifies these tasks using familiar tools to deploy, operate, and manage private-cloud infrastructure. VMware vSAN provides the best-in-class enterprise storage and is the cornerstone of VMWare Cloud Foundation, accelerating customer's multi-cloud journey.

VMware HCI, powered by vSAN, is the cornerstone for modern datacenters whether they are on-premises or in the cloud. vSAN Runs on standard x86 servers from more than 18 OEMs. Deployment options include over 500 vSAN ReadyNode choices, integrated systems such as Dell EMC VxRail 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.

Whether you are a customer deploying traditional, or container-based applications, vSAN delivers developer-ready infrastructure, scales without compromise, simplifies operations, and management tasks as the best HCI solution today – and tomorrow.

Architecture

vSAN is VMware’s software-defined storage solution, built from the ground up for vSphere virtual machines.

It abstracts and aggregates locally attached disks in a vSphere cluster to create a storage solution that can be provisioned and managed from vCenter and the vSphere Client. vSAN is embedded within the hypervisor, hence storage and compute for VMs are delivered from the same x86 server platform running the hypervisor.

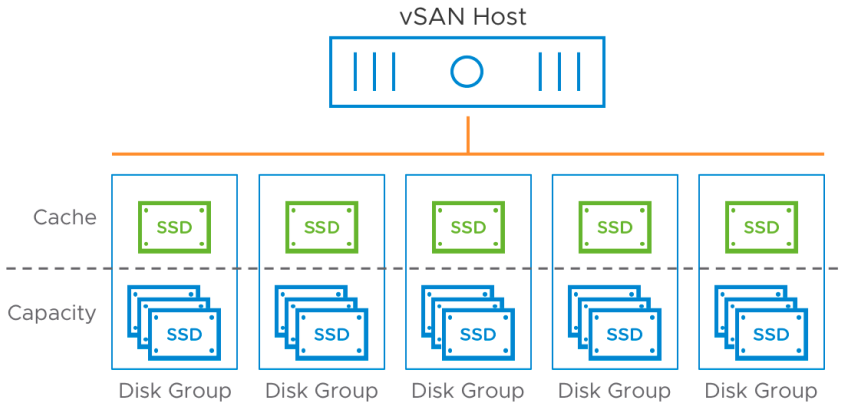
vSAN backed HCI provides a wide array of deployment options that span from a 2-node setup to a standard cluster with the ability to have up to 64 hosts in a cluster. Also, vSAN accommodates a stretched cluster topology to serve as an active-active disaster recovery solution. vSAN includes a capability called HCI Mesh that allows customers to remotely mount a vSAN datastore to other vSAN clusters, disaggregating storage and compute. This allows greater flexibility to scale storage and compute independently.

vSAN integrates with the entire VMware stack, including features such as vMotion, HA, DRS etc. VM storage provisioning and day-to-day management of storage SLAs can be all be controlled through VM-level policies that can be set and modified on-the-fly. vSAN delivers enterprise-class features, scale and performance, making it the ideal storage platform for VMs.

Servers with Local Storage

Each host contains flash drives (all flash configuration) or a combination of magnetic disks and flash drives (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 primarily for writes but can also serve as read cache for the buffered writes. 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.

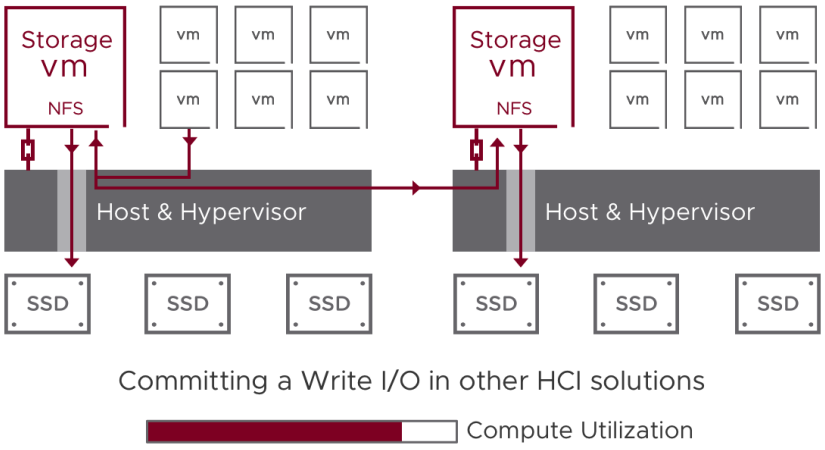
VMware is always looking for ways to not only improve the performance of vSAN but improve the consistency of its performance so that applications can meet their service level requirements.

Storage Controller Virtual Appliance Disadvantages

Storage in a Hyper-Converged 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.

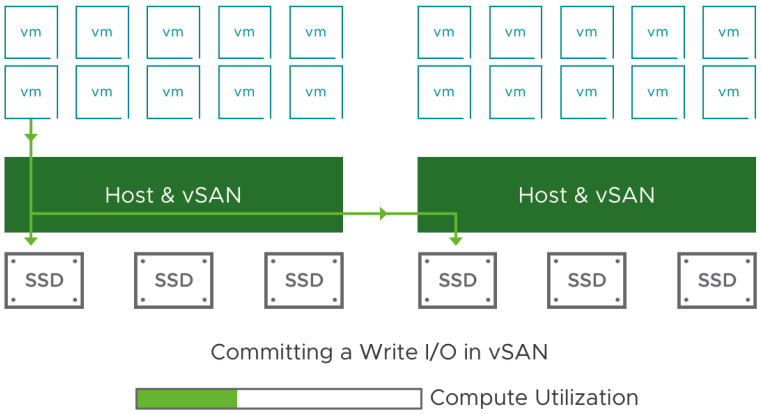


Committing a Write I/O in other HCI solutions

Compute Utilization

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). 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.

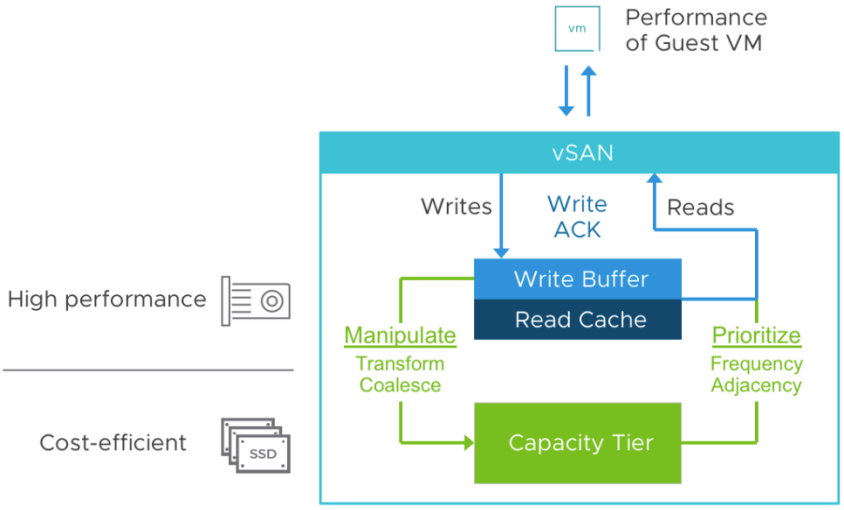


Committing a Write I/O in vSAN

Compute Utilization

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.

vSAN optimizes performance and cost through a two-tier architecture. Write operations are always staged to the buffer and these are eventually destaged, a process by which data is moved from cache tier to capacity tier based on highly optimized algorithms. Read IO are serviced by the Read cache if the reference blocks are available in the cache or retrieved from the capacity tier.

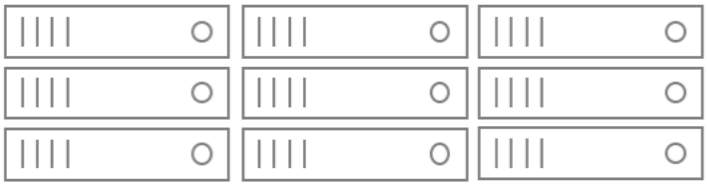


This architecture allows customers to have greater flexibility to adopt newer technology and upgrade hardware systematically.

Cluster Types

Standard Cluster

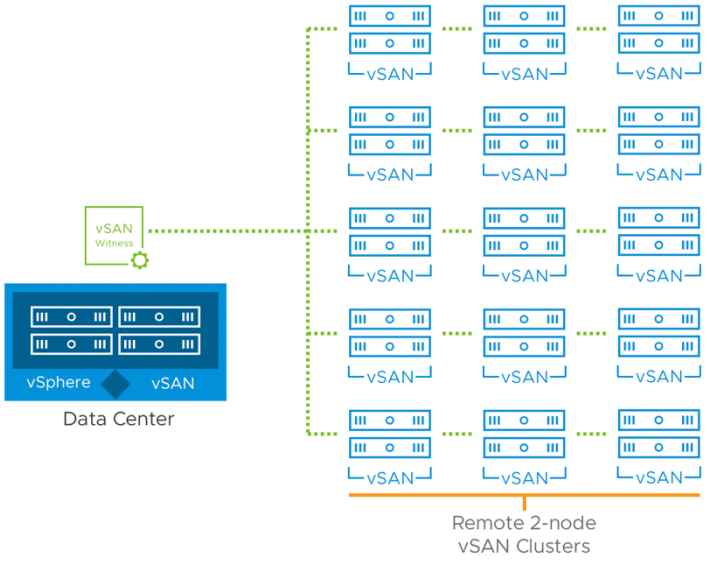
A standard vSAN cluster consists of a minimum of three physical nodes and can be scaled to 64 nodes.



All the hosts in a standard cluster are commonly located at a single location. 10Gb or higher network connectivity is 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.

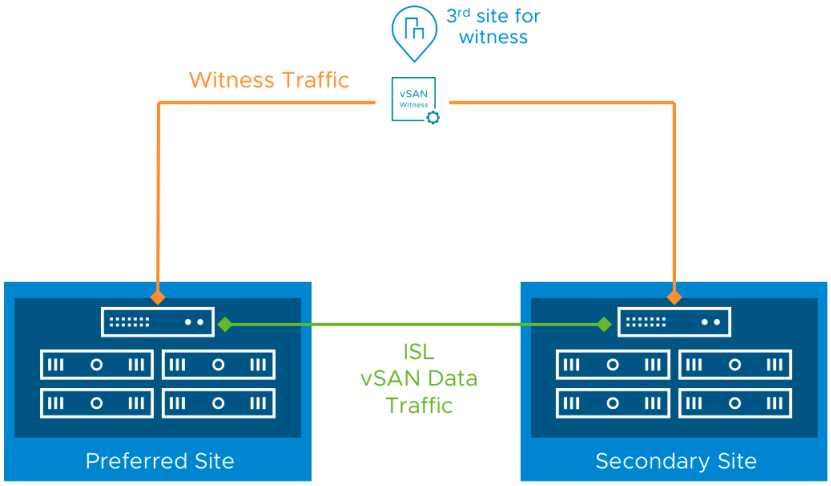


Direct connections between hosts eliminate the need to procure and manage an expensive network switch for a 2-node cluster, which lowers costs—especially in remote office deployments. *While 10Gbps connections may be directly connected, 1Gbps connections might require a crossover cable with older NICs that do not support Auto MDI-X. A "vSAN Witness Host" is required for a 2-node configuration to establish a quorum on certain failure conditions.

Each 2-Node deployment until vSAN 7 required a dedicated witness appliance. vSAN 7 Update 1 introduces the capability to use a shared witness appliance instance across multiple 2-Node deployments. Up to 64 2-node clusters can share a single witness appliance. This enhancement further simplifies design, eases manageability, and operations.

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.



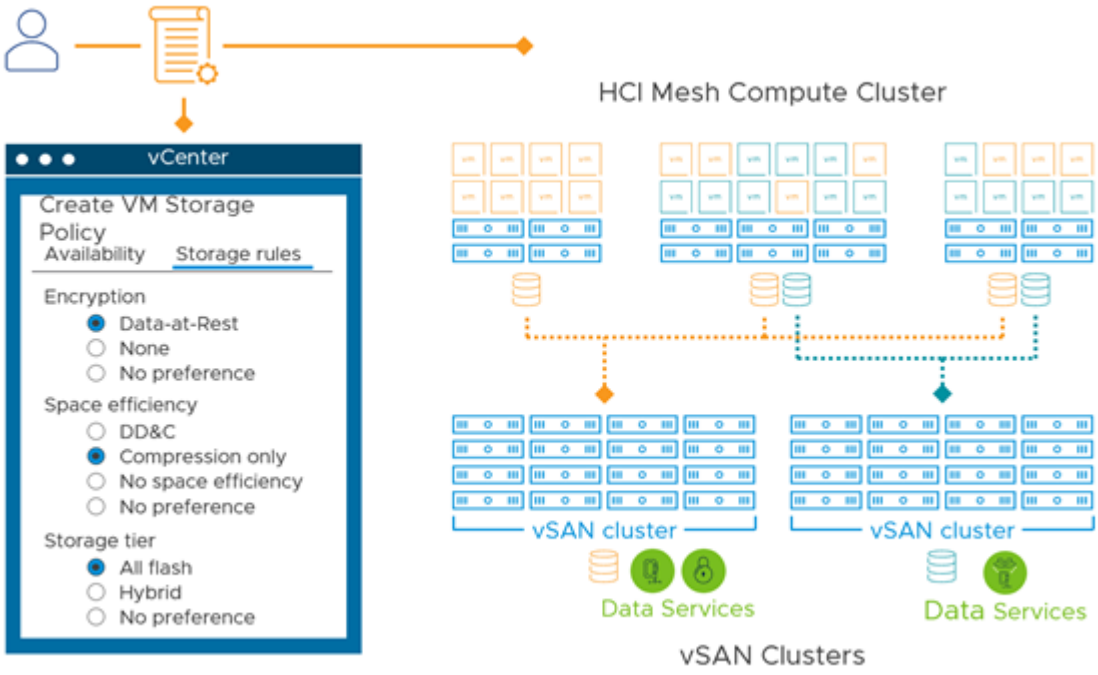
The two sites are well-connected from a network perspective with a round trip time (RTT) latency of no more than five milliseconds (5ms). A vSAN Witness Host is placed at a third site to avoid "split-brain" issues if connectivity is lost between the two Stretched Cluster sites. A vSAN Stretched Cluster 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.

HCI Mesh

vSAN redefined storage consumption as a cluster resource similar to compute and memory in a vSphere cluster. This is done by pooling local storage from hosts in a cluster into a single vSAN datastore for consumption within the same cluster. Extending from this base architecture, HCI Mesh provides additional flexibility to borrow capacity from other vSAN clusters. This capability helps define a relationship between one or more vSAN clusters to effectively share capacity and enable cross-cluster storage consumption. An administrator can mount a vSAN datastore from a server cluster to one or more client clusters; this allows a client cluster to consume capacity from the server cluster.

In vSAN 7 U2, traditional vSphere clusters can mount a remote vSAN datastore. HCI Mesh enables vSAN clusters to share capacity with computer-only clusters, or non-HCI-based clusters. Meaning, that HCI Mesh compute clusters can consume storage resources provided by a remote vSAN cluster, in the same way, that multiple vSphere clusters can connect to a traditional storage array. You can also specify storage rules for recommended data placement, to find a compatible datastore. Scalability for a single remote vSAN datastore has been increased to 128 hosts.

Also, virtual machines can be migrated using vMotion (compute only) across clusters without migrating the data. HCI Mesh enables customers with improved agility to independently scale storage and compute. It uses vSAN's native protocols for optimal efficiency and interoperability between clusters.



HCI Mesh uses a hub and spoke topology to associate the server cluster to its client cluster. Each server cluster can support up to five client servers, and each client cluster can mount up to five remote vSAN datastores.

Hardware Support

Compute

vSAN is natively embedded with vSphere and follows the same compatibility guide outlined in the VMware Compatibility Guide (VCG) for server hardware. There is no distinction between what is required for compute with vSphere and vSAN. This allows customers who already have an investment in vSphere supported hosts to easily add vSAN as a storage platform.

Networking

Networking requirements for vSAN include both hardware and connectivity. vSphere Host networking devices that are on the VMware Compatibility Guide (VCG) are supported for use with their approved versions of vSphere. Hosts participating in a vSAN cluster must be connected to a Layer 2 or Layer 3 network using IPv4 or IPv6. Hybrid Configuration requires at least 1 Gbps bandwidth. Host bandwidth for the vSAN network must be at least 1Gbps for a Hybrid configuration. All-flash configuration requires a minimum of 10 Gbps(shared/dedicated) bandwidth. Based on workload requirements, additional data services, and a futuristic outlook, a 25/40/100 Gbps network infrastructure is optimal and recommended. The network infrastructure for a vSAN environment should be designed with the same redundancy levels 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.

The nature of a distributed storage system like vSAN means that a network connecting the hosts is heavily relied upon for resilient connectivity, performance, and efficiency. vSAN 7 U2 supports clusters configured for RDMA-based networking: RoCE v2 specifically. Transmitting native vSAN protocols directly over RDMA can offer a level of efficiency that is difficult to achieve with traditional TCP-based connectivity over ethernet. The support for RDMA also means that the vSAN hosts have the intelligence to fall back to TCP connectivity if RDMA is not supported on one of the hosts in a cluster.

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

Storage

An ESXi host that contributes storage as part of a vSAN cluster can have a maximum of 5 disk groups. Each disk group has one cache device and one to seven capacity devices. These devices are available in various performance classes based on writes-per-

second throughput. 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 Compatibility List (VCG), vSAN storage devices, such as Host Bus Adapters (HBA), RAID Controllers, and storage devices must be on the VMware Compatibility Guide for vSAN.

Deployment Options

There are three deployment options available for vSAN: ReadyNodes, Appliance, and as-a-service. Customers have the broadest choice of consumption options for HCI with over 500 pre-certified vSAN ReadyNodes from all major server vendors. Alternatively, they can choose Dell EMC VxRAIL, a jointly engineered with full VMware integration. Also, customers could choose HCI-as-a-Service from leading cloud providers, including Amazon Web Services, Microsoft Azure, Google Cloud, Oracle Cloud, IBM Cloud, Alibaba cloud, and several more public clouds.

Although uncommon, customers can choose to build their own custom configuration or repurpose existing hardware. In such cases there should be careful consideration given to ensuring each hardware component is compliant with the VMware Compatibility Guide for vSAN. This requires additional effort and is less preferred over the above deployment options.

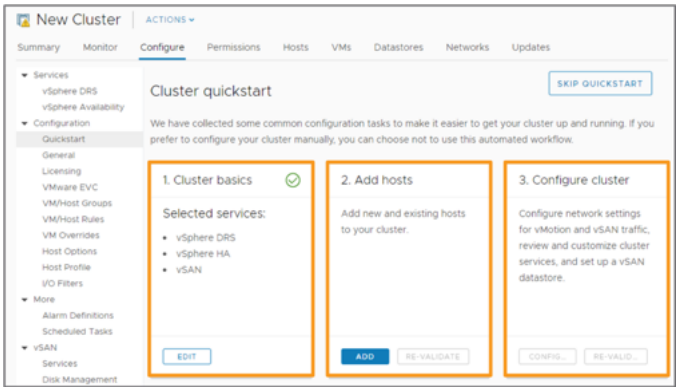
Deployment

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 Quick Start

vSAN deployment and setup are made easy with “Cluster Quickstart”, a guided cluster creation wizard. It is a step-by-step configuration wizard that makes it easier to create a production-ready vSAN cluster. Cluster Quickstart handles the initial deployment and the process of expanding the cluster as needs change. vSAN 7 Update 2 includes vLCM image options into the “Cluster Quickstart” wizard.

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

Add hosts – Add multiple hosts simultaneously

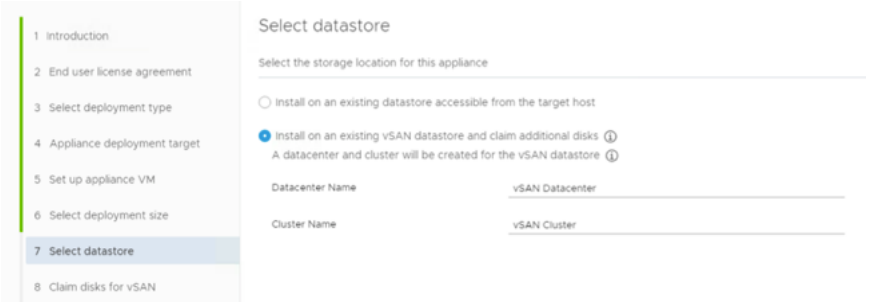
Configure cluster – Configure cluster, network and vSAN datastore settings

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. vSAN 7 Update 2 includes vLCM image options into the Easy Install wizard.

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. 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.

Using the vSAN Cluster Wizard

vSAN cluster wizard provides a simple guided workflow to enable vSAN. The Configure vSAN Wizard initiates the process to enable vSAN in the cluster. This can be a single site, 2-Node, or a Stretched Cluster. Additional services such as Deduplication and Compression and 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 can be used with either Hybrid or All-Flash hardware.

The vSAN Cluster Wizard encompasses all of the required tasks of configuring a vSAN Cluster. These services can also be enabled anytime during the lifecycle of the cluster with ease.

Availability

vSAN is an object-based storage system integrated into VMware vSphere. Virtual Machine residing in a vSAN datastore are comprised of a number of storage objects. These are the most prevalent object types found on a vSAN datastore:

- VM Home namespace, this contains virtual machine configuration files & logs
- Virtual machine swap
- Virtual disk (VMDK)
- Delta disk (snapshot)

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 Data Placement

Each object consists of one or more 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. These rules affect the placement of data that makes up an object to ensure that the requirements defined through the storage policy are adhered to.

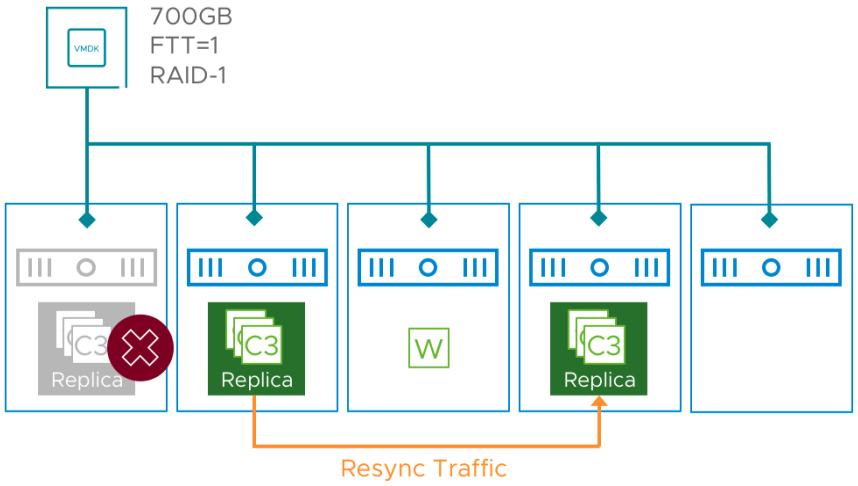
More details on cluster sizing minimums and recommendations can be found in the vSAN Design and Sizing Guide.

VM Swap Object Behavior

VM Swap objects are created when virtual machines are powered on. The swap file size is equal to the difference between allocated memory and reserved memory allocated to the VM. VM swap objects inherit the storage policy assigned to the VM home space 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 failure to tolerate setting 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. vSAN expects data on a host 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.

To restore redundancy, vSAN will begin the repair process for absent components after 60 minutes. For example, an object such as a virtual disk (VMDK file) protected by a RAID-1 mirroring storage policy will create another copy from the healthy version. This process can take a considerable amount of time, based on the amount of data to be copied. The virtual machine is powered-on and accessible as long the number of failures is within the configured threshold.

vSAN 7 Update 2 offers enhanced support for vSphere Proactive HA, where the application state and any potential data stored can be proactively migrated to another host. Depending on the settings of vSphere Proactive HA, it could also proactively place a host in maintenance mode or in quarantine mode depending on the risk that has been detected. A great addition to making sure applications are running at their highest levels of availability.

Repair Process

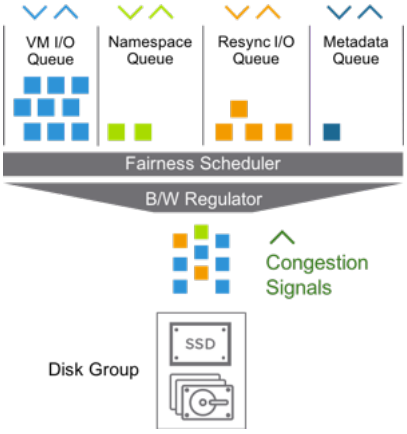
Object repair and rebuild mechanisms are continually optimized; if absent components resume while vSAN is rebuilding another copy, vSAN will determine if it is more efficient to continue building a new copy or update the existing copy that resumed. vSAN will restore redundancy using the most efficient method and cancel the other action. vSAN 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 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.

Resynchronization

Integrated within the hypervisor vSAN has a thorough insight into the I/O types and sources. This intelligence aids in prioritizing the I/O and enables the use of an adaptive mechanism called Adaptive Resync.



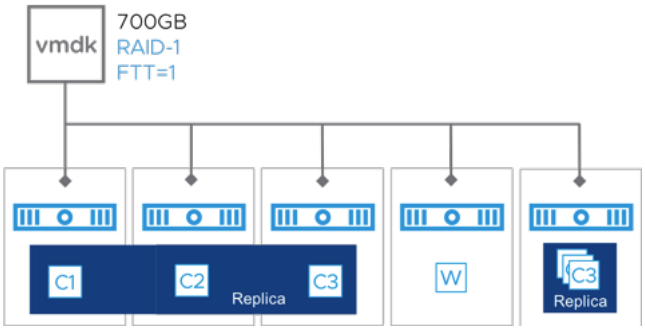
When I/O activity exceeds the sustainable Disk Group bandwidth, Adaptive Resync guarantees bandwidth levels for VM I/O and resynchronization I/O. During times without contention, VM I/O or resynchronization I/O are allowed to use additional bandwidth.

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 and appropriate prioritization to Virtual Machine I/O.

Replica Consolidation

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

vSAN strives to minimize data movement unless required, however some activities may distribute that data across more than one host. This includes rebalance operations that may split the data into different locations. 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?

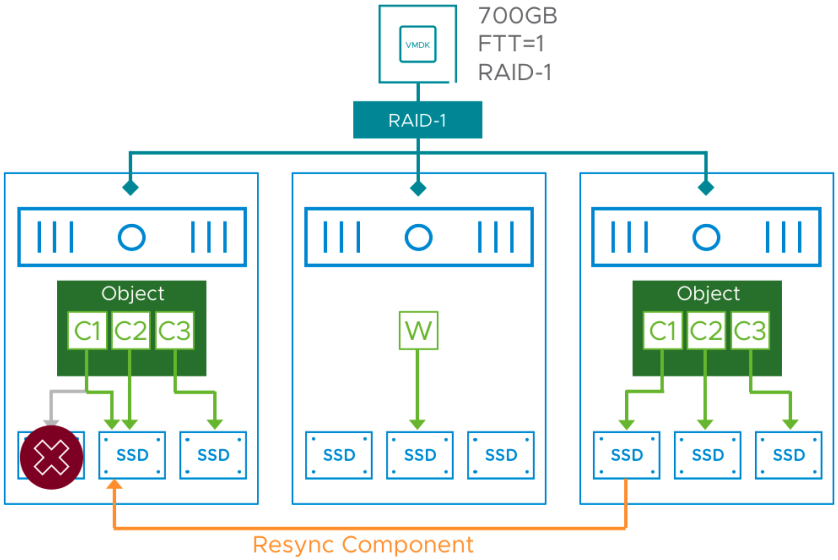


If a fault domain already contains a vSAN component replica, and there is additional capacity for the replica that needs to be evacuated, vSAN has the ability to consolidate them into a single replica. The smallest replicas are moved first, resulting in fewer data rebuilt, and less temporary capacity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

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

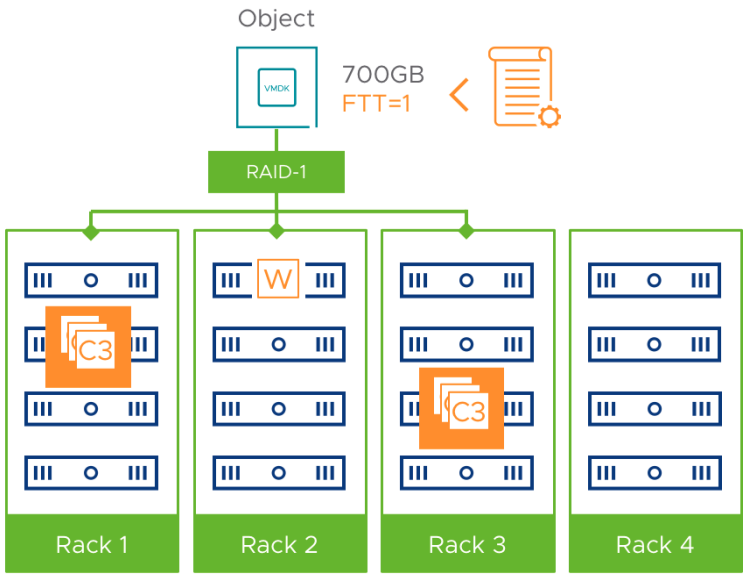
- RAID6 (4+2) requires (2x2+2) or 6 hosts

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.

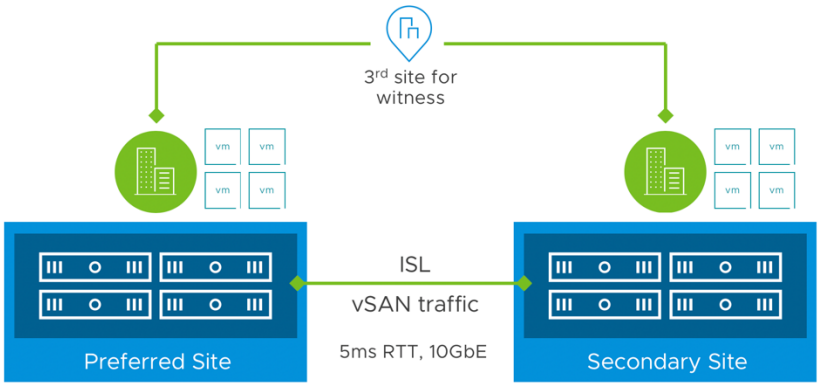
This feature is commonly referred to as "Rack Awareness". The diagram below shows component placement wherein servers in each rack are configured as separate vSAN fault domains.



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 floodwaters 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 20 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.

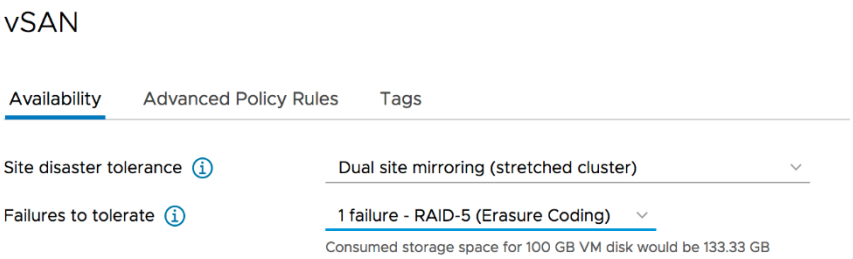
Up to 200ms RTT latency is supported between the witness site and data sites. The bandwidth requirements between the witness site and data sites vary and depend primarily on the number of vSAN objects stored at each site. A minimum bandwidth of 100Mbps is required and the general principle is to have 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.

vSAN Stretched Clusters uses mirroring 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 additionally provides the capability to protect data within a site to sustain failures local to a site either through mirroring or erasure coding. This 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 Client. The figure below shows rules in a storage policy that is part of an all-flash stretched cluster configuration. The Site disaster tolerance is set to Dual site mirroring, 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.



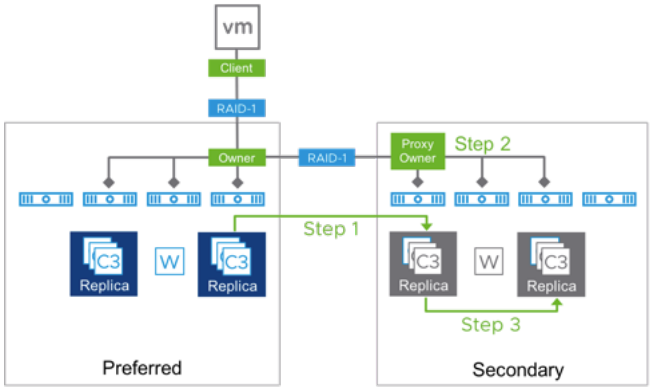
Stretched Cluster Data Sites

A maximum of 40 hosts may be used in a single vSAN Stretched Cluster across the data sites with up to 20 hosts per site since vSAN 7 Update2. 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.

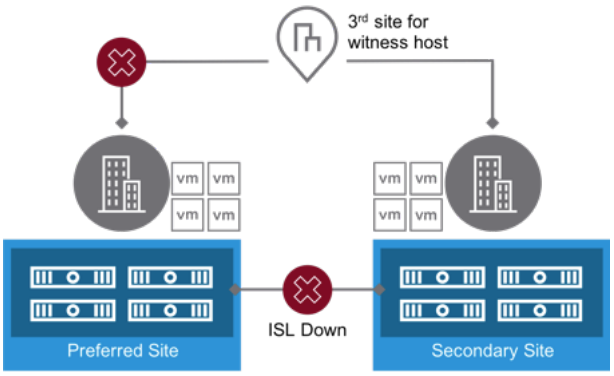
Network bandwidth and round-trip time (RTT) latency are the primary factors that must be considered when deploying a vSAN Stretched Cluster. Because writes are synchronous, they must be committed before they may be acknowledged. When RTT latencies are higher than five milliseconds (5ms) performance issues 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 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 failover 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. vSAN 7 Update 2 has integration with data placement and DRS (Distributed Resource Scheduler) so that after a recovered failure condition, DRS will keep the VM state at the same site until data is fully resynchronized, which ensures that all read operations do not traverse the inter-site link (ISL). Once data is fully resynchronized, DRS moves the VM state to the desired site in accordance with the DRS rules. This improvement dramatically reduces unnecessary read operations occurring across the ISL and frees up ISL resources to continue with its efforts to complete any resynchronizations post-site recovery.



Stretched Cluster Site Affinity

With 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 Cluster Witness Site

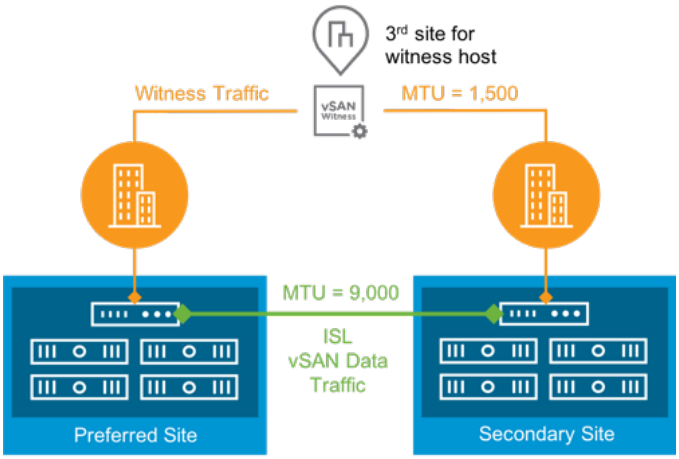
A vSAN Witness host must be deployed to a third site. The vSAN Witness Host allows the cluster achieve quorum when one of the two main data sites are offline as well as acts as "tie-breaker" in scenarios where a network partition occurs between the two data sites. This is sometimes referred to as a "split-brain" 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 principle is to have at least 2Mbps of available bandwidth for every 1000 vSAN objects. The vSAN Stretched Cluster Bandwidth Sizing guide provides more details on networking requirements.

Witness Traffic

Witness Traffic can be separated from data site traffic and supports different MTU sizes.



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 attribute gives additional flexibility and choice in

allowing for a wider variety of customer topology conditions and reduce potential network issues.

Operations

Cluster Operations

vSphere hosts in vSAN clusters are just that, vSphere hosts. Many of the normal host management operations are the same as traditional vSphere clusters. Administrators must take into consideration that each vSAN host is contributing to the overall resources present in the cluster. Effective maintenance of vSAN hosts must take into account the impact to the cluster as a whole.

Maintenance Mode Operations

VMware vSphere includes a feature called Maintenance Mode that is useful for planned downtime activities such as firmware updates, storage device replacement, and software patches. Assuming you have VMware vSphere Distributed Resource Scheduler (DRS) enabled (fully automated), Maintenance Mode will migrate virtual machines from the host entering maintenance mode to other hosts in the cluster. VMware vSAN uses locally attached storage devices. Therefore, when a host that is part of a vSAN cluster enters maintenance mode, the local storage devices in that host cannot contribute to vSAN raw capacity until the host exits maintenance mode. On a similar note, an administrator can choose to evacuate or retain the data on the node entering maintenance mode based on the nature of the maintenance activity.

Evacuate all data to other hosts

This option moves all of the vSAN components from the host entering maintenance mode to other hosts in the vSAN cluster. This option is commonly used when a host will be offline for an extended period of time or permanently decommissioned.

Ensure data accessibility from other hosts

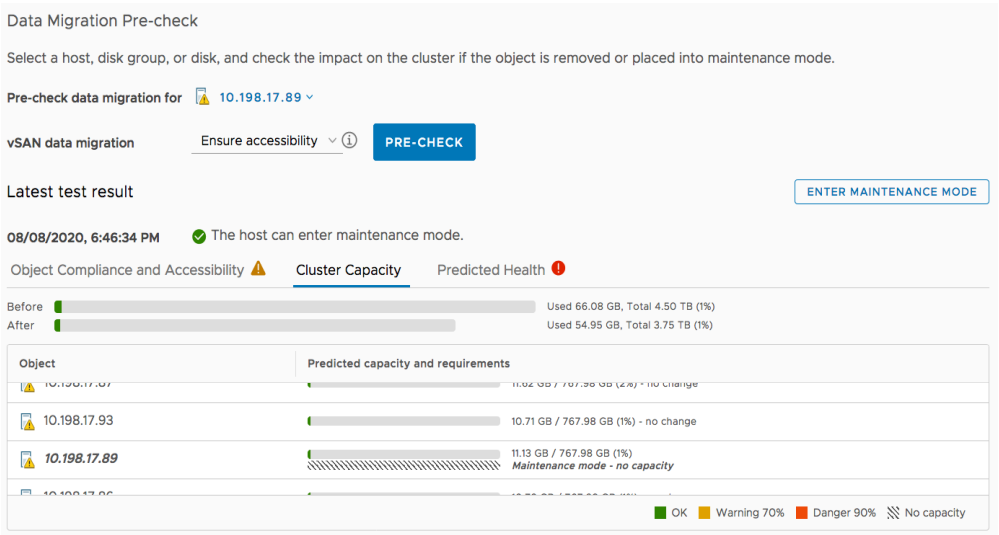
vSAN will verify whether an object remains accessible even though one or more components will be absent due to the host entering maintenance mode. If the object will remain accessible, vSAN will not migrate the component(s). If the object would become inaccessible, vSAN will migrate the necessary number of components to other hosts ensuring that the object remains accessible. To mitigate the reduced availability, an enhanced data durability logic ensures that writes are replicated to an interim object. In the event of failure to the host with the only available replica, the interim object will be used to update the components to the host in maintenance mode. This helps protect against a failure in the interim state of reduced availability.

No data evacuation

Data is not migrated from the host as it enters maintenance mode. This option can also be used when the host will be offline for a short period of time. All objects will remain accessible as long as they have a storage policy assigned where the Primary Level of Failures to Tolerate is set to one or higher.

Predictive Analysis for Maintenance Operations

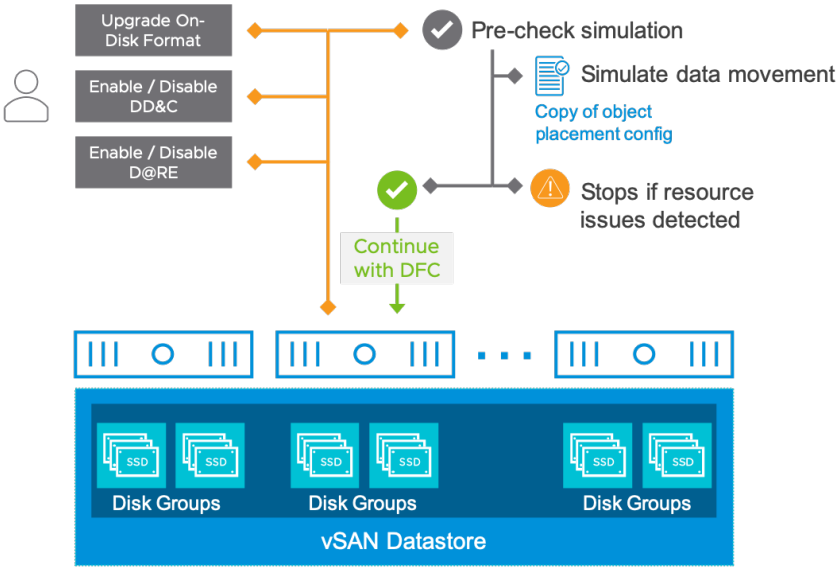
A pre-check engine provides a detailed cluster-wide analysis of the state of the cluster when performing maintenance.



An administrator can have a clear visibility of the impact of a maintenance mode operation at a host, diskgroup or at an individual drive level. A pre-check report shows how object compliance and accessibility, cluster capacity, and the predicted health of the cluster will change when a maintenance task is performed. This gives administrators a clear, informed view of what the results before their taking any actions.

Disk Format Change Pre-Check

An update to the version of the underlying disk format must be upgraded from time to time. Additionally, when changing the state of data services like Deduplication and Compression or Encryption, the underlying disk format may need to be upgraded or modified. Most Disk Format Change (DFC) conversions only require a small metadata update, with no requirement to move object data around. A DFC pre-check runs a simulation to see if the conversion process would complete properly and only proceeds if the conversion can be completed properly.



This pre-check prevents clusters from ending up in an incomplete DFC conversion state.

Lifecycle Management

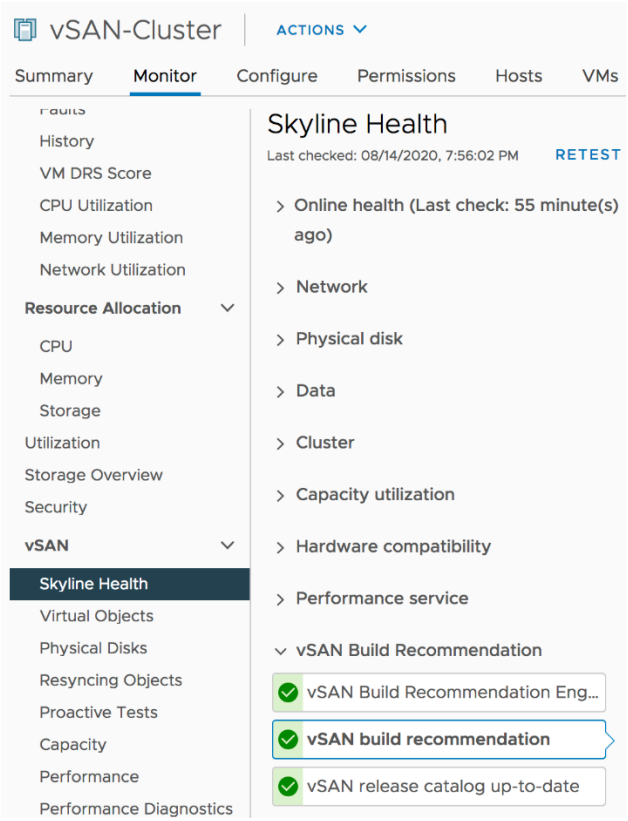
Lifecycle management is a time-consuming task. It is common for admins to maintain their infrastructure with many tools that require specialized skills. VMware customers currently use two different interfaces for day two operations: vSphere Update Manager (VUM) for software and drivers and server vendor-provided utility for firmware updates. In this latest release, VMware HCI sets the foundation for a new, unified mechanism to update software and firmware management that is native to vSphere called vSphere Lifecycle Manager (vLCM).

vSphere Update Manager

VMware vSphere Update Manager™ has been a preferred tool of choice 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.

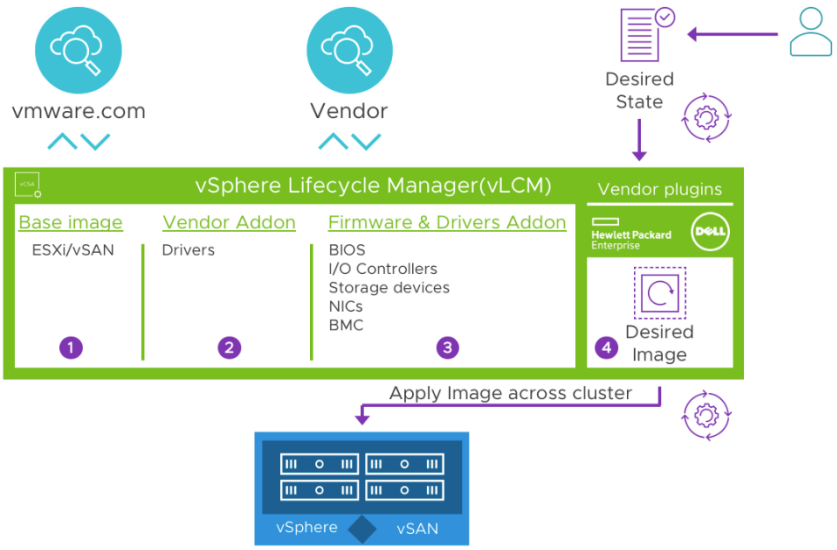


The first step to enable this functionality is entering valid My VMware Portal credentials. The vSAN Health Check produces a warning if credentials have not been entered. A vSAN Cluster baseline group is created using information from the VMware Compatibility Guide and underlying hardware configuration.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

vSAN allows baseline updates to be configured per cluster. Baselines can be created that will properly upgrade clusters to the newest ESXi versions or just patches and updates for the current version of ESXi. This provides better overall platform flexibility for hardware that has not yet been validated for newer versions of vSAN, as well as when vSAN is used with other solutions that have specific vSphere version requirements.

vSphere Lifecycle Manager

vSphere Lifecycle Manager (vLCM) is native to vSphere and provides a powerful framework to simplify cluster lifecycle management. vLCM is next-generation replacement to VUM. While VUM is highly efficient in managing software stack in terms of patches and upgrades. vLCM provides complete visibility and upgrade capabilities of the entire server stack including device firmware. vLCM is built off a desired-state model that provides lifecycle management for the hypervisor and the full server stack of drivers and firmware for your hyperconverged infrastructure. vLCM can be used to apply a desired image at the cluster level, monitor compliance, and remediate the cluster if there is a drift. This eliminates the need to monitor compliance for individual components and helps maintain a consistent state for the entire cluster in adherence to the VCG.



vLCM is a service that runs in the vCenter Server. In order to update the full server-stack firmware, a server vendor plugin called a Hardware Support Manager (HSM) must be installed. This plugin is provided by the associated hardware vendor. Broadly vLCM supports HPE, Dell Lenovo & Hitachi hardware. The precise make and model that is supported to be managed through vLCM can be validated through VMware Compatibility Guide.

As of vSAN 7 Update 2, vLCM supports lifecycle management of vSphere with Tanzu enabled clusters, using NSX-T networking. vLCM supports and is fully aware of NSX-T managed environments, vSAN fault domains and stretched clusters. Awareness of vSAN topologies ensures a preferred order of host updates that complete an entire fault domain prior to moving onto the next. Additional pre-checks and the support of more clusters that can be remediated concurrently improve the robustness and scalability of vLCM.

The following table provides a comparison of capabilities between VUM and vLCM:

Management Tasks	**vSphere Update Manager (VUM)**	**vSphere Lifecycle Manager (vLCM)**
Upgrade and patch ESXi hosts	Yes	Yes
Install and update third party software on hosts	Yes	Yes
Upgrade virtual machine hardware and VMware Tools	Yes	Yes
Update firmware of all ESXi hosts in a cluster	No	Yes
One desired image to manage entire cluster	No	Yes
Check hardware compatibility of hosts against vSAN Hardware Compatibility List	Yes	Yes
Checks for drift detection	No	Yes
Shared witness upgrade	Yes	No

Storage Operations

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.

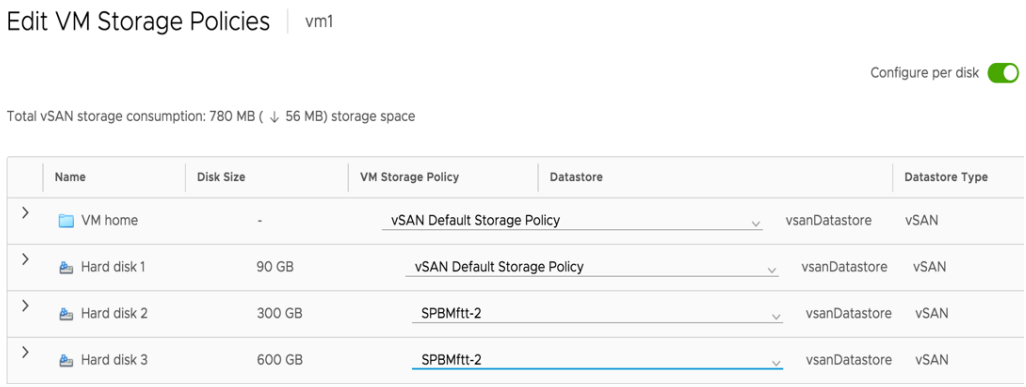
Storage Policy Based Management

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, date placement techniques to improve performance. A storage policy contains one or more rules that define service levels.

Storage policies are primarily created and managed through the vCenter Server. 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.

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 demonstrates the level of granularity at which a storage policy can be assigned. In the below example, the VM home namespace and Hard disk 1 have been assigned a default storage policy whereas the Hard disk 2 and Hard disk 3 have been assigned a policy that can tolerate 2 failures.

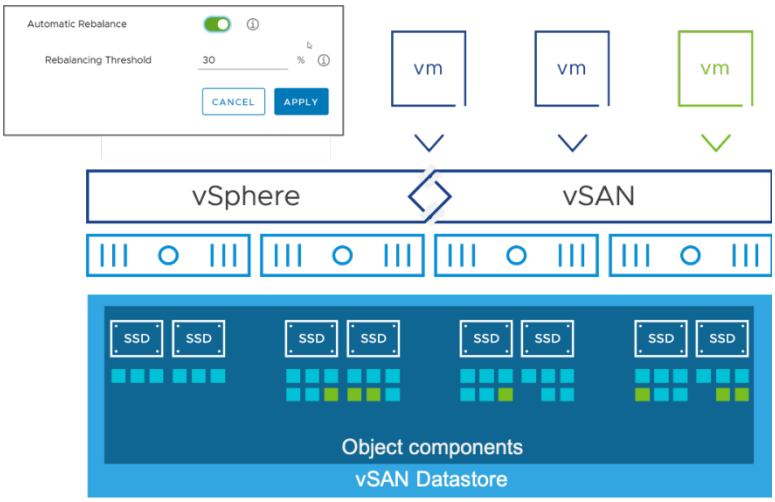


Each Storage Policy can be easily changed and applied to all the vSAN Objects that are assigned to each of these policies. Rather than changing Storage Policies.

Balanced Capacity Distribution

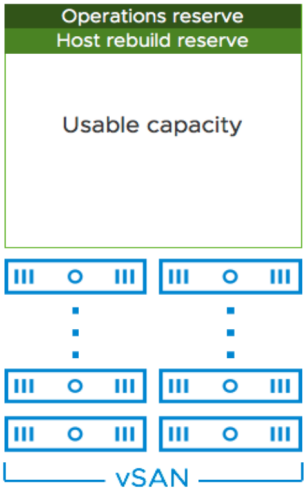
When deploying data to vSAN clusters there are many factors that are considered specific to placement across the cluster. Storage Policies determine where copies of data are placed. Some vSAN deployments are better suited to have a more even distribution of data. A good example of this would be a VMware Horizon virtual desktop deployment on vSAN tens, hundreds, or thousands of very similar virtual desktops. The majority of the virtual desktops likely have the same uniform configuration. In contrast, other environments that have a large mix of workload types, each with their own storage profile often result in a less uniform distribution of data. As more workloads are deployed, or destroyed and redeployed, the overall capacity becomes less balanced.

vSAN automatically balances capacity when any of the capacity devices reach a threshold of 80% utilization. This is termed as reactive rebalancing. Administrators can also choose to maintain a stricter and enforced balance across devices through proactive rebalancing. This capability allows an administrator to set a threshold of variance across capacity devices, vSAN invokes proactive balancing when the threshold is reached.



Capacity Reservation

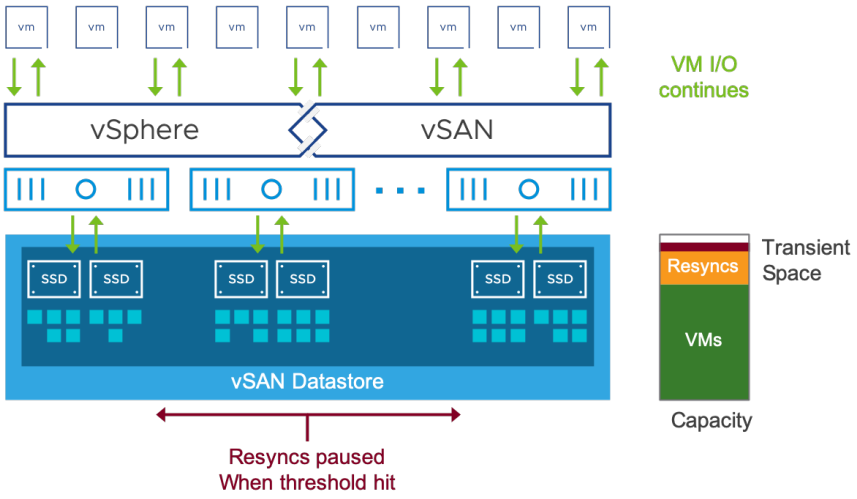
VMs and containerized workloads tend to be deployed, deleted, or modified at an increased frequency that specifically impacts capacity. A certain amount of transient capacity is required to accommodate such transient changes and perform internal operations. vSAN allows administrators to granularly configure Operations reserve and Host rebuild reserve to accommodate sufficient space for internal operations and failover capacity, respectively.



These reservations act as soft thresholds that prevent the provisioning and powering on VMs if the operation could consume the reserved capacity. vSAN does not prevent powered on VM operations, such as I/O, from the guest operating system or applications from consuming the space even after the threshold limits are reached.

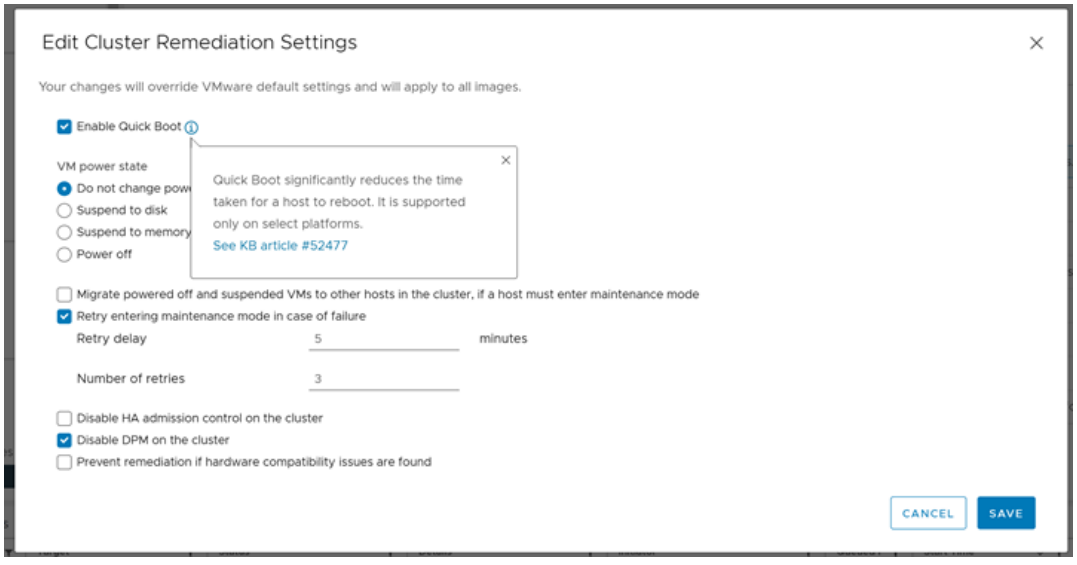
Capacity Aware Operations

vSAN actively monitors free capacity and adjusts operations accordingly, such as cases where resyncs could exceed recommended thresholds. In scenarios such as this, resync operations are paused until sufficient capacity is available, therefore preventing I/O disruptions from disk full situations.



Quick Boot – Suspend to memory

Host updates are a necessary task in the data center. When a host is contributing CPU, memory, and storage resources, the more quickly a host can be updated the better. Less time offline means more time in production. This is where vSphere Quick Boot comes into play, where host restarts during an upgrade can be accelerated by bypassing the hardware and firmware initialization process of a server. vSAN 7 U2 provides better integration and coordination for hosts using Quick Boot to speed up the host update process. By suspending the VMs to memory, and better integration with the Quick Boot workflow, the amount of data moved during a rolling upgrade is drastically reduced due to reduced VM migrations, and a smaller amount of resynchronization data. When the circumstances are suitable, Quick boot can deliver a much more efficient host update process. The Quick Boot feature is not enabled on hosts by default. Administrators will need to enable this manually, after ensuring that their hardware and hypervisor settings are compatible with the feature.



Health, Support, and Troubleshooting

Organizations rely on vendor support more than ever to maintain data center performance through knowledgeable and accessible teams, advanced analytics and artificial intelligence (AI) technologies. VMware incorporates a wide array of such innovations through vSAN ReadyCare program.

vSAN ReadyCare

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.

vSAN has invested significantly in delivering new capabilities that can identify potential issues and recommend solutions before a support request is required.

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.

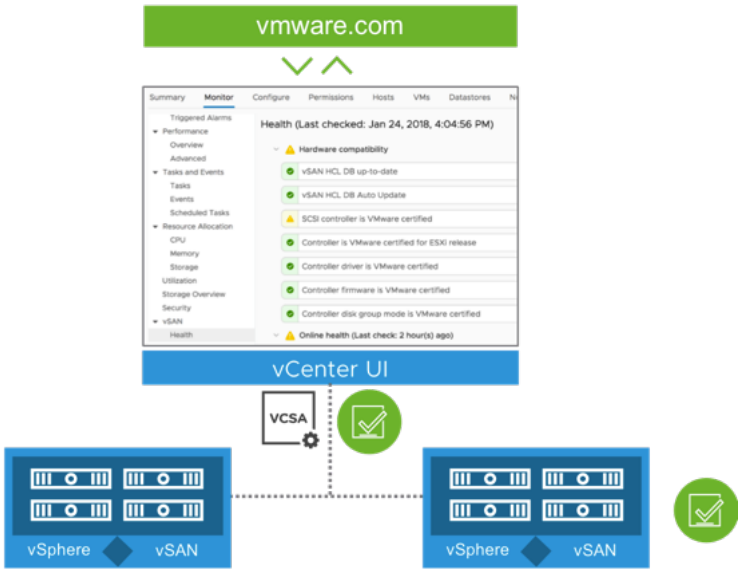
Here is a look at some of the innovations and capabilities available today with vSAN helping us deliver the enhanced support experience of vSAN ReadyCare:

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. 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 deactivated or vCenter has no Internet accessibility.



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.

Health checks history

In vSAN 7 Update 2 the Health checks history allows administrators to easily view a timeline of discrete alerts, which is especially helpful to understand a series of events that may have occurred. The user interface provides the ability to understand relationships with other alerts, providing insight to the administrator of the core issue that could be the cause of the alerts.

Skyline Health

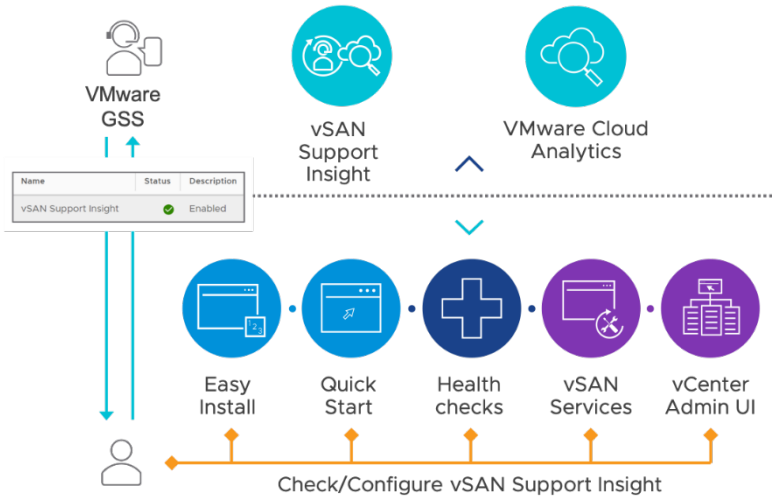
VMware Skyline™ is a proactive support service aligned with VMware Global Support Services. VMware Skyline automatically and securely collects, aggregates, and analyzes product usage data which proactively identifies potential problems and helps VMware Technical Support Engineers improve the resolution time.

This enables richer, more informed interactions between customers and VMware without extensive time investments by Technical Support Engineers. These capabilities transform support operations from reactive, break/fix to a proactive, predictive, and prescriptive experience that produces an even greater return on your VMware support investment.

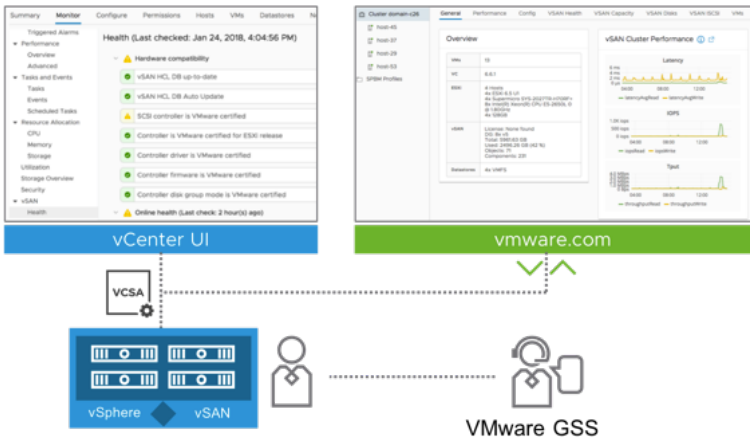
Customers that participate in the Customer Experience Improvement Program (CEIP) also get added benefit by proactively providing VMware Global Support Services (GSS) with some data about their environment.

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.

For those customers who have not chosen to configure CEIP can easily enable it later though various vSAN workflows.



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.

Skyline Health Diagnostic (SHD) tool for isolated environments

The Skyline Health Diagnostic (SHD) tool for health check management and diagnostics for isolated environments is available in vSAN 7 Update 2. The Skyline Health Diagnostics tool is a self-service tool that brings some of the benefits of Skyline health directly to an isolated environment. It is run by an administrator at a frequency they desire. It will scan critical log bundles to detect issues and give notifications and recommendations to important issues and their related KB articles. This tool helps our customers with isolated environments resolve their issues faster without the need to contact GSS.

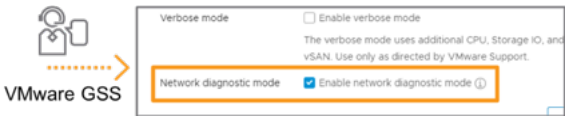
Performance for Support Diagnostics

"Performance for Support" dashboards natively within the UI 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 immediately available to GSS personnel to better support the environment. GSS Support personnel 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.

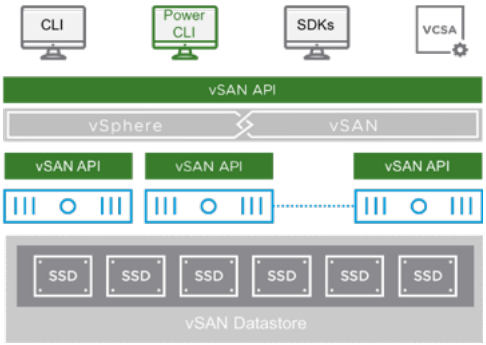
Additionally, an on-demand network diagnostics mode can temporarily collect granular network intelligence, that can be submitted with log bundles for deeper visibility at both the host and network layers.



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.

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

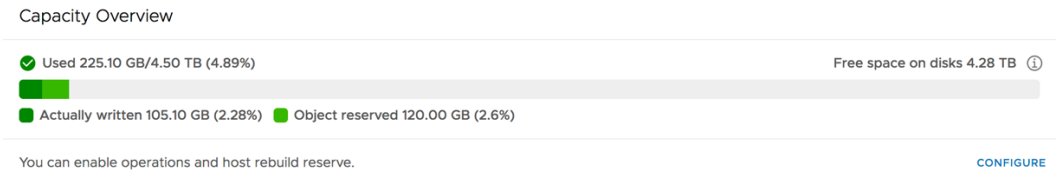
PowerCLI includes cmdlets for performance monitoring, some cluster configuration information tasks such as upgrades, and vSAN iSCSI operations. A Host-level API can query cluster-level information. S.M.A.R.T. device data can also be obtained through the vSAN API. SDKs are available for several programming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

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.

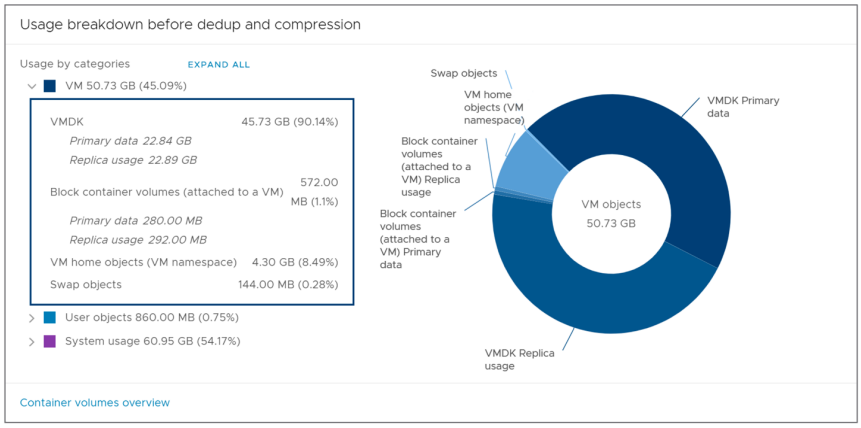
The vSAN Capacity Overview quickly and easily shows the amount of capacity being consumed on a vSAN datastore.



Administrators can also 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 a 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. Deduplication and Compression Overview shows the amount of space that would be required if Deduplication and Compression were to be deactivated.

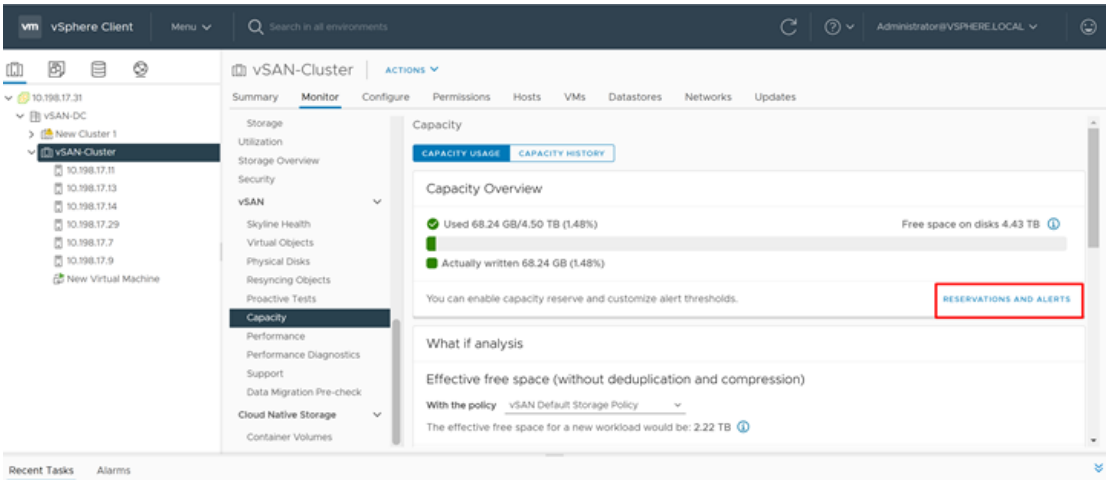
An updated and easier-to-read capacity breakdown provides more details on the distribution of object types.

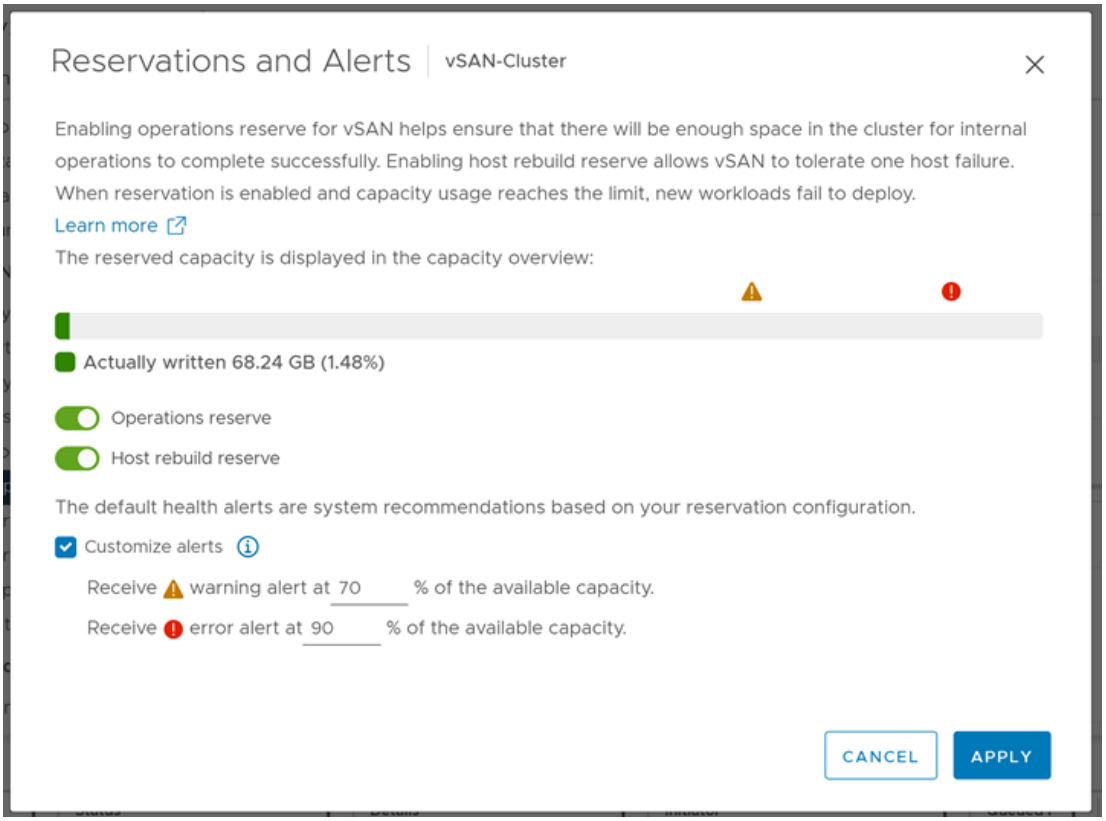


Note: Percentages are of used capacity, not of total capacity.

In vSAN 7 Update 2 administrators can see how oversubscribed capacity is for the cluster. vSAN is inherently thin provisioned, meaning that only the used space of an object is counted against the capacity usage. Oversubscription visibility helps the administrator understand how much storage has been allocated, so they can easily see the capacity required in a worst-case

scenario and adhere to their own sizing practices. vSAN 7 U2 also provides customizable warning and error alert thresholds directly in the Capacity Management UI in vCenter Server. Redundant alerting for capacity thresholds are eliminated to help clarify and simplify the condition reported to the administrator.

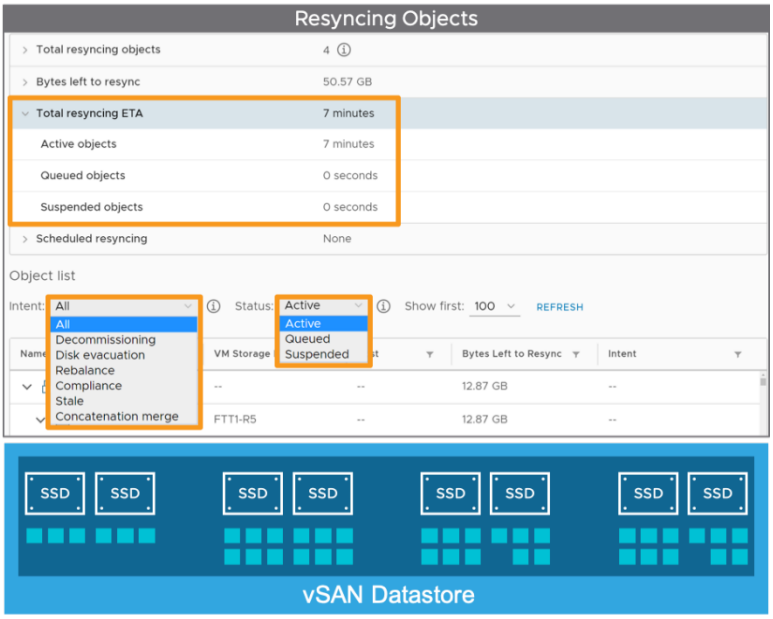




Resynchronization Operations

vSAN Resyncing Objects dashboard provides administrators with a thorough insight of current resync operations and their estimated time to completion. In addition, details of resync operations that are waiting in the queue can also be retrieved.

Types of resyncs can be easily filtered on the cause of the resync, such as rebalancing, decommissioning, or component merging. All of these improvements add up to more visibility, and more control.

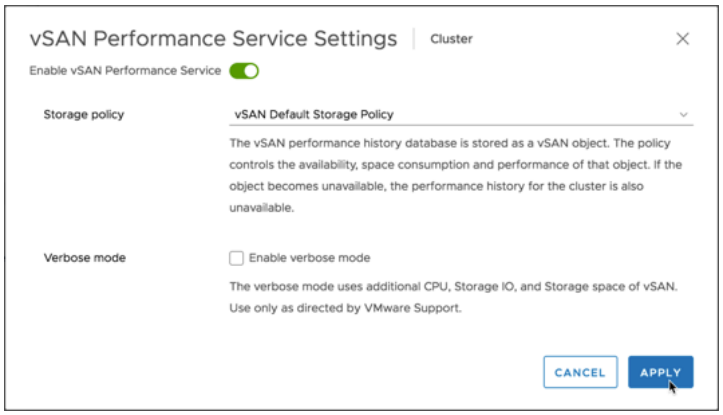


Performance Service

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.



Changes, available in vSAN 7 Update 1, were made in the hypervisor to better accommodate the architecture of AMD-based chipsets. Improvements were also made in vSAN (regardless of chipset used) to help reduce CPU resources during I/O activity for objects using the RAID-5 or RAID-6 data placement scheme, this is especially beneficial for workloads issuing large sequential writes. vSAN7 U2 includes enhancements to help I/O commit to the buffer tier with higher degrees of parallelization. All of these add up to fewer impediments as I/O traverses the storage stack, and a reduction of CPU utilization which can drive up the return on investment.

Performance Metrics

Performance metrics aid in monitoring health, adherence to SLAs, and troubleshooting. vSAN provides insightful and quantifiable metrics with granularity. These metrics can be monitored at Cluster, Host, Virtual Machine, and at a virtual disk level.

Cluster Metrics

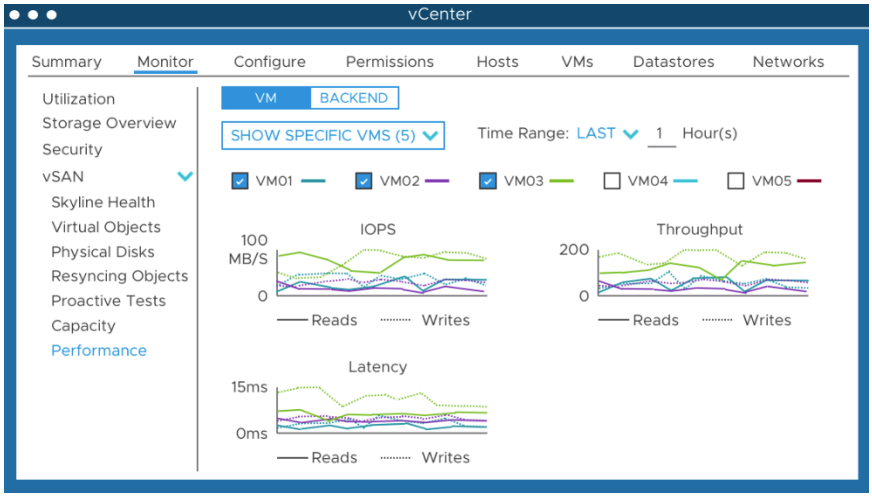
Cluster level graphs show IOPs, throughput, latency, congestion, and outstanding I/O. “vSAN – Virtual Machine Consumption” graphs show metrics generated by virtual machines across the cluster. In addition to normal virtual machines reads and writes, “vSAN – Backend” consumption adds traffic such as metadata 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

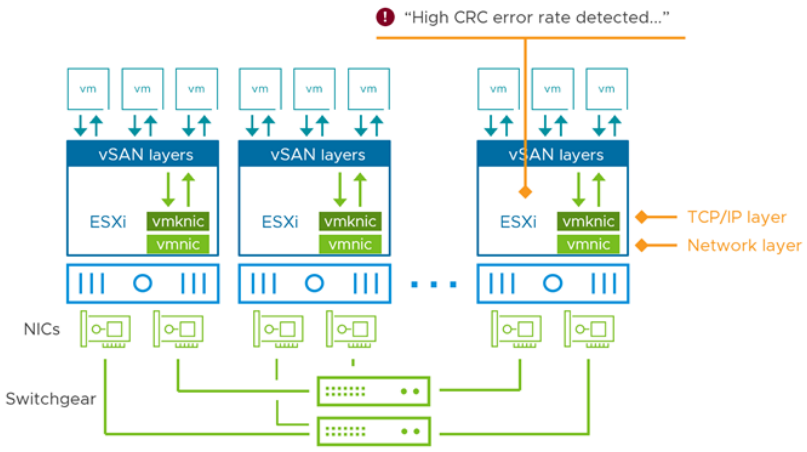
vSAN provides greater flexibility and granularity to monitor specific Virtual Machine performance metrics. This is particularly helpful in a larger infrastructure and densely populated clusters. The administrator can choose VMs for a targeted analysis or perform a comparative study across multiple VMs. This helps expedite troubleshooting and isolating performance issues.



Virtual machine metrics include IOPS, throughput, and latency. It is also possible to view information at the virtual disk level.

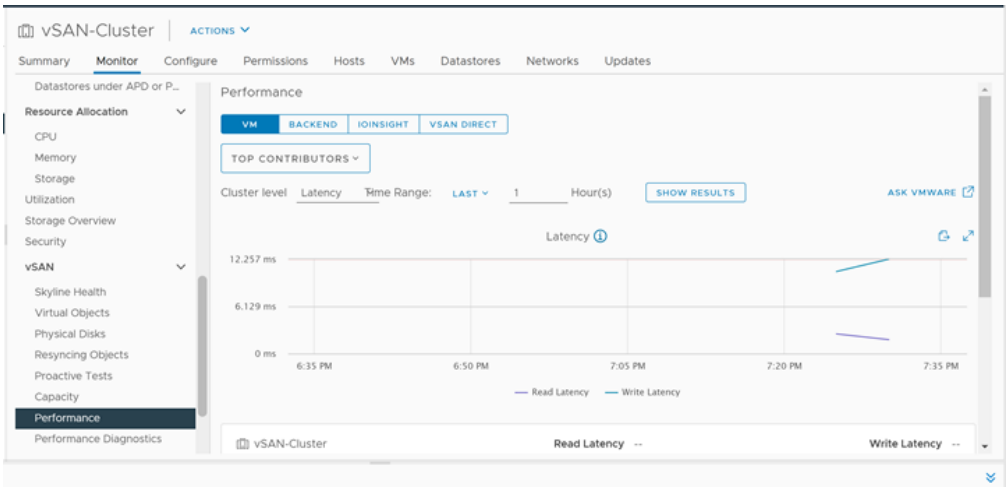
Enhanced network monitoring

Several metrics are included to monitor the physical networking layer, including CRC and carrier errors, transmit and receive errors, and pauses. Visibility is added at the TCP/IP layers including ARP drops and TCP zero frames. Not only do these metrics show up in time-based performance graphs, but health alarms ensure an administrator knows they are approaching or surpassing critical thresholds.



A separate view can be set to observe latency, throughput, or IOPS, and can view either the top VMs, or the top disk groups used -

the "top contributers". These will be shown clearly in an ordered list with a customizable time-based view to understand the history of the metric desired.

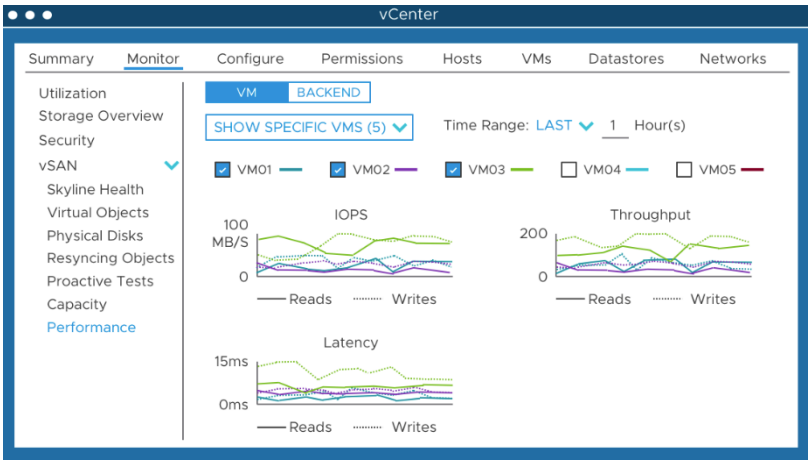


Performance Analytics

In addition to metrics available in the UI, vSAN provides additional utilities for a deeper insight and analysis of workloads. These utilities are typically used for targeted analytics or troubleshooting.

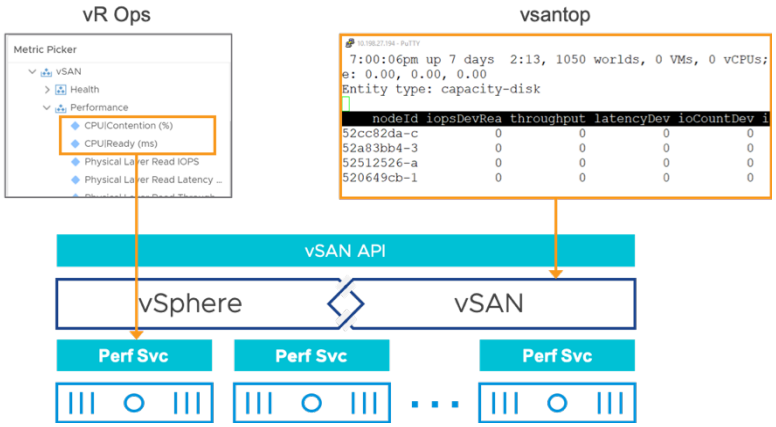
vSAN IO Insight

vSAN IO Insight provides more detailed insight into workload characteristics such as I/O size, Read/Write ratio, IO throughput and latency distribution (Random and Sequential). The size of storage I/O can have a profound impact on the perceived level of performance of the VM and presenting I/O size distributions for a discrete workload can help identify if the I/O sizes are a contributing factor to latency spikes.



vsantop

vsantop is another utility built with an awareness of vSAN architecture to retrieve focused metrics at a granular interval. This utility is focused on monitoring vSAN performance metrics at an individual host level.



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.

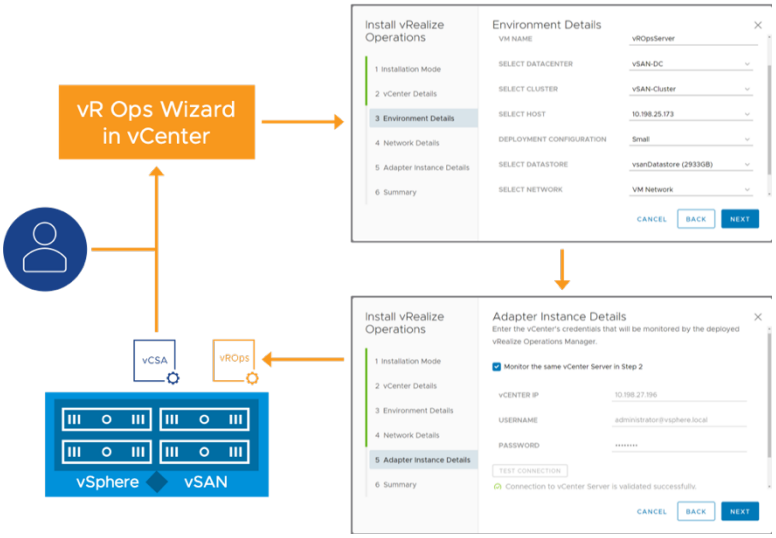
HCIBench has API-level integration with Performance Diagnostics. Administrators run a benchmark test and HCIBench will save the time range in vCenter Server. Detailed results of the test including supporting graphs are easily retrieved from the Performance Diagnostics section in the vSphere Web Client. 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.

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.

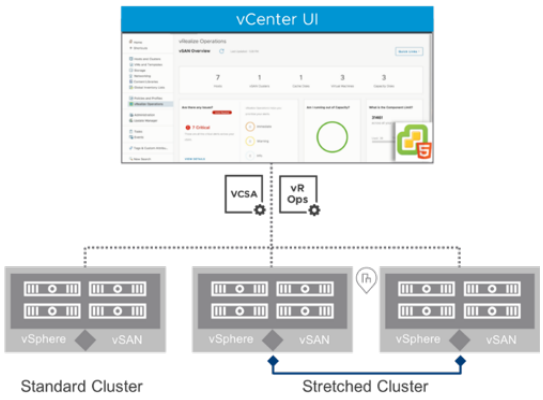
vRealize Operations Management Packs could be added to extend the capabilities of vRealize Operations by including prebuilt dashboards that focus on design, management, and operations for a variety of solutions and use cases.



vRealize Operations within vCenter provides an easy way for customers to see basic vRealize intelligence with vCenter. vCenter 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 vRealize Operations instance is 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.

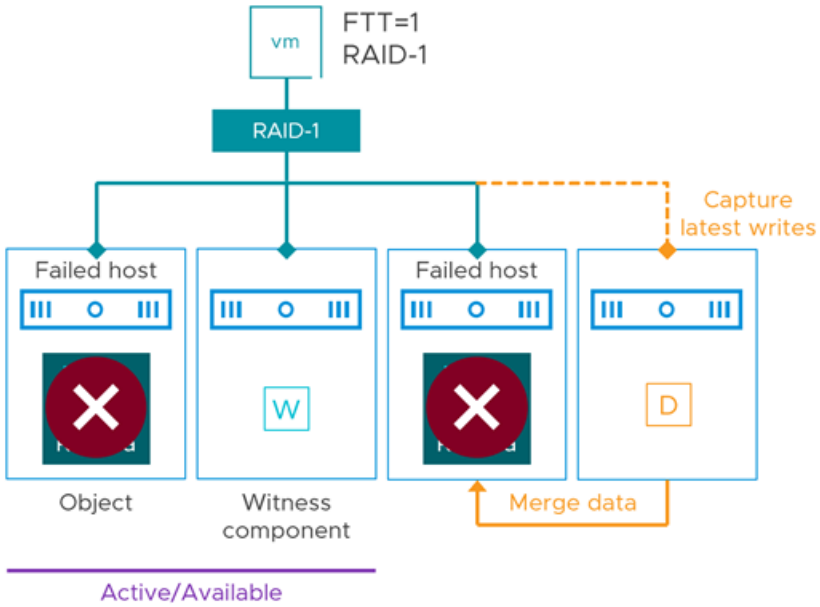
vRealize Operations dashboards have the ability to differentiate between normal and stretched vSAN clusters, displaying appropriate intelligence for each.



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.

Enhanced Data Durability During Unplanned Events

vSAN 7 U2 ensures the latest written data is saved redundantly in the event of an unplanned transient error or outage. When an unplanned outage occurs on a host, vSAN 7 U2 and later version immediately write all incremental updates to another host in addition to the other host holding the active object replica. This helps ensure the durability of the changed data if an additional outage occurs on the other host holding the active object replica. This builds off the capability first introduced in vSAN 7 U1 that used this technique for planned maintenance events. These data durability improvements also have an additional benefit: Improving the time in which data is resynchronized to a stale object.



Data Services

Space Efficiency

Space efficiency features such 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.

Deduplication and Compression

Deduplication and Compression feature 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 deactivated by default and can be enabled using a simple drop-down menu. 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.

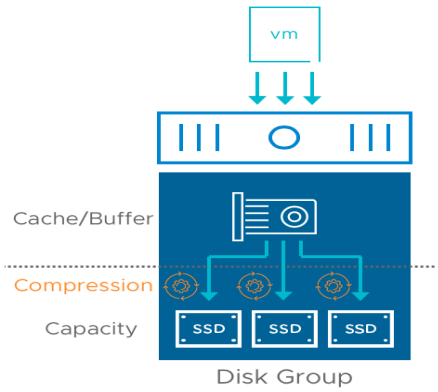
Note: A rolling format of all disks in the vSAN cluster is required when deduplication and compression are enabled on an existing cluster. This can take a considerable amount of time. However, this process does not incur virtual machine downtime.

Deduplication and compression are supported only with all flash vSAN configurations. The processes of deduplication and compression on any storage platform incur overhead, including vSAN. The extreme performance and low latency capabilities of flash devices easily outweigh the additional resource requirements of deduplication and compression. Furthermore, the space efficiency generated by deduplication and compression lowers the cost-per usable-GB of all flash.

Compression-only

Compression-only feature is functional subset of Deduplication & Compression that is intended to provide a compression-only capability. This allows customers to have greater flexibility in choosing the right amount of space efficiency and minimize the performance trade-off.

Compression-only feature is a cluster-wide setting. While enabled, the compression logic is implemented at capacity device level. This reduces the failure domain to the specific device as opposed to an entire diskgroup as in the case of Deduplication and Compression.

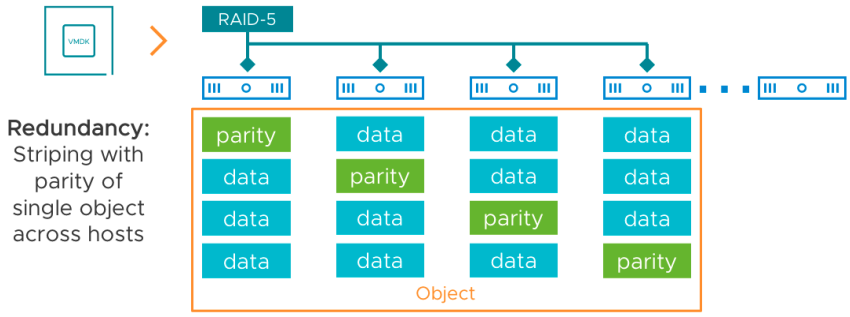


Erasure Coding

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 virtual disk object. 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.

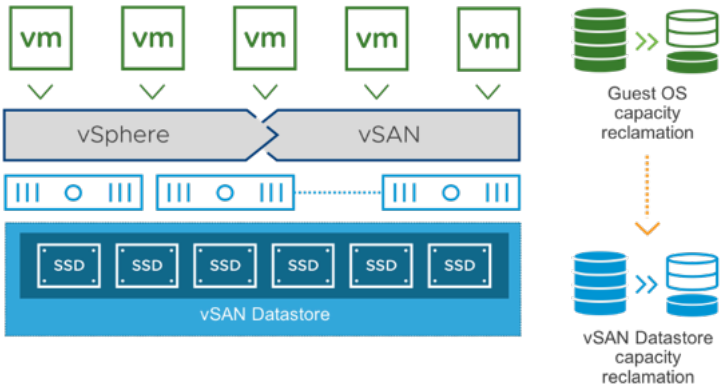


RAID-6 erasure coding requires a minimum of six hosts. For the same virtual disk object as in the previous example, a RAID-6 erasure coding rule creates four data components and two parity components. This configuration can survive the loss of two disks or hosts simultaneously (FTT=2).

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.

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.



When enabled vSAN 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 better storage capacity utilization in vSAN environments Administrators may find in some cases, dramatic space savings in their production vSAN environments.

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.

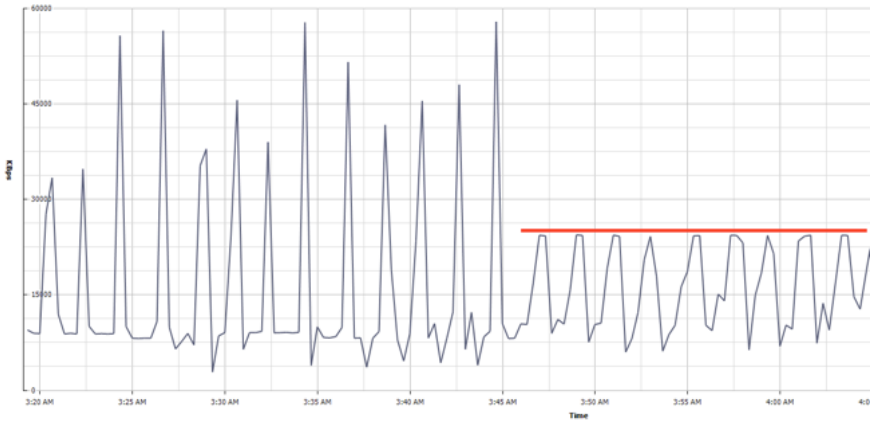
The last step is adding initiator names or an initiator group, which controls access to the target. It is possible to add individual names or create groups for ease of management.

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 possible to utilize vSAN iSCSI Target Service.

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.

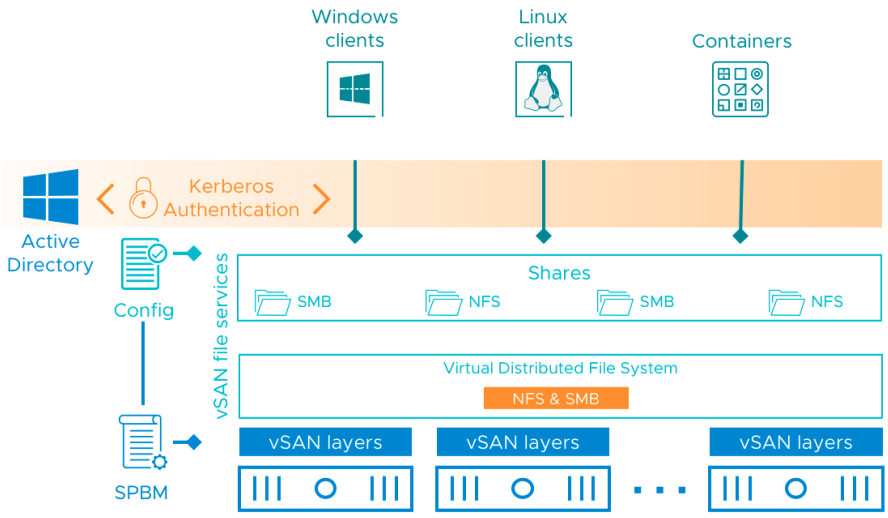


The IOPS limit eliminates possible performance impact to the other virtual machines. The reports take longer, but they are still finished in plenty of time for delivery the next day.

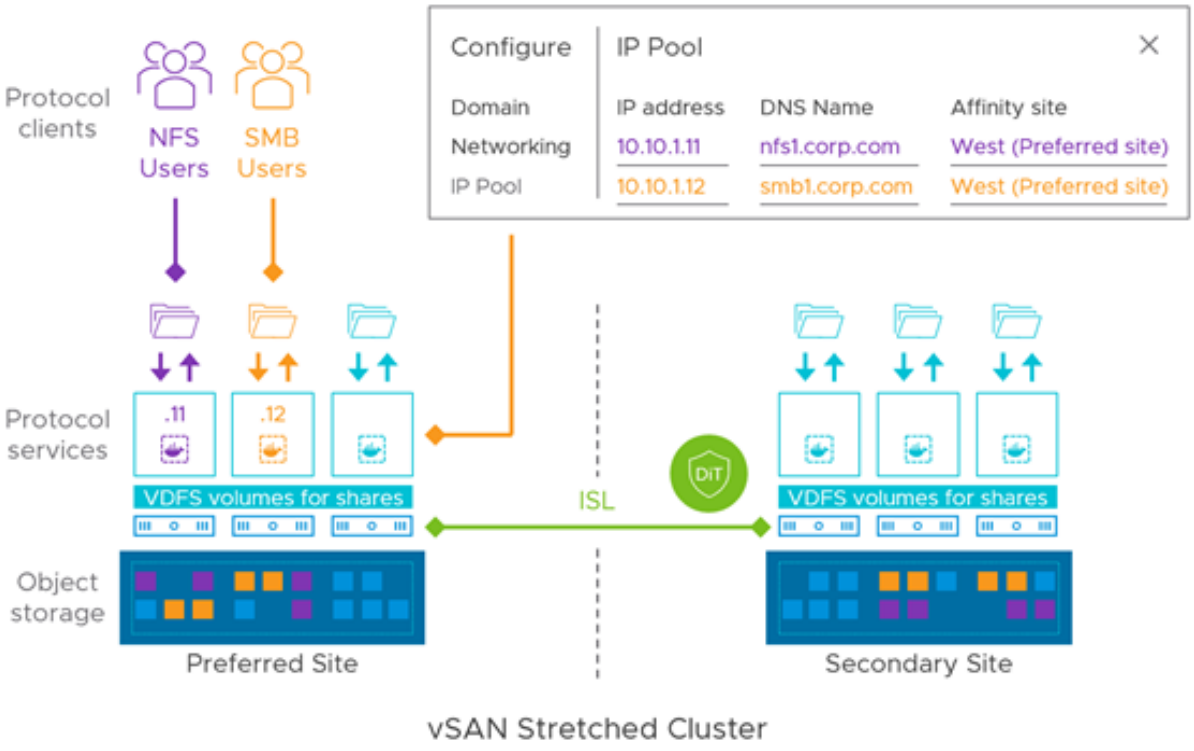
Keep in mind storage policies can be dynamically created, modified, and assigned to virtual machines. If an IOPS limit is proving to be too restrictive, simply modify the existing policy or create a new policy with a different IOPS limit and assign it to the virtual machines. The new or updated policy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

Native File Services

vSAN File Service provides file share capabilities natively and supports NFSv3, NFSv4.1, SMB v2.1 and SMB v3. Enabling file services in vSAN is similar to enabling other cluster-level features such as iSCSI services, encryption, deduplication, and compression. Once enabled, a set of containers on each of the hosts are deployed. These containers act as the primary delivery vehicle to provision file services. An abstraction layer comprising of Virtual Distributed File System (VDFS) provides the underlying scalable filesystem by aggregating vSAN objects to provide resilient file server endpoints and a control plane for deployment, management, and monitoring. File shares are integrated into the existing vSAN Storage Policy Based Management on a per-share basis. vSAN File Service supports Kerberos based authentication when using Microsoft Active Directory.



File services in vSAN 7 Update 2 can be used in vSAN stretched clusters as well as vSAN 2-Node topologies, which can make it ideal for those edge locations also in need of a file server. Data-in-Transit encryption, as well as the space reclamation technique known as UNMAP, are also supported. A snapshotting mechanism for point-in-time recovery of files is available through API. Finally, vSAN 7 U2 optimizes some of the metadata handling and data path for more efficient transactions, especially with small files.



Administrators can set a soft quota to generate alerts when a set threshold is exceeded and configure a hard quota on shares to restrict users from exceeding the allocated share size. The entire lifecycle of provisioning and managing file services can be seamlessly performed through the vCenter UI. This feature helps address a broader set of use cases requiring file services.

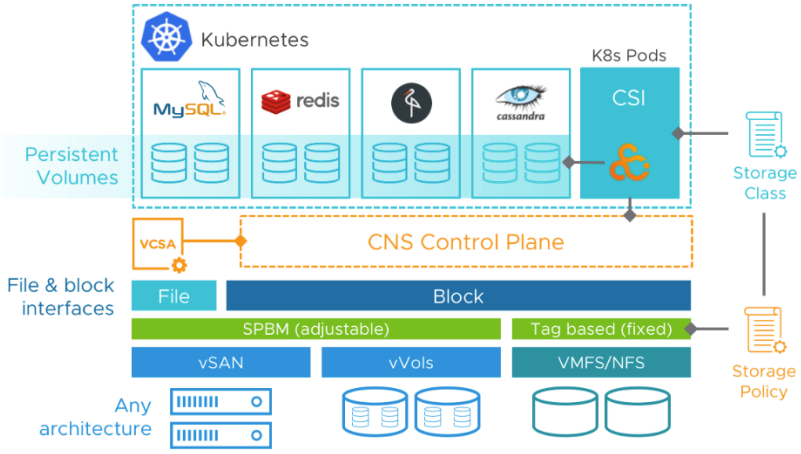
Cloud Native Storage

Cloud-native applications in the initial phases were mostly stateless and did not require persistent storage. In subsequent stages of evolution, several cloud-native applications benefited or required persistent storage. As more developers adopted newer

technologies such as containers and Kubernetes to develop, build, deploy, run and manage their applications, it became increasingly important that an equally agile Cloud-Native Storage be made available to these applications. vSAN enables a platform to seamlessly manage and provisioning persistent volumes to containerized workloads with the same ease and consistency as you would manage VMs. vSAN's architecture simplifies manageability, operations, and monitoring capabilities while providing enterprise-class storage performance and availability.

Persistent Volumes

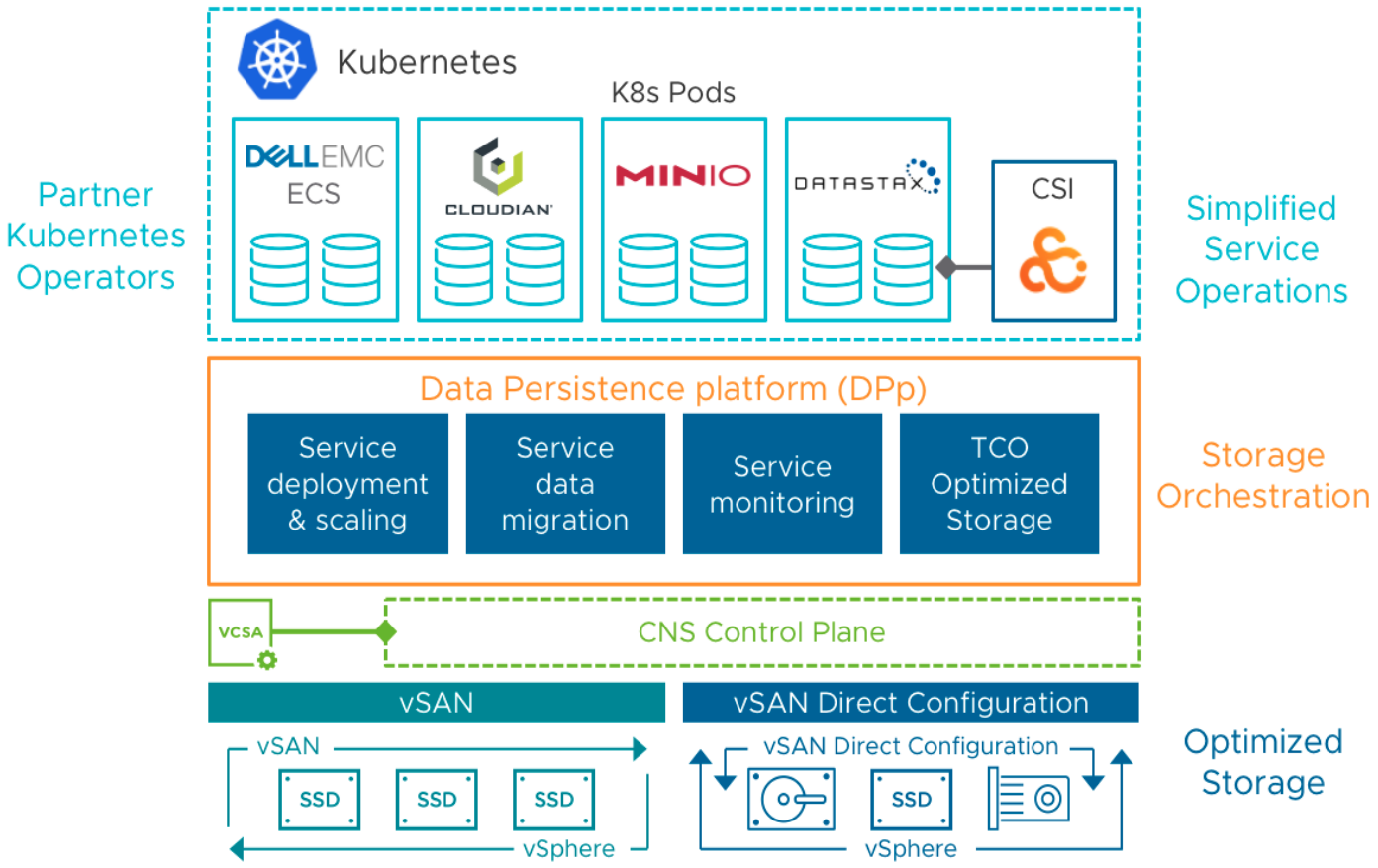
With VMware stack, Cloud Native Storage is natively built-in. vSAN provides the best-in-class enterprise storage platform for modern, cloud-native applications on Kubernetes, in vSphere. Administrators can dynamically provision policy-driven persistent volumes to Kubernetes based containers. vSAN supports both block and file-based persistent volumes.



Kubernetes uses a Storage Class to define different tiers of storage and to describe different types of requirements for storage backing the persistent volume. The requirements specified in storage class is mapped to an associated storage policy to provision persistent volumes with the requested attributes. The ability to manage persistent volumes through storage policies ensures operational consistency for the vSphere administrator. vCenter Server provides a unified interface to provision, manage, monitor, and troubleshoot persistent volumes.

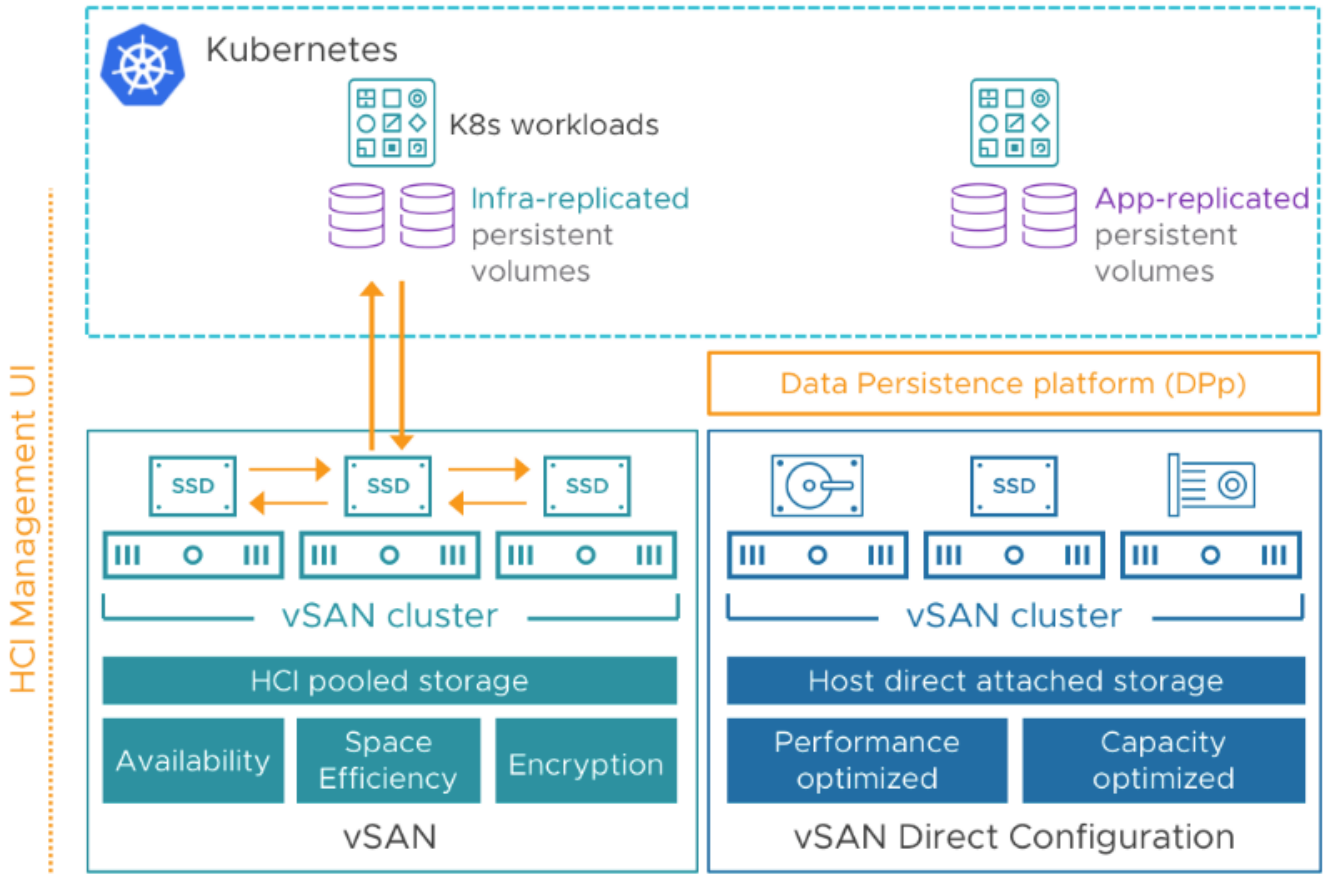
vSAN Data Persistence platform (DPp)

vSAN 7 Update 1 introduces a framework to simplify the deployment and operations of modern applications relying on cloud-native architecture. Cloud-native applications have unique infrastructure needs due to their shared-nothing architecture (SNA) and integrated data services. The vSAN Data Persistence platform (DPp) allows modern databases and stateful services to plug into this framework through APIs to maintain an efficient persistence of these distributed applications and data. The APIs allow partners to take advantage and build robust solutions and orchestrate the persistence of the application and data.



vSAN Direct

vSAN Direct Configuration provides performance optimized data storage for shared-nothing applications (SNA) like Cassandra and MongoDB that do not need data services. The applications are administered through the same vCenter management interface but use an optimized data path to access designated storage devices on the hosts in the vSAN cluster but not part of the vSAN datastore. This ability aids in reducing raw storage overhead for stateful cloud native applications. vSAN Direct can also be tailored to suit applications that prioritize high storage density over performance. The vSAN ReadyNode program will be updated to provide configurations to suit these value-based use cases. vSAN Direct is available as part of vSAN Data Persistence platform(DPp) through VMware VCF with Tanzu.

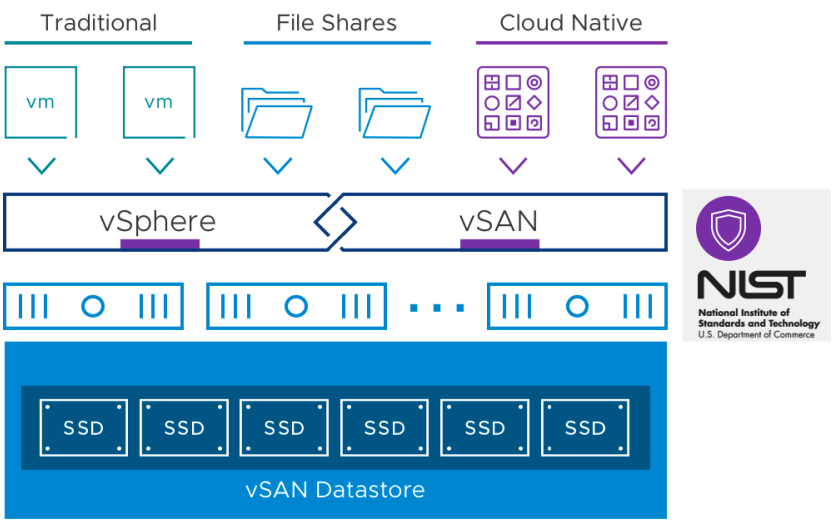


vSAN 7 U2 provides enhancements for Kubernetes-powered workloads. Migration from the legacy vSphere Cloud Provider (VCP) to the Container Storage Interface (CSI) is supported. The CSI driver enables Kubernetes to provision and manage persistent volumes when running on vSphere. Using the CSI helps administrators more effectively run, monitor, and manage container applications and their persistent volumes in their environment. Persistent volumes can be resized without the need to take it offline, eliminating any interruption.

Security

Native VMkernel Cryptographic Module

VMware incorporates 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.

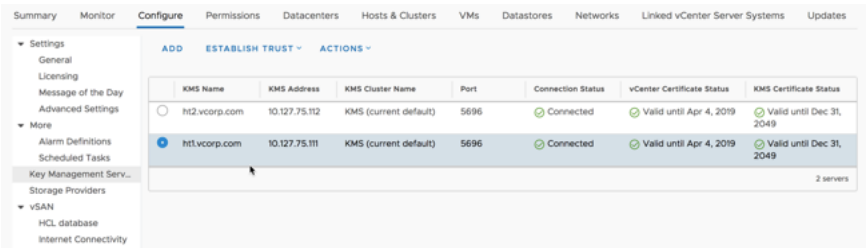
Because the VMware VMkernel Cryptographic Module is part of the ESXi kernel, it can easily provide FIPS 140-2 approved cryptographic services to various VMware products and services.

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.

Key Management

Key management is a core requirement for being able to use vSAN Data-at-rest 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.

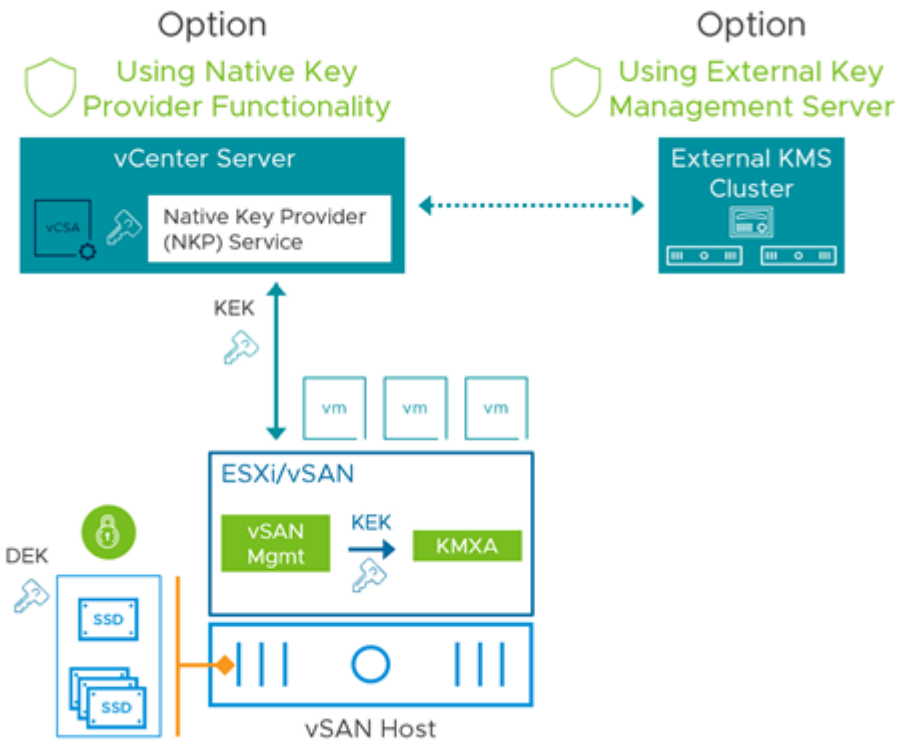


For a complete list of validated KMS vendors, their solutions, and links to their documentation, reference the VMware Hardware Compatibility List.

vSphere and vSAN Native Key Provider

"Native Key Provider" feature simplifies the key management for environments using encryption. For vSAN, the embedded KMS is ideal for Edge or 2-Node topologies and is a great example of VMware's approach to intrinsic security. The Native Key Provider is intended to help customers who were not already enabling encryption to do so in a secure and supported manner. This vSAN 7

Update 2 feature makes vSAN Data-at-rest protection – easy, flexible, and secure. Customers requiring key management for other infrastructure will need to implement a KMS. NKP is purpose-built to serve only vSphere.



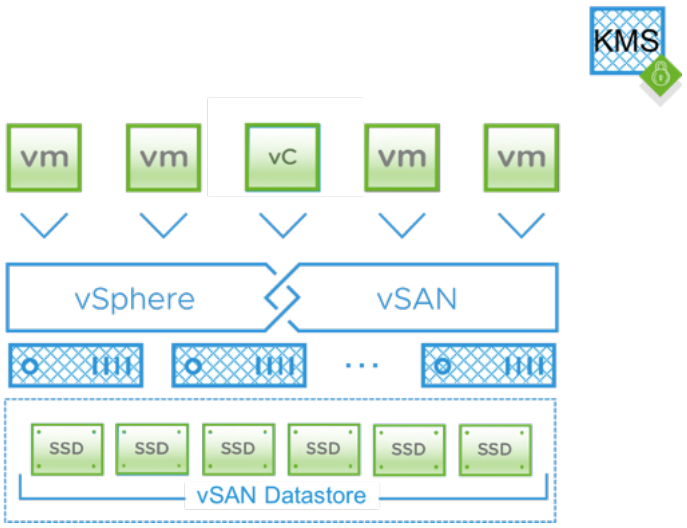
vSAN Encryption

vSAN delivers Data-At-Rest and Data-In-Transit encryption to ensure data security based on FIPS 140-2 validated Cryptographic modules. The services are rendered at a cluster-wide option. Both services can be enabled together or independently.

Data-At-Rest Encryption

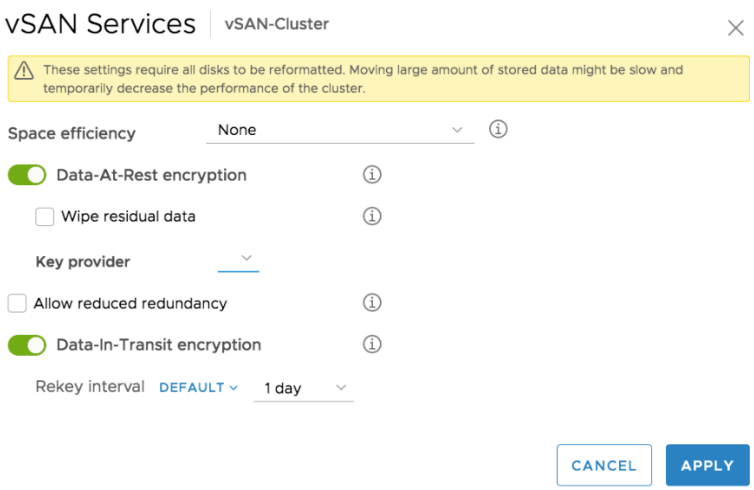
The VMware VMkernel Cryptographic Module is used by vSAN to provide Data at Rest Encryption to protect data on the disk groups and their devices in a vSAN cluster.

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.

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.

Data-In-Transit Encryption

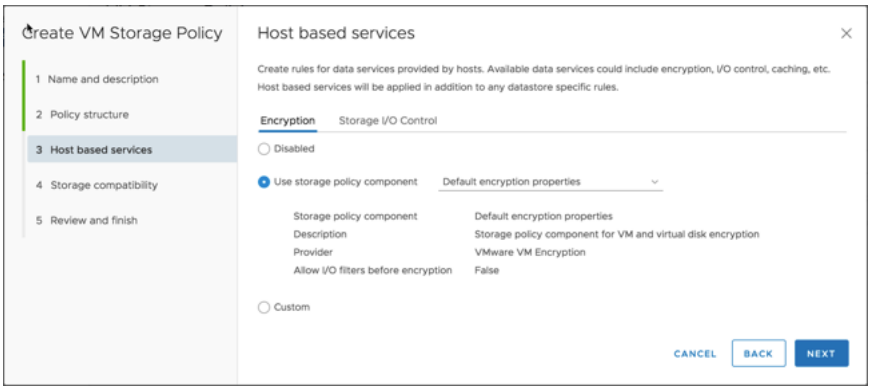
vSAN Data-in-Transit Encryption securely encrypts all vSAN traffic in transit across hosts using FIPS 140-2 validated Cryptographic modules. AES-GCM secure cipher is used for this encryption. This feature can be enabled independently from vSAN data-at-rest encryption and does not require a Key Management solutions (KMS). This helps address regulatory standards and compliance that is prevalent across specific industry segments and deployments involving multi-tenancy for added security.

A rekey interval for all hosts in the cluster can be specified at the cluster level in the configuration for Data-in-Transit Encryption. Health checks ensure verify that all hosts in cluster have a consistent encryption configuration, and rekey interval and alert the administrator if there is an anomaly.

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.

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.

Secure Disk wipe

vSAN facilitates security throughout the entire lifecycle, from provisioning to decommissioning a host or a drive. Secure Disk wipe capability completes this cycle by allowing administrators to securely erase data prior to a hosts or drive by removed or repurposed. This feature is based on NIST standards and rendered through PowerCLI or API.

Compliance

vSAN is in compliance with Voluntary Product Accessibility Template (VPAT), that defines conformance of information and communication technology with United State Rehabilitation Act. Section 508 of VPAT provides specific guidelines to promote standards of accessibility for people with limitations in vision, hearing, or other physical disabilities. vSAN compliance with VPAT ensures that the product is eligible for use in organizations that require all purchases to adhere to these regulatory requirements.

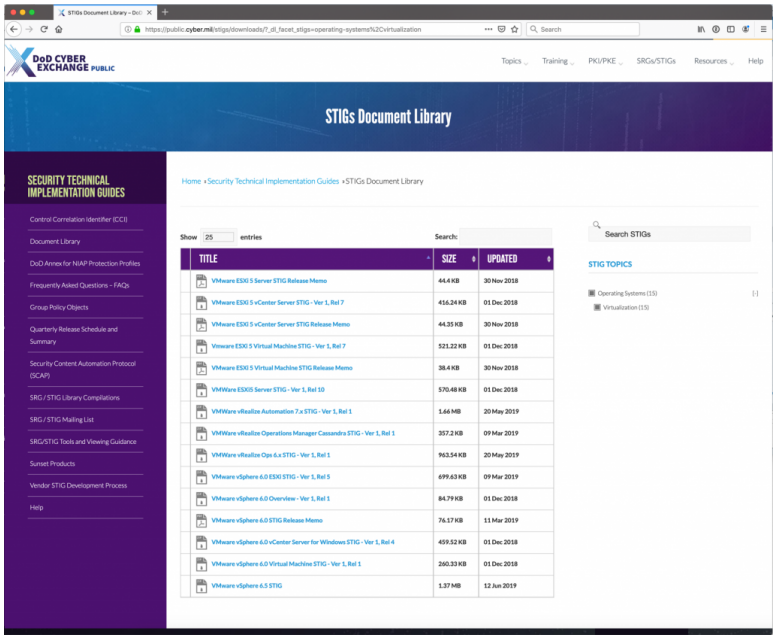
vSAN is native to the vSphere hypervisor and, because of that tight integration, shares the robust security and compliance benefits realized by the vSphere platform. 2-factor authentication methods, such as RSA SecurID and Common Access Card (CAC), are supported by vCenter Server, vSphere, and vSAN.

In April of 2017, VMware announced that VMware vSAN has been added to the VMware vSphere 6.0 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.

Again, in June of 2019, VMware announced that VMware vSAN has been added to the VMware vSphere 6.5 STIG Framework. At the time of this writing, VMware HCI is still the only HCI Solution that is part of a DISA STIG.

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 DoD Cyber Exchange 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 DoD Cyber Exchange website.

Summary

Hyper-Converged 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.

References

Additional Documentation

For more information about VMware vSAN, please visit the product pages at http://www.vmware.com/products/vsan.html

Below are some links to online documentation:

Cloud Platform on TechZone:
https://core.vmware.com

Virtual Blocks:
http://virtualblocks.com/

VMware vSAN Community: https://communities.vmware.com/community/vmtn/vsan

VMware PowerCLI:
https://code.vmware.com/web/dp/tool/vmware-powercli/

VMware API Explorer:
https://code.vmware.com/apis

Support Knowledge base:
https://kb.vmware.com/

VMware Contact Information

For additional information or to purchase VMware Virtual SAN, VMware's global network of solutions providers is ready to assist. If you would like to contact VMware directly, you can reach a sales representative at 1-877-4VMWARE (650-475-5000 outside North America) or email sales@vmware.com. When emailing, please include the state, country, and company name from which you are inquiring.

VMware, Inc. 3401 Hillview Avenue Palo Alto CA 94304 USA Tel 877-486-9273 Fax 650-427-5001 www.vmware.com