vSAN Planning and Deployment

Update 3

VMware vSphere 7.0
VMware vSAN 7.0

vmware®

You can find the most up-to-date technical documentation on the VMware website at:

https://docs.vmware.com/

VMware, Inc.
3401 Hillview Ave.
Palo Alto, CA 94304
www.vmware.com

Contents

About vSAN Planning and Deployment

vSAN Planning and Deployment describes how to design and deploy a vSAN cluster in a vSphere environment. The information includes system requirements, sizing guidelines, and suggested best practices.

At VMware, we value inclusion. To foster this principle within our customer, partner, and internal community, we create content using inclusive language.

Intended Audience

This manual is intended for anyone who wants to design and deploy a vSAN cluster in a VMware vSphere environment. The information in this manual is written for experienced system administrators who are familiar with virtual machine technology and virtual datacenter operations. This manual assumes familiarity with VMware vSphere, including VMware ESXi, vCenter Server, and the vSphere Client.

For more information about vSAN features and how to configure a vSAN cluster, see *Administering VMware vSAN*.

For more information about monitoring a vSAN cluster and fixing problems, see the *vSAN Monitoring and Troubleshooting Guide*.

Introduction to vSAN

1

VMware vSAN is a distributed layer of software that runs natively as a part of the ESXi hypervisor. vSAN aggregates local or direct-attached capacity devices of a host cluster and creates a single storage pool shared across all hosts in the vSAN cluster.

While supporting VMware features that require shared storage, such as HA, vMotion, and DRS, vSAN eliminates the need for external shared storage and simplifies storage configuration and virtual machine provisioning activities.

This chapter includes the following topics:

vSAN Concepts

VMware vSAN uses a software-defined approach that creates shared storage for virtual machines. It virtualizes the local physical storage resources of ESXi hosts and turns them into pools of storage that can be divided and assigned to virtual machines and applications according to their quality-of-service requirements. vSAN is implemented directly in the ESXi hypervisor.

You can configure vSAN to work as either a hybrid or all-flash cluster. In hybrid clusters, flash devices are used for the cache layer and magnetic disks are used for the storage capacity layer. In all-flash clusters, flash devices are used for both cache and capacity.

You can activate vSAN on your existing host clusters and when you create new clusters. vSAN aggregates all local capacity devices into a single datastore shared by all hosts in the vSAN cluster. You can expand the datastore by adding capacity devices or hosts with capacity devices to the cluster. vSAN works best when all ESXi hosts in the cluster share similar or identical

configurations across all cluster members, including similar or identical storage configurations. This consistent configuration balances virtual machine storage components across all devices and hosts in the cluster. Hosts without any local devices also can participate and run their virtual machines on the vSAN datastore.

If a host contributes its local storage devices to the vSAN datastore, it must provide at least one device for flash cache and at least one device for capacity. Capacity devices are also called data disks.

The devices on the contributing host form one or more disk groups. Each disk group contains one flash cache device, and one or multiple capacity devices for persistent storage. Each host can be configured to use multiple disk groups.

For best practices, capacity considerations, and general recommendations about designing and sizing a vSAN cluster, see the *VMware vSAN Design and Sizing Guide*.

Characteristics of vSAN

This topic summarizes characteristics that apply to vSAN, its clusters, and datastores.

vSAN provides numerous benefits to your environment.

Table 1-1. vSAN Features

Supported Features	Description
Shared storage support	vSAN supports VMware features that require shared storage, such as HA, vMotion, and DRS. For example, if a host becomes overloaded, DRS can migrate virtual machines to other hosts in the cluster.
On-disk format	vSAN on-disk virtual file format provides highly scalable snapshot and clone management support per vSAN cluster. For information about the number of virtual machine snapshots and clones supported per vSAN cluster, see the *Configuration Maximums* documentation.
All-flash and hybrid configurations	vSAN can be configured for all-flash or hybrid cluster.
Fault domains	vSAN supports configuring fault domains to protect hosts from rack or chassis failures when the vSAN cluster spans across multiple racks or blade server chassis in a data center.
iSCSI target service	vSAN iSCSI target service enables hosts and physical workloads that reside outside the vSAN cluster to access the vSAN datastore.
Stretched cluster	vSAN supports stretched clusters that span across two geographic locations.

Table 1-1. vSAN Features (continued)

Supported Features	Description
Support for Windows Server Failover Clusters (WSFC)	vSAN 6.7 Update 3 and later releases support SCSI-3 Persistent Reservations (SCSI3-PR) on a virtual disk level required by Windows Server Failover Cluster (WSFC) to arbitrate an access to a shared disk between nodes. Support of SCSI-3 PRs enables configuration of WSFC with a disk resource shared between VMs natively on vSAN datastores. Currently the following configurations are supported: ■ Up to 6 application nodes per cluster. ■ Up to 64 shared virtual disks per node. **Note** Microsoft SQL Server 2012 or later running on Microsoft Windows Server 2012 or later has been qualified on vSAN.
vSAN health service	vSAN health service includes preconfigured health check tests to monitor, troubleshoot, diagnose the cause of cluster component problems, and identify any potential risk.
vSAN performance service	vSAN performance service includes statistical charts used to monitor IOPS, throughput, latency, and congestion. You can monitor performance of a vSAN cluster, host, disk group, disk, and VMs.
Integration with vSphere storage features	vSAN integrates with vSphere data management features traditionally used with VMFS and NFS storage. These features include snapshots, linked clones, and vSphere Replication.
Virtual Machine Storage Policies	vSAN works with VM storage policies to support a VM-centric approach to storage management. If you do not assign a storage policy to the virtual machine during deployment, the vSAN Default Storage Policy is automatically assigned to the VM.
Rapid provisioning	vSAN enables rapid provisioning of storage in the vCenter Server® during virtual machine creation and deployment operations.
Deduplication and compression	vSAN performs block-level deduplication and compression to save storage space. When you enable deduplication and compression on a vSAN all-flash cluster, redundant data within each disk group is reduced. Deduplication and compression is a cluster-wide setting, but the functions are applied on a disk group basis. Compression-only vSAN is applied on a per-disk basis.
Data at rest encryption	vSAN provides data at rest encryption. Data is encrypted after all other processing, such as deduplication, is performed. Data at rest encryption protects data on storage devices, in case a device is removed from the cluster.
SDK support	The VMware vSAN SDK for Java is an extension of the VMware vSphere Management SDK. It includes documentation, libraries and code examples that help developers automate installation, configuration, monitoring, and troubleshooting of vSAN.

vSAN Terms and Definitions

vSAN introduces specific terms and definitions that are important to understand.

Before you get started with vSAN, review the key vSAN terms and definitions.

Disk Group

A disk group is a unit of physical storage capacity on a host and a group of physical devices that provide performance and capacity to the vSAN cluster. On each ESXi host that contributes its local devices to a vSAN cluster, devices are organized into disk groups.

Each disk group must have one flash cache device and one or multiple capacity devices. The devices used for caching cannot be shared across disk groups, and cannot be used for other purposes. A single caching device must be dedicated to a single disk group. In hybrid clusters, flash devices are used for the cache layer and magnetic disks are used for the storage capacity layer. In an all-flash cluster, flash devices are used for both cache and capacity. For information about creating and managing disk groups, see *Administering VMware vSAN*.

Consumed Capacity

Consumed capacity is the amount of physical capacity consumed by one or more virtual machines at any point. Many factors determine consumed capacity, including the consumed size of your VMDKs, protection replicas, and so on. When calculating for cache sizing, do not consider the capacity used for protection replicas.

Object-Based Storage

vSAN stores and manages data in the form of flexible data containers called objects. An object is a logical volume that has its data and metadata distributed across the cluster. For example, every VMDK is an object, as is every snapshot. When you provision a virtual machine on a vSAN datastore, vSAN creates a set of objects comprised of multiple components for each virtual disk. It also creates the VM home namespace, which is a container object that stores all metadata files of your virtual machine. Based on the assigned virtual machine storage policy, vSAN provisions and manages each object individually, which might also involve creating a RAID configuration for every object.

When vSAN creates an object for a virtual disk and determines how to distribute the object in the cluster, it considers the following factors:

- vSAN verifies that the virtual disk requirements are applied according to the specified virtual machine storage policy settings.
- vSAN verifies that the correct cluster resources are used at the time of provisioning. For example, based on the protection policy, vSAN determines how many replicas to create. The performance policy determines the amount of flash read cache allocated for each replica and how many stripes to create for each replica and where to place them in the cluster.

- vSAN continually monitors and reports the policy compliance status of the virtual disk. If you find any noncompliant policy status, you must troubleshoot and resolve the underlying problem.

  Note When required, you can edit VM storage policy settings. Changing the storage policy settings does not affect virtual machine access. vSAN actively throttles the storage and network resources used for reconfiguration to minimize the impact of object reconfiguration to normal workloads. When you change VM storage policy settings, vSAN might initiate an object recreation process and subsequent resynchronization. See *vSAN Monitoring and Troubleshooting*.

- vSAN verifies that the required protection components, such as mirrors and witnesses, are placed on separate hosts or fault domains. For example, to rebuild components during a failure, vSAN looks for ESXi hosts that satisfy the placement rules where protection components of virtual machine objects must be placed on two different hosts, or across fault domains.

vSAN Datastore

After you enable vSAN on a cluster, a single vSAN datastore is created. It appears as another type of datastore in the list of datastores that might be available, including Virtual Volume, VMFS, and NFS. A single vSAN datastore can provide different service levels for each virtual machine or each virtual disk. In vCenter Server®, storage characteristics of the vSAN datastore appear as a set of capabilities. You can reference these capabilities when defining a storage policy for virtual machines. When you later deploy virtual machines, vSAN uses this policy to place virtual machines in the optimal manner based on the requirements of each virtual machine. For general information about using storage policies, see the *vSphere Storage* documentation.

A vSAN datastore has specific characteristics to consider.

- vSAN provides a single vSAN datastore accessible to all hosts in the cluster, whether or not they contribute storage to the cluster. Each host can also mount any other datastores, including Virtual Volumes, VMFS, or NFS.
- You can use Storage vMotion to move virtual machines between vSAN datastores, NFS datastores, and VMFS datastores.
- Only magnetic disks and flash devices used for capacity can contribute to the datastore capacity. The devices used for flash cache are not counted as part of the datastore.

Objects and Components

Each object is composed of a set of components, determined by capabilities that are in use in the VM Storage Policy. For example, with **Primary level of failures to tolerate** set to 1, vSAN ensures that the protection components, such as replicas and witnesses, are placed on separate hosts in the vSAN cluster, where each replica is an object component. In addition, in the same policy, if the **Number of disk stripes per object** configured to two or more, vSAN also stripes the object across multiple capacity devices and each stripe is considered a component of the specified object. When needed, vSAN might also break large objects into multiple components.

A vSAN datastore contains the following object types:

VM Home Namespace

The virtual machine home directory where all virtual machine configuration files are stored, such as `.vmx`, log files, vmdks, and snapshot delta description files.

VMDK

A virtual machine disk or `.vmdk` file that stores the contents of the virtual machine's hard disk drive.

VM Swap Object

Created when a virtual machine is powered on.

Snapshot Delta VMDKs

Created when virtual machine snapshots are taken.

Memory object

Created when the snapshot memory option is selected when creating or suspending a virtual machine.

Virtual Machine Compliance Status: Compliant and Noncompliant

A virtual machine is considered noncompliant when one or more of its objects fail to meet the requirements of its assigned storage policy. For example, the status might become noncompliant when one of the mirror copies is inaccessible. If your virtual machines are in compliance with the requirements defined in the storage policy, the status of your virtual machines is compliant. From the **Physical Disk Placement** tab on the **Virtual Disks** page, you can verify the virtual machine object compliance status. For information about troubleshooting a vSAN cluster, see *vSAN Monitoring and Troubleshooting*.

Component State: Degraded and Absent States

vSAN acknowledges the following failure states for components:

- Degraded. A component is Degraded when vSAN detects a permanent component failure and determines that the failed component cannot recover to its original working state. As a result, vSAN starts to rebuild the degraded components immediately. This state might occur when a component is on a failed device.
- Absent. A component is Absent when vSAN detects a temporary component failure where components, including all its data, might recover and return vSAN to its original state. This state might occur when you are restarting hosts or if you unplug a device from a vSAN host. vSAN starts to rebuild the components in absent status after waiting for 60 minutes.

Object State: Healthy and Unhealthy

Depending on the type and number of failures in the cluster, an object might be in one of the following states:

- Healthy. When at least one full RAID 1 mirror is available, or the minimum required number of data segments are available, the object is considered healthy.
- Unhealthy. An object is considered unhealthy when no full mirror is available or the minimum required number of data segments are unavailable for RAID 5 or RAID 6 objects. If fewer than 50 percent of an object's votes are available, the object is unhealthy. Multiple failures in the cluster can cause objects to become unhealthy. When the operational status of an object is considered unhealthy, it impacts the availability of the associated VM.

Witness

A witness is a component that contains only metadata and does not contain any actual application data. It serves as a tiebreaker when a decision must be made regarding the availability of the surviving datastore components, after a potential failure. A witness consumes approximately 2 MB of space for metadata on the vSAN datastore when using on-disk format 1.0, and 4 MB for on-disk format for version 2.0 and later.

vSAN 6.0 and later maintains a quorum by using an asymmetrical voting system where each component might have more than one vote to decide the availability of objects. Greater than 50 percent of the votes that make up a VM’s storage object must be accessible at all times for the object to be considered available. When 50 percent or fewer votes are accessible to all hosts, the object is no longer accessible to the vSAN datastore. Inaccessible objects can impact the availability of the associated VM.

Storage Policy-Based Management (SPBM)

When you use vSAN, you can define virtual machine storage requirements, such as performance and availability, in the form of a policy. vSAN ensures that the virtual machines deployed to vSAN datastores are assigned at least one virtual machine storage policy. When you know the storage requirements of your virtual machines, you can define storage policies and assign the policies

to your virtual machines. If you do not apply a storage policy when deploying virtual machines, vSAN automatically assigns a default vSAN policy with **Primary level of failures to tolerate** set to 1, a single disk stripe for each object, and thin provisioned virtual disk. For best results, define your own virtual machine storage policies, even if the requirements of your policies are the same as those defined in the default storage policy. For information about working with vSAN storage policies, see *Administering VMware vSAN*.

vSphere PowerCLI

VMware vSphere PowerCLI adds command-line scripting support for vSAN, to help you automate configuration and management tasks. vSphere PowerCLI provides a Windows PowerShell interface to the vSphere API. PowerCLI includes cmdlets for administering vSAN components. For information about using vSphere PowerCLI, see *vSphere PowerCLI Documentation*.

vSAN and Traditional Storage

Although vSAN shares many characteristics with traditional storage arrays, the overall behavior and function of vSAN is different. For example, vSAN can manage and work only with ESXi hosts and a single vSAN instance can support only one cluster.

vSAN and traditional storage also differ in the following key ways:

- vSAN does not require external networked storage for storing virtual machine files remotely, such as on a Fibre Channel (FC) or Storage Area Network (SAN).
- Using traditional storage, the storage administrator preallocates storage space on different storage systems. vSAN automatically turns the local physical storage resources of the ESXi hosts into a single pool of storage. These pools can be divided and assigned to virtual machines and applications according to their quality-of-service requirements.
- vSAN does not behave like traditional storage volumes based on LUNs or NFS shares. The iSCSI target service uses LUNs to enable an initiator on a remote host to transport block-level data to a storage device in the vSAN cluster.
- Some standard storage protocols, such as FCP, do not apply to vSAN.
- vSAN is highly integrated with vSphere. You do not need dedicated plug-ins or a storage console for vSAN, compared to traditional storage. You can deploy, manage, and monitor vSAN by using the vSphere Client.
- A dedicated storage administrator does not need to manage vSAN. Instead a vSphere administrator can manage a vSAN environment.
- With vSAN, VM storage policies are automatically assigned when you deploy new VMs. The storage policies can be changed dynamically as needed.

Building a vSAN Cluster

If you are considering vSAN, you can choose from more than one configuration solution for deploying a vSAN cluster.

Depending on your requirement, you can deploy vSAN in one of the following ways.

vSAN ReadyNode

The vSAN ReadyNode is a preconfigured solution of the vSAN software provided by VMware partners, such as Cisco, Dell, Fujitsu, IBM, and Supermicro. This solution includes validated server configuration in a tested, certified hardware form factor for vSAN deployment that is recommended by the server OEM and VMware. For information about the vSAN ReadyNode solution for a specific partner, visit the VMware Partner website.

User-Defined vSAN Cluster

You can build a vSAN cluster by selecting individual software and hardware components, such as drivers, firmware, and storage I/O controllers that are listed in the vSAN Compatibility Guide (VCG) website at http://www.vmware.com/resources/compatibility/search.php. You can choose any servers, storage I/O controllers, capacity and flash cache devices, memory, any number of cores you must have per CPU, that are certified and listed on the VCG website. Review the compatibility information on the VCG website before choosing software and hardware components, drivers, firmware, and storage I/O controllers that vSAN supports. When designing a vSAN cluster, use only devices, firmware, and drivers that are listed on the VCG website. Using software and hardware versions that are not listed in the VCG might cause cluster failure or unexpected data loss. For information about designing a vSAN cluster, see Chapter 3 Designing and Sizing a vSAN Cluster.

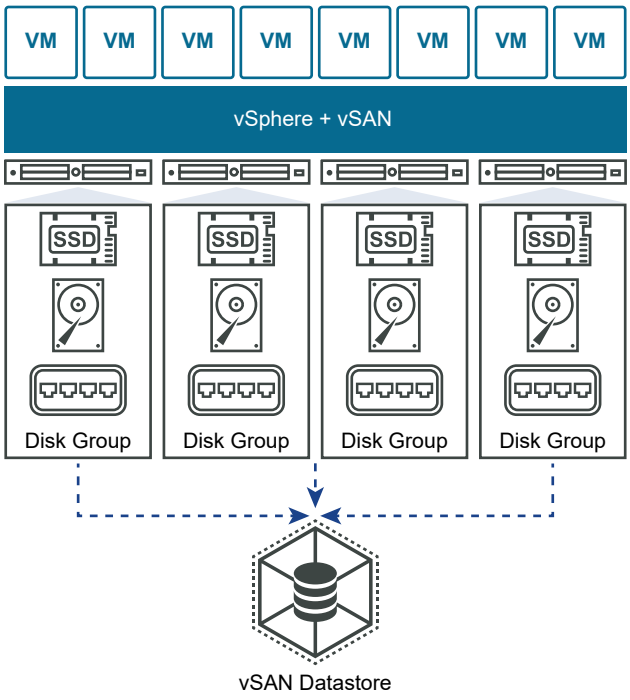
vSAN Deployment Options

This section covers the different supported deployment options that are supported for vSAN clusters.

Standard vSAN Cluster

A standard vSAN cluster consists of a minimum of three hosts. Typically, all hosts in a standard vSAN cluster reside at the same location, and are connected on the same Layer 2 network. All-flash configurations require 10 Gb network connections, and this also is recommended for hybrid configurations.

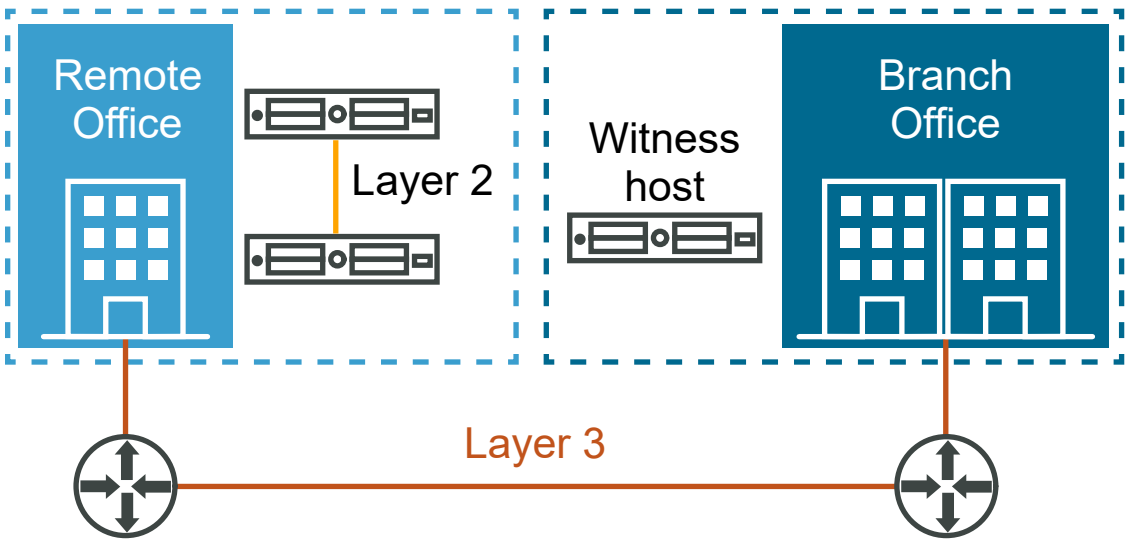
For more information, see Chapter 5 Creating a vSAN Cluster.



Two-Node vSAN Cluster

Two-node vSAN clusters are often used for remote office/branch office environments, typically running a small number of workloads that require high availability. A two-node vSAN cluster consists of two hosts at the same location, connected to the same network switch or directly connected. You can configure a two-node vSAN cluster that uses a third host as a witness, which can be located remotely from the branch office. Usually the witness resides at the main site, along with the vCenter Server.

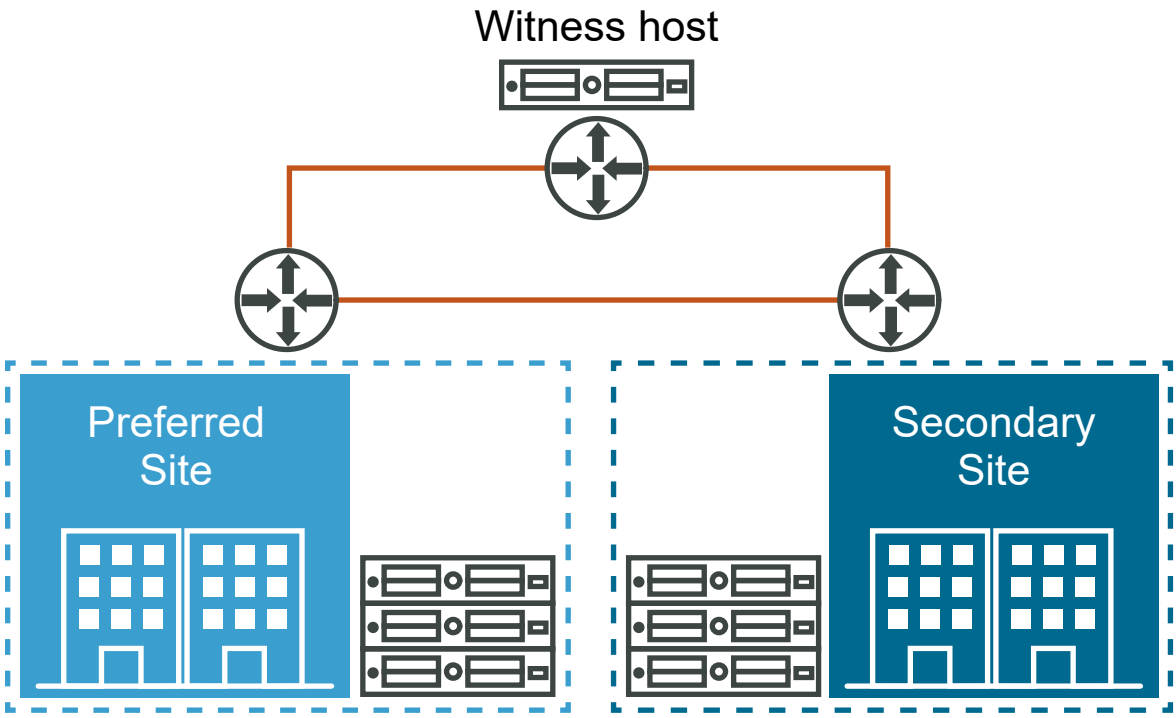
For more information, see Introduction to Stretched Clusters .



vSAN 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 must have a network latency of no more than five milliseconds (5 ms). A vSAN witness host resides at a third site to provide the witness function. The witness also acts as tie-breaker in scenarios where a network partition occurs between the two data sites. Only metadata such as witness components is stored on the witness.

For more information, see Introduction to Stretched Clusters .



Integrating with Other VMware Software

After you have vSAN up and running, it is integrated with the rest of the VMware software stack. You can do most of what you can do with traditional storage by using vSphere components and features including vSphere vMotion, snapshots, clones, Distributed Resource Scheduler (DRS), vSphere High Availability, vCenter Site Recovery Manager, and more.

Integrating with vSphere HA

You can enable vSphere HA and vSAN on the same cluster. As with traditional datastores, vSphere HA provides the same level of protection for virtual machines on vSAN datastores. This level of protection imposes specific restrictions when vSphere HA and vSAN interact. For specific considerations about integrating vSphere HA and vSAN, see Using vSAN and vSphere HA.

Integrating with VMware Horizon View

You can integrate vSAN with VMware Horizon View. When integrated, vSAN provides the following benefits to virtual desktop environments:

- High-performance storage with automatic caching
- Storage policy-based management, for automatic remediation

For information about integrating vSAN with VMware Horizon, see the *VMware Horizon with View* documentation. For designing and sizing VMware Horizon View for vSAN, see the *Designing and Sizing Guide for Horizon View*.

Limitations of vSAN

This topic discusses the limitations of vSAN.

When working with vSAN, consider the following limitations:

- vSAN does not support hosts participating in multiple vSAN clusters. However, a vSAN host can access other external storage resources that are shared across clusters.
- vSAN does not support vSphere DPM and Storage I/O Control.
- vSAN does not support SE Sparse disks.
- vSAN does not support RDM, VMFS, diagnostic partition, and other device access features.

Requirements for Enabling vSAN

2

Before you activate vSAN, verify that your environment meets all requirements.

This chapter includes the following topics:

Hardware Requirements for vSAN

Verify that the ESXi hosts in your organization meet the vSAN hardware requirements.

Storage Device Requirements

All capacity devices, drivers, and firmware versions in your vSAN configuration must be certified and listed in the vSAN section of the *VMware Compatibility Guide*.

Table 2-1. Storage Device Requirements for vSAN Hosts

Storage Component	Requirements
Cache	■ One SAS or SATA solid-state disk (SSD) or PCIe flash device. ■ Before calculating the **Primary level of failures to tolerate**, check the size of the flash caching device in each disk group. For hybrid cluster, it must provide at least 10 percent of the anticipated storage consumed on the capacity devices, not including replicas such as mirrors. For guidance on determining the cache ratio for all-flash clusters, refer to Designing vSAN Disk groups – All Flash Cache Ratio Update. ■ vSphere Flash Read Cache must not use any of the flash devices reserved for vSAN cache. ■ The cache flash devices must not be formatted with VMFS or another file system.
Virtual machine data storage	■ For hybrid group configuration, make sure that at least one SAS or NL-SAS magnetic disk is available. ■ For all-flash disk group configuration, make sure at least one SAS, or SATA solid-state disk (SSD), or PCIe flash device.
Storage controllers	One SAS or SATA host bus adapter (HBA), or a RAID controller that is in passthrough mode or RAID 0 mode. To avoid issues, consider these points when the same storage controller is backing both vSAN and non-vSAN disks: Do not mix the controller mode for vSAN and non-vSAN disks to avoid handling the disks inconsistently, which can negatively impact vSAN operation. If the vSAN disks are in RAID mode, the non-vSAN disks must also be in RAID mode. When you use non-vSAN disks for VMFS, use the VMFS datastore only for scratch, logging, and core dumps. Do not run virtual machines from a disk or RAID group that shares its controller with vSAN disks or RAID groups. Do not passthrough non-vSAN disks to virtual machine guests as Raw Device Mappings (RDMs). For more information, see https://kb.vmware.com/s/article/2129050. To learn about controller supported features, such as passthrough and RAID, refer to the vSAN HCL: https://www.vmware.com/resources/compatibility/search.php?deviceCategory=vsan

Memory

The memory requirements for vSAN depend on the number of disk groups and devices that the ESXi hypervisor must manage. For more information, see the VMware knowledge base article at https://kb.vmware.com/s/article/2113954.

Flash Boot Devices

During installation, the ESXi installer creates a coredump partition on the boot device. The default size of the coredump partition satisfies most installation requirements.

- If the memory of the ESXi host has 512 GB of memory or less, you can boot the host from a USB, SD, or SATADOM device. When you boot a vSAN host from a USB device or SD card, the size of the boot device must be at least 4 GB.
- If the memory of the ESXi host has more than 512 GB, consider the following guidelines.
  - You can boot the host from a SATADOM or disk device with a size of at least 16 GB. When you use a SATADOM device, use a single-level cell (SLC) device.
  - If you are using vSAN 6.5 or later, you must resize the coredump partition on ESXi hosts to boot from USB/SD devices. For more information, see the VMware knowledge base article at http://kb.vmware.com/kb/2147881.

When you boot an ESXi 6.0 or later host from USB device or from SD card, vSAN trace logs are written to RAMDisk. These logs are automatically offloaded to persistent media during shutdown or system crash (panic). This is the only support method for handling vSAN traces when booting an ESXi from a USB stick or SD card. If a power failure occurs, vSAN trace logs are not preserved.

When you boot an ESXi 6.0 or later host from a SATADOM device, vSAN trace logs are written directly to the SATADOM device. Therefore it is important that the SATADOM device meets the specifications outlined in this guide.

Cluster Requirements for vSAN

Verify that a host cluster meets the requirements for enabling vSAN.

- All capacity devices, drivers, and firmware versions in your vSAN configuration must be certified and listed in the vSAN section of the *VMware Compatibility Guide*.
- A standard vSAN cluster must contain a minimum of three hosts that contribute capacity to the cluster. A two host vSAN cluster consists of two data hosts and an external witness host. For information about the considerations for a three-host cluster, see Design Considerations for a vSAN Cluster.
- A host that resides in a vSAN cluster must not participate in other clusters.

Software Requirements for vSAN

Verify that the vSphere components in your environment meet the software version requirements for using vSAN.

To use the full set of vSAN capabilities, the ESXi hosts that participate in vSAN clusters must be version 7.0 Update 1 or later. During the vSAN upgrade from previous versions, you can keep the current on-disk format version, but you cannot use many of the new features. vSAN 7.0 Update 1 and later software supports all on-disk formats.

Networking Requirements for vSAN

Verify that the network infrastructure and the networking configuration on the ESXi hosts meet the minimum networking requirements for vSAN.

Table 2-2. Networking Requirements for vSAN

Networking Component	Requirement
Host Bandwidth	Each host must have minimum bandwidth dedicated to vSAN. ■ Dedicated 1 Gbps for hybrid configurations ■ Dedicated or shared 10 Gbps for all-flash configurations For information about networking considerations in vSAN, see Designing the vSAN Network.
Connection between hosts	Each host in the vSAN cluster, regardless of whether it contributes capacity, must have a VMkernel network adapter for vSAN traffic. See Set Up a VMkernel Network for vSAN.
Host network	All hosts in your vSAN cluster must be connected to a vSAN Layer 2 or Layer 3 network.
IPv4 and IPv6 support	The vSAN network supports both IPv4 and IPv6.
Network latency	■ Maximum of 1 ms RTT for standard (non-stretched) vSAN clusters between all hosts in the cluster ■ Maximum of 5 ms RTT between the two main sites for stretched clusters ■ Maximum of 200 ms RTT from a main site to the vSAN witness host

License Requirements

Verify that you have a valid license for vSAN.

Using vSAN in production environments requires a special license that you assign to the vSAN clusters.

You can assign a standard vSAN license to the cluster, or a license that covers advanced functions. Advanced features include RAID 5/6 erasure coding, and deduplication and compression. An enterprise license is required for encryption and stretched clusters. For information about assigning licenses, see Configure License Settings for a vSAN Cluster.

The capacity of the license must cover the total number of CPUs in the cluster.

Designing and Sizing a vSAN Cluster

3

For best performance and use, plan the capabilities and configuration of your hosts and their storage devices before you deploy vSAN in a vSphere environment. Carefully consider certain host and networking configurations within the vSAN cluster.

The *Administering VMware vSAN* documentation examines the key points about designing and sizing a vSAN cluster. For detailed instructions about designing and sizing a vSAN cluster, see *VMware vSAN Design and Sizing Guide*.

This chapter includes the following topics:

Designing and Sizing vSAN Storage Components

Plan capacity and cache based on the expected consumption. Consider the requirements for availability and endurance.

- Planning Capacity in vSAN

  You can size the capacity of a vSAN datastore to accommodate the virtual machines (VMs) files in the cluster and to handle failures and maintenance operations.

- Design Considerations for Flash Caching Devices in vSAN

  Plan the configuration of flash devices for vSAN cache and all-flash capacity to provide high performance and required storage space, and to accommodate future growth.

- Design Considerations for Flash Capacity Devices in vSAN

  Plan the configuration of flash capacity devices for vSAN all-flash configurations to provide high performance and required storage space, and to accommodate future growth.

- Design Considerations for Magnetic Disks in vSAN

  Plan the size and number of magnetic disks for capacity in hybrid configurations by following the requirements for storage space and performance.

- Design Considerations for Storage Controllers in vSAN

  Include storage controllers on the hosts of a vSAN cluster that best satisfy the requirements for performance and availability.

Planning Capacity in vSAN

You can size the capacity of a vSAN datastore to accommodate the virtual machines (VMs) files in the cluster and to handle failures and maintenance operations.

Raw Capacity

Use this formula to determine the raw capacity of a vSAN datastore. Multiply the total number of disk groups in the cluster by the size of the capacity devices in those disk groups. Subtract the overhead required by the vSAN on-disk format.

Failures to Tolerate

When you plan the capacity of the vSAN datastore, not including the number of virtual machines and the size of their VMDK files, you must consider the **Failures to tolerate** of the virtual machine storage policies for the cluster.

The **Failures to tolerate** has an important role when you plan and size storage capacity for vSAN. Based on the availability requirements of a virtual machine, the setting might result in doubled consumption or more, compared with the consumption of a virtual machine and its individual devices.

For example, if the **Failure tolerance method** is set to **RAID-1 (Mirroring) - Performance** and the **Failures to tolerate** (FTT) is set to 1, virtual machines can use about 50 percent of the raw capacity. If the FTT is set to 2, the usable capacity is about 33 percent. If the FTT is set to 3, the usable capacity is about 25 percent.

But if the **Failure tolerance method** is set to **RAID-5/6 (Erasure Coding) - Capacity** and the FTT is set to 1, virtual machines can use about 75 percent of the raw capacity. If the FTT is set to 2, the usable capacity is about 67 percent. For more information about RAID 5/6, see *Administering VMware vSAN*.

For information about the attributes in a vSAN storage policy, see *Administering VMware vSAN*.

Calculating Required Capacity

Plan the capacity required for the virtual machines in a cluster with RAID 1 mirroring based on the following criteria:

1 Calculate the storage space that the virtual machines in the vSAN cluster are expected to consume.

```
expected overall consumption = number of VMs in the cluster * expected percentage of
consumption per VMDK
```

2 Consider the **Failures to tolerate** attribute configured in the storage policies for the virtual machines in the cluster. This attribute directly impacts the number of replicas of a VMDK file on hosts in the cluster.

```
datastore capacity = expected overall consumption * (FTT + 1)
```

3 Estimate the overhead requirement of the vSAN on-disk format.

- On-disk format version 3.0 and later adds an extra overhead, typically no more than 1-2 percent capacity per device. Deduplication and compression with software checksum enabled require extra overhead of approximately 6.2 percent capacity per device.
- On-disk format version 2.0 adds an extra overhead, typically no more than 1-2 percent capacity per device.
- On-disk format version 1.0 adds an extra overhead of approximately 1 GB per capacity device.

Capacity Sizing Guidelines

- Keep at least 30 percent unused space to prevent vSAN from rebalancing the storage load. vSAN rebalances the components across the cluster whenever the consumption on a single capacity device reaches 80 percent or more. The rebalance operation might impact the performance of applications. To avoid these issues, keep storage consumption to less than 70 percent. vSAN 7.0 Update 1 and later enables you to manage unused capacity using operations reserve and host rebuild reserve.
- Plan extra capacity to handle any potential failure or replacement of capacity devices, disk groups, and hosts. When a capacity device is not reachable, vSAN recovers the components from another device in the cluster. When a flash cache device fails or is removed, vSAN recovers the components from the entire disk group.
- Reserve extra capacity to make sure that vSAN recovers components after a host failure or when a host enters maintenance mode. For example, provision hosts with enough capacity so that you have sufficient free capacity left for components to rebuild after a host failure or during maintenance. This extra space is important when you have more than three hosts, so you have sufficient free capacity to rebuild the failed components. If a host fails, the

rebuilding takes place on the storage available on another host, so that another failure can be tolerated. However, in a three-host cluster, vSAN does not perform the rebuild operation if the **Failures to tolerate** is set to 1 because when one host fails, only two hosts remain in the cluster. To tolerate a rebuild after a failure, you must have at least three surviving hosts.

- Provide enough temporary storage space for changes in the vSAN VM storage policy. When you dynamically change a VM storage policy, vSAN might create a new RAID tree layout of the object. When vSAN instantiates and synchronizes a new layout, the object may consume extra space temporarily. Keep some temporary storage space in the cluster to handle such changes.
- If you plan to use advanced features, such as software checksum or deduplication and compression, reserve extra capacity to handle the operational overhead.
- Include additional hosts other than what is specified in the storage policy to provide durability during failures. For more information, see *Administering VMware vSAN Guide*.

Considerations for Virtual Machine Objects

When you plan the storage capacity in the vSAN datastore, consider the space required in the datastore for the VM home namespace objects, snapshots, and swap files.

- VM Home Namespace. You can assign a storage policy specifically to the home namespace object for a virtual machine. To prevent unnecessary allocation of capacity and cache storage, vSAN applies only the **Failures to tolerate** and the **Force provisioning** settings from the policy on the VM home namespace. Plan storage space to meet the requirements for a storage policy assigned to a VM Home Namespace whose **Failures to tolerate** is greater than 0.
- Snapshots. Delta devices inherit the policy of the base VMDK file. Plan extra space according to the expected size and number of snapshots, and to the settings in the vSAN storage policies.

  The space that is required might be different. Its size depends on how often the virtual machine changes data and how long a snapshot is attached to the virtual machine.
- Swap files. In vSAN 6.7 and later, virtual machine swap files inherit the storage policy of the VM Namespace.

Design Considerations for Flash Caching Devices in vSAN

Plan the configuration of flash devices for vSAN cache and all-flash capacity to provide high performance and required storage space, and to accommodate future growth.

Choosing Between PCIe or SSD Flash Devices

Choose SSD flash devices according to the requirements for performance, capacity, write endurance, and cost of the vSAN storage.

- Compatibility. The model of the SSD devices must be listed in the vSAN section of the *VMware Compatibility Guide*.

- Performance. PCIe devices generally have faster performance than SATA devices.
- Capacity. The maximum capacity that is available for PCIe devices is generally greater than the maximum capacity that is currently listed for SATA devices for vSAN in the *VMware Compatibility Guide*.
- Write endurance. The write endurance of the SSD devices must meet the requirements for capacity or for cache in all-flash configurations, and for cache in hybrid configurations.

  For information about the write endurance requirements for all-flash and hybrid configurations, see the *VMware vSAN Design and Sizing Guide*. For information about the write endurance class of SSD devices, see the vSAN section of the *VMware Compatibility Guide*.
- Cost. PCIe devices generally have higher cost than SSD devices.

Flash Devices as vSAN Cache

Design the configuration of flash cache for vSAN for write endurance, performance, and potential growth based on these considerations.

Table 3-1. Sizing vSAN Cache

Storage Configuration	Considerations
All-flash and hybrid configurations	■ A higher cache-to-capacity ratio eases future capacity growth. Oversizing cache enables you to add more capacity to an existing disk group without the need to increase the size of the cache. ■ Flash caching devices must have high write endurance. ■ Replacing a flash caching device is more complicated than replacing a capacity device because such an operation impacts the entire disk group. ■ If you add more flash devices to increase the size of the cache, you must create more disk groups. The ratio between flash cache devices and disk groups is always 1:1. A configuration of multiple disk groups provides the following advantages: ■ Reduced risk of failure. If a single caching device fails, fewer capacity devices are affected. ■ Potentially improved performance if you deploy multiple disk groups that contain smaller flash caching devices. However, when you configure multiple disk groups, the memory consumption of the hosts increases.
All-flash configurations	In all-flash configurations, vSAN uses the cache layer for write caching only. The write cache must be able to handle high write activities. This approach extends the life of capacity flash that might be less expensive and might have lower write endurance. For guidance on determining the cache ratio for all-flash clusters, refer to Designing vSAN Disk groups – All Flash Cache Ratio Update.
Hybrid configurations	The flash caching device must provide at least 10 percent of the anticipated storage that virtual machines are expected to consume, not including replicas such as mirrors. The **Primary level of failures to tolerate** attribute from the VM storage policy does not impact the size of the cache. If the read cache reservation is configured in the active VM storage policy, the hosts in the vSAN cluster must have sufficient cache to satisfy the reservation during a post-failure rebuild or maintenance operation. If the available read cache is not sufficient to satisfy the reservation, the rebuild or maintenance operation fails. Use read cache reservation only if you must meet a specific, known performance requirement for a particular workload. The use of snapshots consumes cache resources. If you plan to use several snapshots, consider dedicating more cache than the conventional 10 percent cache-to-consumed-capacity ratio.

Design Considerations for Flash Capacity Devices in vSAN

Plan the configuration of flash capacity devices for vSAN all-flash configurations to provide high performance and required storage space, and to accommodate future growth.

Choosing Between PCIe or SSD Flash Devices

Choose SSD flash devices according to the requirements for performance, capacity, write endurance, and cost of the vSAN storage.

- Compatibility. The model of the SSD devices must be listed in the vSAN section of the *VMware Compatibility Guide*.
- Performance. PCIe devices generally have faster performance than SATA devices.
- Capacity. The maximum capacity that is available for PCIe devices is generally greater than the maximum capacity that is currently listed for SATA devices for vSAN in the *VMware Compatibility Guide*.
- Write endurance. The write endurance of the SSD devices must meet the requirements for capacity or for cache in all-flash configurations, and for cache in hybrid configurations.

  For information about the write endurance requirements for all-flash and hybrid configurations, see the *VMware vSAN Design and Sizing Guide*. For information about the write endurance class of SSD devices, see the vSAN section of the *VMware Compatibility Guide*.
- Cost. PCIe devices generally have higher cost than SSD devices.

Flash Devices as vSAN Capacity

In all-flash configurations, vSAN does not use cache for read operations and does not apply the read-cache reservation setting from the VM storage policy. For cache, you can use a small amount of more expensive flash that has high write endurance. For capacity, you can use flash that is less expensive and has lower write endurance.

Plan a configuration of flash capacity devices by following these guidelines:

- For better performance of vSAN, use more disk groups of smaller flash capacity devices.
- For balanced performance and predictable behavior, use the same type and model of flash capacity devices.

Design Considerations for Magnetic Disks in vSAN

Plan the size and number of magnetic disks for capacity in hybrid configurations by following the requirements for storage space and performance.

SAS and NL-SAS Magnetic Devices

Use SAS or NL-SAS magnetic devices by following the requirements for performance, capacity, and cost of the vSAN storage.

- Compatibility. The model of the magnetic disk must be certified and listed in the vSAN section of the *VMware Compatibility Guide*.
- Performance. SAS and NL-SAS devices have faster performance.

- Capacity. The capacity of SAS or NL-SAS magnetic disks for vSAN is available in the vSAN section of the *VMware Compatibility Guide*. Consider using a larger number of smaller devices instead of a smaller number of larger devices.
- Cost. SAS and NL-SAS devices can be expensive.

Magnetic Disks as vSAN Capacity

Plan a magnetic disk configuration by following these guidelines:

- For better performance of vSAN, use many magnetic disks that have smaller capacity.

  You must have enough magnetic disks that provide adequate aggregated performance for transferring data between cache and capacity. Using more small devices provides better performance than using fewer large devices. Using multiple magnetic disk spindles can speed up the destaging process.

  In environments that contain many virtual machines, the number of magnetic disks is also important for read operations when data is not available in the read cache and vSAN reads it from the magnetic disk. In environments that contain a small number of virtual machines, the disk number impacts read operations if the **Number of disk stripes per object** in the active VM storage policy is greater than one.
- For balanced performance and predictable behavior, use the same type and model of magnetic disks in a vSAN datastore.
- Dedicate a high enough number of magnetic disks to satisfy the value of the **Primary level of failures to tolerate** and the **Number of disk stripes per object** attributes in the defined storage policies. For information about the VM storage policies for vSAN, see *Administering VMware vSAN*.

Design Considerations for Storage Controllers in vSAN

Include storage controllers on the hosts of a vSAN cluster that best satisfy the requirements for performance and availability.

- Use storage controller models, and driver and firmware versions that are listed in the *VMware Compatibility Guide*. Search for vSAN in the *VMware Compatibility Guide*.
- Use multiple storage controllers, if possible, to improve performance and to isolate a potential controller failure to only a subset of disk groups.
- Use storage controllers that have the highest queue depths in the *VMware Compatibility Guide*. Using controllers with high queue depth improves performance. For example, when vSAN is rebuilding components after a failure or when a host enters maintenance mode.
- Use storage controllers in passthrough mode for best performance of vSAN. Storage controllers in RAID 0 mode require higher configuration and maintenance efforts compared to storage controllers in passthrough mode.
- Disable caching on the controller, or set caching to 100 percent Read.

Designing and Sizing vSAN Hosts

Plan the configuration of the hosts in the vSAN cluster for best performance and availability.

Memory and CPU

Size the memory and the CPU of the hosts in the vSAN cluster based on the following considerations.

Table 3-2. Sizing Memory and CPU of vSAN Hosts

Compute Resource	Considerations
Memory	■ Memory per virtual machine ■ Memory per host, based on the expected number of virtual machines ■ At least 32-GB memory for fully operational vSAN with 5 disk groups per host and 7 capacity devices per disk group Hosts that have 512-GB memory or less can boot from a USB, SD, or SATADOM device. If the memory of the host is greater than 512 GB, boot the host from a SATADOM or disk device. For more information, see the VMware knowledge base article at https://kb.vmware.com/s/article/2113954
CPU	■ Sockets per host ■ Cores per socket ■ Number of vCPUs based on the expected number of virtual machines ■ vCPU-to-core ratio ■ 10% CPU overhead for vSAN

Host Networking

Provide more bandwidth for vSAN traffic to improve performance.

- If you plan to use hosts that have 1-GbE adapters, dedicate adapters for vSAN only. For all-flash configurations, plan hosts that have dedicated or shared 10-GbE adapters.
- If you plan to use 10-GbE adapters, they can be shared with other traffic types for both hybrid and all-flash configurations.
- If a 10-GbE adapter is shared with other traffic types, use a vSphere Distributed Switch for vSAN traffic to isolate the traffic by using Network I/O Control and VLANs.
- Create a team of physical adapters for vSAN traffic for redundancy.

Multiple Disk Groups

If the flash cache or storage controller stops responding, an entire disk group can fail. As a result, vSAN rebuilds all components for the failed disk group from another location in the cluster.

Use of multiple disk groups, with each disk group providing less capacity, provides the following benefits and disadvantages:

- Benefits
  - Performance is improved because the datastore has more aggregated cache, and I/O operations are faster.
  - Risk of failure is spread among multiple disk groups.
  - If a disk group fails, vSAN rebuilds fewer components, so performance is improved.
- Disadvantages
  - Costs are increased because two or more caching devices are required.
  - More memory is required to handle more disk groups.
  - Multiple storage controllers are required to reduce the risk of a single point of failure.

Drive Bays

For easy maintenance, consider hosts whose drive bays and PCIe slots are at the front of the server body.

Hot Plug and Swap of Devices

Consider the storage controller passthrough mode support for easy hot plugging or replacement of magnetic disks and flash capacity devices on a host. If a controller works in RAID 0 mode, you must perform additional steps before the host can discover the new drive.

Design Considerations for a vSAN Cluster

Design the configuration of hosts and management nodes for best availability and tolerance to consumption growth.

Sizing the vSAN Cluster for Failures to Tolerate

You configure the **Failures to tolerate** (FTT) attribute in the VM storage policies to handle host failures. The number of hosts required for the cluster is calculated as follows: `2 * FTT + 1`. The more failures the cluster is configured to tolerate, the more capacity hosts are required.

If the cluster hosts are connected in rack servers, you can organize the hosts into fault domains to improve resilience against issues such as top-of-rack switch failures and loss of server rack power. See Designing and Sizing vSAN Fault Domains .

Limitations of a Two-Host or Three-Host Cluster Configuration

In a three-host configuration, you can tolerate only one host failure by setting the number of failures to tolerate to 1. vSAN saves each of the two required replicas of virtual machine data on separate hosts. The witness object is on a third host. Because of the small number of hosts in the cluster, the following limitations exist:

- When a host fails, vSAN cannot rebuild data on another host to protect against another failure.
- If a host must enter maintenance mode, vSAN cannot evacuate data from the host to maintain policy compliance. While the host is in maintenance mode, data is exposed to a potential failure or inaccessibility if an additional failure occurs.

  You can use only the **Ensure data accessibility** data evacuation option. **Ensure data accessibility** guarantees that the object remains available during data migration, although it might be at risk if another failure occurs. vSAN objects on two-host or three-host clusters are not policy compliant. When the host exists maintenance mode, objects are rebuilt to ensure policy compliance.

  In any situation where two-host or three-host cluster has an inaccessible host or disk group, vSAN objects are at risk of becoming inaccessible should another failure occur.

Balanced and Unbalanced Cluster Configuration

vSAN works best on hosts with uniform configurations, including storage configurations.

Using hosts with different configurations has the following disadvantages in a vSAN cluster:

- Reduced predictability of storage performance because vSAN does not store the same number of components on each host.
- Different maintenance procedures.
- Reduced performance on hosts in the cluster that have smaller or different types of cache devices.

Deploying vCenter Server on vSAN

If the vCenter Server becomes unavailable, vSAN continues to operate normally and virtual machines continue to run.

If vCenter Server is deployed on the vSAN datastore, and a problem occurs in the vSAN cluster, you can use a Web browser to access each ESXi host and monitor vSAN through the vSphere Host Client. vSAN health information is visible in the Host Client, and also through esxcli commands.

Designing the vSAN Network

Consider networking features that can provide availability, security, and bandwidth guarantee in a vSAN cluster.

For details about the vSAN network configuration, see the *vSAN Network Design Guide*.

Networking Failover and Load Balancing

vSAN uses the teaming and failover policy that is configured on the backing virtual switch for network redundancy only. vSAN does not use NIC teaming for load balancing.

If you plan to configure a NIC team for availability, consider these failover configurations.

Teaming Algorithm	Failover Configuration of the Adapters in the Team
Route based on originating virtual port	Active/Passive
Route based on IP hash	Active/Active with static EtherChannel for the standard switch and LACP port channel for the distributed switch
Route based on physical network adapter load	Active/Active

vSAN supports IP-hash load balancing, but cannot guarantee improvement in performance for all configurations. You can benefit from IP hash when vSAN is among its many consumers. In this case, IP hash performs load balancing. If vSAN is the only consumer, you might observe no improvement. This behavior specifically applies to 1-GbE environments. For example, if you use four 1-GbE physical adapters with IP hash for vSAN, you might not be able to use more than 1 Gbps. This behavior also applies to all NIC teaming policies that VMware supports.

vSAN does not support multiple VMkernel adapters on the same subnet. You can use different VMkernel adapters on different subnets, such as another VLAN or separate physical fabric. Providing availability by using several VMkernel adapters has configuration costs that involve vSphere and the network infrastructure. You can increase network availability by teaming physical network adapters.

Using Unicast in vSAN Network

In vSAN 6.6 and later releases, multicast is not required on the physical switches that support the vSAN cluster. You can design a simple unicast network for vSAN. Earlier releases of vSAN rely on multicast to enable heartbeat and to exchange metadata between hosts in the cluster. If some hosts in your vSAN cluster are running earlier versions of software, a multicast network is still required. For more information about using multicast in a vSAN cluster, refer to an earlier version of *Administering VMware vSAN*.

Note The following configuration is not supported: vCenter Server deployed on a vSAN 6.6 cluster that is using IP addresses from DHCP without reservations. You can use DHCP with reservations, because the assigned IP addresses are bound to the MAC addresses of VMkernel ports.

Using RDMA

vSAN 7.0 Update 2 and later releases can use Remote Direct Memory Access (RDMA). RDMA typically has lower CPU utilization and less I/O latency. If your hosts support the RoCE v2 protocol, you can enable RDMA through the vSAN network service in vSphere Client.

Consider the following guidelines when designing vSAN over RDMA:

- Each vSAN host must have a vSAN certified RDMA-capable NIC, as listed in the vSAN section of the VMware Compatibility Guide. Use only the same model network adapters from the same vendor on each end of the connection. Configure the DCBx mode to IEEE.
- All hosts must support RDMA. If any host loses RDMA support, the entire vSAN cluster switches to TCP.
- The network must be lossless. Configure network switches to use Data Center Bridging with Priority Flow Control. Configure a lossless traffic class for vSAN traffic marked at priority level 3.
- vSAN with RDMA does not support LACP or IP-hash-based NIC teaming. vSAN with RDMA does support NIC failover.
- All hosts must be on the same subnet. vSAN with RDMA supports up to 32 hosts.

Allocating Bandwidth for vSAN by Using Network I/O Control

vSAN traffic can share 10-GbE physical network adapters with other system traffic types, such as vSphere vMotion traffic, vSphere HA traffic, and virtual machine traffic. To guarantee the amount of bandwidth required for vSAN, use vSphere Network I/O Control in the vSphere Distributed Switch.

In vSphere Network I/O Control, you can configure reservation and shares for the vSAN outgoing traffic.

- Set a reservation so that Network I/O Control guarantees that minimum bandwidth is available on the physical adapter for vSAN.
- Set shares so that when the physical adapter assigned for vSAN becomes saturated, certain bandwidth is available to vSAN and to prevent vSAN from consuming the entire capacity of the physical adapter during rebuild and synchronization operations. For example, the physical adapter might become saturated when another physical adapter in the team fails and all traffic in the port group is transferred to the other adapters in the team.

For example, on a 10-GbE physical adapter that handles traffic for vSAN, vSphere vMotion, and virtual machines, you can configure certain bandwidth and shares.

Table 3-3. Example Network I/O Control Configuration for a Physical Adapter That Handles vSAN

Traffic Type	Reservation, Gbps	Shares
vSAN	1	100
vSphere vMotion	0.5	70
Virtual machine	0.5	30

If the 10-GbE adapter becomes saturated, Network I/O Control allocates 5 Gbps to vSAN on the physical adapter.

For information about using vSphere Network I/O Control to configure bandwidth allocation for vSAN traffic, see the *vSphere Networking* documentation.

Marking vSAN Traffic

Priority tagging is a mechanism to indicate to the connected network devices that vSAN traffic has high Quality of Service (QoS) demands. You can assign vSAN traffic to a certain class and mark the traffic accordingly with a Class of Service (CoS) value from 0 (low priority) to 7 (high priority). Use the traffic filtering and marking policy of vSphere Distributed Switch to configure priority levels.

Segmenting vSAN Traffic in a VLAN

Consider isolating vSAN traffic in a VLAN for enhanced security and performance, especially if you share the capacity of the backing physical adapter among several traffic types.

Jumbo Frames

If you plan to use jumbo frames with vSAN to improve CPU performance, verify that jumbo frames are enabled on all network devices and hosts in the cluster.

By default, the TCP segmentation offload (TSO) and large receive offload (LRO) features are enabled on ESXi. Consider whether using jumbo frames improves the performance enough to justify the cost of enabling them on all nodes on the network.

Creating Static Routes for vSAN Networking

You might need to create static routes in your vSAN environment.

In traditional configurations, where vSphere uses a single default gateway, all routed traffic attempts to reach its destination through this gateway.

Note vSAN 7.0 and later enables you to override the default gateway for the vSAN VMkernel adapter on each host, and configure a gateway address for the vSAN network.

However, certain vSAN deployments might require static routing. For example, deployments where the witness is on a different network, or the stretched cluster deployment, where both the data sites and the witness host are on different networks.

To configure static routing on your ESXi hosts, use the esxcli command:

`esxcli network ip route ipv4 add -g gateway-to-use –n remote-network`

remote-network is the remote network that your host must access, and *gateway-to-use* is the interface to use when traffic is sent to the remote network.

For information about network design for stretched clusters, see *Administering VMware vSAN*.

Best Practices for vSAN Networking

Consider networking best practices for vSAN to improve performance and throughput.

- For hybrid configurations, dedicate at least 1-GbE physical network adapter. Place vSAN traffic on a dedicated or shared 10-GbE physical adapter for best networking performance.
- For all-flash configurations, use a dedicated or shared 10-GbE physical network adapter.
- Provision one additional physical NIC as a failover NIC.
- If you use a shared 10-GbE network adapter, place the vSAN traffic on a distributed switch and configure Network I/O Control to guarantee bandwidth to vSAN.

Designing and Sizing vSAN Fault Domains

vSAN fault domains can spread redundancy components across the servers in separate computing racks. In this way, you can protect the environment from a rack-level failure such as loss of power or connectivity.

Fault Domain Constructs

vSAN requires at least three fault domains to support FTT=1. Each fault domain consists of one or more hosts. Fault domain definitions must acknowledge physical hardware constructs that might represent a potential zone of failure, for example, an individual computing rack enclosure.

If possible, use at least four fault domains. Three fault domains do not support certain data evacuation modes, and vSAN is unable to reprotect data after a failure. In this case, you need an additional fault domain with capacity for rebuilding, which you cannot provide with only three fault domains.

If fault domains are enabled, vSAN applies the active virtual machine storage policy to the fault domains instead of the individual hosts.

Calculate the number of fault domains in a cluster based on the **Failures to tolerate** (FTT) attribute from the storage policies that you plan to assign to virtual machines.

```
number of fault domains = 2 * FTT + 1
```

If a host is not a member of a fault domain, vSAN interprets it as a stand-alone fault domain.

Using Fault Domains Against Failures of Several Hosts

Consider a cluster that contains four server racks, each with two hosts. If the **Failures to tolerate** is set to one and fault domains are not enabled, vSAN might store both replicas of an object with hosts in the same rack enclosure. In this way, applications might be exposed to a potential data loss on a rack-level failure. When you configure hosts that could potentially fail together into separate fault domains, vSAN ensures that each protection component (replicas and witnesses) is placed in a separate fault domain.

If you add hosts and capacity, you can use the existing fault domain configuration or you can define fault domains.

For balanced storage load and fault tolerance when using fault domains, consider the following guidelines:

- Provide enough fault domains to satisfy the **Failures to tolerate** that are configured in the storage policies.

  Define at least three fault domains. Define a minimum of four domains for best protection.
- Assign the same number of hosts to each fault domain.
- Use hosts that have uniform configurations.
- Dedicate one fault domain of free capacity for rebuilding data after a failure, if possible.

Using Boot Devices and vSAN

Starting an ESXi installation that is a part of a vSAN cluster from a flash device imposes certain restrictions.

When you boot a vSAN host from a USB/SD device, you must use a high-quality USB or SD flash drive of 4 GB or larger.

When you boot a vSAN host from a SATADOM device, you must use single-level cell (SLC) device. The size of the boot device must be at least 16 GB.

During installation, the ESXi installer creates a coredump partition on the boot device. The default size of the coredump partition satisfies most installation requirements.

- If the memory of the ESXi host has 512 GB of memory or less, you can boot the host from a USB, SD, or SATADOM device.
- If the memory of the ESXi host has more than 512 GB, consider the following guidelines.
  - You can boot the host from a SATADOM or disk device with a size of at least 16 GB. When you use a SATADOM device, use a single-level cell (SLC) device.
  - If you are using vSAN 6.5 or later, you must resize the coredump partition on ESXi hosts to boot from USB/SD devices. For more information, see the VMware knowledge base article at http://kb.vmware.com/kb/2147881.

Hosts that boot from a disk have a local VMFS. If you have a disk with VMFS that runs VMs, you must separate the disk for an ESXi boot that is not for vSAN. In this case you need separate controllers.

Log Information and Boot Devices in vSAN

When you boot ESXi from a USB or SD device, log information and stack traces are lost on host reboot. They are lost because the scratch partition is on a RAM drive. Use persistent storage for logs, stack traces, and memory dumps.

Do not store log information on the vSAN datastore. This configuration is not supported because a failure in the vSAN cluster could impact the accessibility of log information.

Consider the following options for persistent log storage:

- Use a storage device that is not used for vSAN and is formatted with VMFS or NFS.
- Configure the ESXi Dump Collector and vSphere Syslog Collector on the host to send memory dumps and system logs to vCenter Server.

For information about setting up the scratch partition with a persistent location, see the *vCenter Server Installation and Setup* documentation.

Persistent Logging in a vSAN Cluster

Provide storage for persistence of the logs from the hosts in the vSAN cluster.

If you install ESXi on a USB or SD device and you allocate local storage to vSAN, you might not have enough local storage or datastore space left for persistent logging.

To avoid potential loss of log information, configure the ESXi Dump Collector and vSphere Syslog Collector to redirect ESXi memory dumps and system logs to a network server.

For more information about configuring the vSphere Syslog Collector, see http://kb.vmware.com/kb/2021652.

For more information about configuring the ESXi Dump Collector, see https://kb.vmware.com/s/article/2002954.

Preparing a New or Existing Cluster for vSAN

4

Before you enable vSAN on a cluster and start using it as virtual machine storage, provide the infrastructure that is required for correct operation of vSAN.

This chapter includes the following topics:

Selecting or Verifying the Compatibility of Storage Devices

An important step before you deploy vSAN is to verify that your storage devices, drivers, and firmware are compatible with vSAN by consulting the *VMware Compatibility Guide*.

You can choose from several options for vSAN compatibility.

- Use a vSAN ReadyNode server, a physical server that OEM vendors and VMware validate for vSAN compatibility.

- Assemble a node by selecting individual components from validated device models.

***VMware Compatibility Guide* Section**	**Component Type for Verification**
Systems	Physical server that runs ESXi.
vSAN	■ Magnetic disk SAS model for hybrid configurations. ■ Flash device model that is listed in the *VMware Compatibility Guide*. Certain models of PCIe flash devices can also work with vSAN. Consider also write endurance and performance class. ■ Storage controller model that supports passthrough. vSAN can work with storage controllers that are configured for RAID 0 mode if each storage device is represented as an individual RAID 0 group.

Preparing Storage

Provide enough disk space for vSAN and for the virtualized workloads that use the vSAN datastore.

Preparing Storage Devices

Use flash devices and magnetic disks based on the requirements for vSAN.

Verify that the cluster has the capacity to accommodate anticipated virtual machine consumption and the **Failures to tolerate** in the storage policy for the virtual machines.

The storage devices must meet the following requirements so that vSAN can claim them:

- The storage devices are local to the ESXi hosts. vSAN cannot claim remote devices.
- The storage devices do not have any existing partition information.
- On the same host, you cannot have both all-flash and hybrid disk groups.

Prepare Devices for Disk Groups

Each disk group provides one flash caching device and at least one magnetic disk or one flash capacity device. For hybrid clusters, the capacity of the flash caching device must be at least 10 percent of the anticipated consumed storage on the capacity device, without the protection copies. For guidance on determining the cache ratio for all-flash clusters, refer to Designing vSAN Disk groups – All Flash Cache Ratio Update.

vSAN requires at least one disk group on a host that contributes storage to a cluster that consists of at least three hosts. Use hosts that have uniform configuration for best performance of vSAN.

Raw and Usable Capacity

Provide raw storage capacity that is greater than the capacity for virtual machines to handle certain cases.

- Do not include the size of the flash caching devices as capacity. These devices do not contribute storage and are used as cache unless you have added flash devices for storage.
- Provide enough space to handle the **Failures to tolerate** (FTT) value in a virtual machine storage policy. A FTT that is greater than 0 extends the device footprint. If the FTT is set to 1, the footprint is double. If the FTT is set to 2, the footprint is triple, and so on.
- Verify whether the vSAN datastore has enough space for an operation by examining the space on the individual hosts rather than on the consolidated vSAN datastore object. For example, when you evacuate a host, all free space in the datastore might be on the host that you are evacuating. The cluster is not able to accommodate the evacuation to another host.
- Provide enough space to prevent the datastore from running out of capacity, if workloads that have thinly provisioned storage start consuming a large amount of storage.
- Verify that the physical storage can accommodate the reprotection and maintenance mode of the hosts in the vSAN cluster.
- Consider the vSAN overhead to the usable storage space.
  - On-disk format version 1.0 adds an extra overhead of approximately 1 GB per capacity device.
  - On-disk format version 2.0 adds an extra overhead, typically no more than 1-2 percent capacity per device.
  - On-disk format version 3.0 and later adds an extra overhead, typically no more than 1-2 percent capacity per device. Deduplication and compression with software checksum enabled require extra overhead of approximately 6.2 percent capacity per device.

For more information about planning the capacity of vSAN datastores, see the *VMware vSAN Design and Sizing Guide*.

vSAN Policy Impact on Capacity

The vSAN storage policy for virtual machines affects the capacity devices in several ways.

Table 4-1. vSAN VM Policy and Raw Capacity

Aspects of Policy Influence	Description
Policy changes	■ The **Failures to tolerate** (FTT) influences the physical storage space that you must supply for virtual machines. The greater the FTT is for higher availability, the more space you must provide. When FTT is set to 1, it imposes two replicas of the VMDK file of a virtual machine. With FTT set to 1, a VMDK file that is 50 GB requires 100-GB space on different hosts. If the FTT is changed to 2, you must have enough space to support three replicas of the VMDK across the hosts in the cluster, or 150 GB. ■ Some policy changes, such as a new number of disk stripes per object, require temporary resources. vSAN recreates the objects affected by the change. For a certain time, the physical storage must accommodate the old and new objects.
Available space for reprotecting or maintenance mode	When you place a host in maintenance mode or you clone a virtual machine, the datastore might not be able to evacuate the virtual machine objects, although the vSAN datastore indicates that enough space is available. This lack of space can occur if the free space is on the host that is being placed in maintenance mode.

Mark Flash Devices as Capacity Using ESXCLI

You can manually mark the flash devices on each host as capacity devices using esxcli.

Prerequisites

Verify that you are using vSAN 6.5 or later.

Procedure

1 To learn the name of the flash device that you want to mark as capacity, run the following command on each host.

a In the ESXi Shell, run the `esxcli storage core device list` command.

b Locate the device name at the top of the command output and write the name down.

The command takes the following options:

Table 4-2. Command Options

Options	Description
`-d|--disk=str`	The name of the device that you want to tag as a capacity device. For example, `mpx.vmhba1:C0:T4:L0`
`-t|--tag=str`	Specify the tag that you want to add or remove. For example, the `capacityFlash` tag is used for marking a flash device for capacity.

The command lists all device information identified by ESXi.

2 In the output, verify that the `Is SSD` attribute for the device is `true`.

3 To tag a flash device as capacity, run the `esxcli vsan storage tag add -d <device name> -t capacityFlash` command.

   For example, the `esxcli vsan storage tag add -t capacityFlash -d mpx.vmhba1:C0:T4:L0` command, where `mpx.vmhba1:C0:T4:L0` is the device name.

4 Verify whether the flash device is marked as capacity.

   a In the output, identify whether the `IsCapacityFlash` attribute for the device is set to `1`.

Example: Command Output

You can run the `vdq -q -d <device name>` command to verify the `IsCapacityFlash` attribute. For example, running the `vdq -q -d mpx.vmhba1:C0:T4:L0` command, returns the following output.

```
\{
"Name"     : "mpx.vmhba1:C0:T4:L0",
"VSANUUID" : "",
"State"    : "Eligible for use by VSAN",
"ChecksumSupport": "0",
"Reason"   : "None",
"IsSSD"    : "1",
"IsCapacityFlash": "1",
"IsPDL"    : "0",
    \},
```

Untag Flash Devices Used as Capacity Using ESXCLI

You can untag flash devices that are used as capacity devices, so that they are available for caching.

Procedure

1 To untag a flash device marked as capacity, run the `esxcli vsan storage tag remove -d <device name> -t capacityFlash` command. For example, the `esxcli vsan storage tag remove -t capacityFlash -d mpx.vmhba1:C0:T4:L0` command, where `mpx.vmhba1:C0:T4:L0` is the device name.

2 Verify whether the flash device is untagged.

   a In the output, identify whether the `IsCapacityFlash` attribute for the device is set to `0`.

Example: Command Output

You can run the `vdq -q -d <device name>` command to verify the `IsCapacityFlash` attribute. For example, running the `vdq -q -d mpx.vmhba1:C0:T4:L0` command, returns the following output.

```
[
    \{
"Name"     : "mpx.vmhba1:C0:T4:L0",
"VSANUUID" : "",
"State"    : "Eligible for use by VSAN",
"ChecksumSupport": "0",
"Reason"   : "None",
"IsSSD"    : "1",
"IsCapacityFlash": "0",
"IsPDL"    : "0",
    \},
```

Mark Flash Devices as Capacity Using RVC

Run the `vsan.host_claim_disks_differently` RVC command to mark storage devices as flash, capacity flash, or magnetic disk (HDD).

You can use the RVC tool to tag flash devices as capacity devices either individually, or in a batch by specifying the model of the device. When you want to tag flash devices as capacity devices, you can include them in all-flash disk groups.

Note The `vsan.host_claim_disks_differently` command does not check the device type before tagging them. The command tags any device that you append with the `capacity_flash` command option, including the magnetic disks and devices that are already in use. Make sure that you verify the device status before tagging.

For information about the RVC commands for vSAN management, see the *RVC Command Reference Guide*.

Prerequisites

- Verify that you are using vSAN version 6.5 or later.
- Verify that SSH is enabled on the vCenter Server.

Procedure

1. Open an SSH connection to the vCenter Server.
2. Log in to the vCenter Server by using a local account that has administrator privilege.
3. Start the RVC by running the following command.

   ```
   rvc local_user_name@target_vCenter_Server
   ```

For example, to use the same vCenter Server to mark flash devices for capacity as a user root, run the following command:

```
rvc root@localhost
```

4 Enter the password for the user name.

5 Navigate to the *vcenter_server*/*data_center*/computers/*cluster*/hosts directory in the vSphere infrastructure.

6 Run the `vsan.host_claim_disks_differently` command with the `--claim-type capacity_flash --model` *model_name* options to mark all flash devices of the same model as capacity on all hosts in the cluster.

```
vsan.host_claim_disks_differently --claim-type capacity_flash --model model_name *
```

What to do next

Enable vSAN on the cluster and claim capacity devices.

Providing Memory for vSAN

You must provision hosts with memory according to the maximum number of devices and disk groups that you intend to map to vSAN.

To satisfy the case of the maximum number of devices and disk groups, you must provision hosts with 32 GB of memory for system operations. For information about the maximum device configuration, see the *vSphere Configuration Maximums* documentation.

Preparing Your Hosts for vSAN

As a part of the preparation for enabling vSAN, review the requirements and recommendations about the configuration of hosts for the cluster.

- Verify that the storage devices on the hosts, and the driver and firmware versions for them, are listed in the vSAN section of the *VMware Compatibility Guide*.
- Make sure that a minimum of three hosts contribute storage to the vSAN datastore.
- For maintenance and remediation operations on failure, add at least four hosts to the cluster.
- Designate hosts that have uniform configuration for best storage balance in the cluster.
- Do not add hosts that have only compute resources to the cluster to avoid unbalanced distribution of storage components on the hosts that contribute storage. Virtual machines that require much storage space and run on compute-only hosts might store a great number of components on individual capacity hosts. As a result, the storage performance in the cluster might be lower.

- Do not configure aggressive CPU power management policies on the hosts for saving power. Certain applications that are sensitive to CPU speed latency might have low performance. For information about CPU power management policies, see the *vSphere Resource Management* documentation.
- If your cluster contains blade servers, consider extending the capacity of the datastore with an external storage enclosure that is connected to the blade servers. Make sure the storage enclosure is listed in the vSAN section of the *VMware Compatibility Guide*.
- Consider the configuration of the workloads that you place on a hybrid or all-flash disk configuration.
  - For high levels of predictable performance, provide a cluster of all-flash disk groups.
  - For balance between performance and cost, provide a cluster of hybrid disk groups.

vSAN and vCenter Server Compatibility

Synchronize the versions of vCenter Server and of ESXi to avoid potential faults because of differences in the vSAN support in vCenter Server and ESXi.

For best integration between vSAN components on vCenter Server and ESXi, deploy the latest version of the two vSphere components. See the *vCenter Server Installation and Setup* and *vSphere Upgrade* documentation.

Preparing Storage Controllers

Configure the storage controller on a host according to the requirements of vSAN.

Verify that the storage controllers on the vSAN hosts satisfy certain requirements for mode, driver, and firmware version, queue depth, caching and advanced features.

Table 4-3. Examining Storage Controller Configuration for vSAN

Storage Controller Feature	Storage Controller Requirement
Required mode	■ Review the vSAN requirements in the *VMware Compatibility Guide* for the required mode, passthrough or RAID 0, of the controller. ■ If both passthrough and RAID 0 modes are supported, configure passthrough mode instead of RAID0. RAID 0 introduces complexity for disk replacement.
RAID mode	■ In the case of RAID 0, create one RAID volume per physical disk device. ■ Do not enable a RAID mode other than the mode listed in the *VMware Compatibility Guide*. ■ Do not enable controller spanning.

Table 4-3. Examining Storage Controller Configuration for vSAN (continued)

Storage Controller Feature	Storage Controller Requirement
Driver and firmware version	■ Use the latest driver and firmware version for the controller according to *VMware Compatibility Guide*. ■ If you use the in-box controller driver, verify that the driver is certified for vSAN. OEM ESXi releases might contain drivers that are not certified and listed in the *VMware Compatibility Guide*.
Queue depth	Verify that the queue depth of the controller is 256 or higher. Higher queue depth provides improved performance.
Cache	Disable the storage controller cache, or set it to 100 percent read if disabling cache is not possible.
Advanced features	Disable advanced features, for example, HP SSD Smart Path.

Configuring vSAN Network

Before you enable vSAN on a cluster and ESXi hosts, you must construct the necessary network to carry the vSAN communication.

vSAN provides a distributed storage solution, which implies exchanging data across the ESXi hosts that participate in the cluster. Preparing the network for installing vSAN includes certain configuration aspects.

For information about network design guidelines, see Designing the vSAN Network.

Placing Hosts in the Same Subnet

Hosts must be connected in the same subnet for best networking performance. In vSAN 6.0 and later, you can also connect hosts in the same Layer 3 network if necessary.

Dedicating Network Bandwidth on a Physical Adapter

Allocate at least 1 Gbps bandwidth for vSAN. You might use one of the following configuration options:

- Dedicate 1-GbE physical adapters for a hybrid host configuration.
- Use dedicated or shared 10-GbE physical adapters for all-flash configurations.
- Use dedicated or shared 10-GbE physical adapters for hybrid configurations if possible.
- Direct vSAN traffic on a 10-GbE physical adapter that handles other system traffic and use vSphere Network I/O Control on a distributed switch to reserve bandwidth for vSAN.

Configuring a Port Group on a Virtual Switch

Configure a port group on a virtual switch for vSAN.

- Assign the physical adapter for vSAN to the port group as an active uplink.

When you need a NIC team for network availability, select a teaming algorithm based on the connection of the physical adapters to the switch.

- If designed, assign vSAN traffic to a VLAN by enabling tagging in the virtual switch.

Examining the Firewall on a Host for vSAN

vSAN sends messages on certain ports on each host in the cluster. Verify that the host firewalls allow traffic on these ports.

When you enable vSAN on a cluster, all required ports are added to ESXi firewall rules and configured automatically. There is no need for an administrator to open any firewall ports or enable any firewall services manually.

You can view open ports for incoming and outgoing connections. Select the ESXi host, and click **Configure > Security Profile**.

Considerations about the vSAN License

When you prepare your cluster for vSAN, review the requirements of the vSAN license.

- Make sure that you obtained a valid license for full host configuration control in the cluster. The license should be different from the one that you used for evaluation purposes.

  After the license or the evaluation period of a vSAN expires, you can continue to use the current configuration of vSAN resources. However, you cannot add capacity to a disk group or create disk groups.

- If the cluster consists of all-flash disk groups, verify that the all-flash feature is available under your license.
- If the vSAN cluster uses advanced features such as deduplication and compression or stretched cluster, verify that the feature is available under your license.
- Consider the CPU capacity of the vSAN license across the cluster when adding and removing hosts to the cluster.

  vSAN licenses have per CPU capacity. When you assign a vSAN license to a cluster, the amount of license capacity used is equal to the total number of CPUs on the hosts that participate in the cluster.

  For more information about vSAN licensing editions and potential licensing scenarios, see the VMware VSAN Licensing Guide.

Creating a vSAN Cluster

5

You can activate vSAN when you create a cluster or enable vSAN on your existing clusters.

This chapter includes the following topics:

Characteristics of a vSAN Cluster

Before working on a vSAN environment, be aware of the characteristics of a vSAN cluster.

A vSAN cluster includes the following characteristics:

- You can have multiple vSAN clusters for each vCenter Server instance. You can use a single vCenter Server to manage more than one vSAN cluster.
- vSAN consumes all devices, including flash cache and capacity devices, and does not share devices with other features.
- vSAN clusters can include hosts with or without capacity devices. The minimum requirement is three hosts with capacity devices. For best results, create a vSAN cluster with uniformly configured hosts.
- If a host contributes capacity, it must have at least one flash cache device and one capacity device.

- In hybrid clusters, the magnetic disks are used for capacity and flash devices for read and write cache. vSAN allocates 70 percent of all available cache for read cache and 30 percent of available cache for the write buffer. In a hybrid configuration, the flash devices serve as a read cache and a write buffer.
- In all-flash clusters, one designated flash device is used as a write cache, additional flash devices are used for capacity. In all-flash clusters, all read requests come directly from the flash pool capacity.
- Only local or direct-attached capacity devices can participate in a vSAN cluster. vSAN cannot consume other external storage, such as SAN or NAS, attached to cluster.

To learn about the characteristics of a vSAN cluster configured through Quickstart, see Using Quickstart to Configure and Expand a vSAN Cluster.

For best practices about designing and sizing a vSAN cluster, see Chapter 3 Designing and Sizing a vSAN Cluster.

Before Creating a vSAN Cluster

This topic provides a checklist of software and hardware requirements for creating a vSAN cluster. You can also use the checklist to verify that the cluster meets the guidelines and basic requirements.

Requirements for vSAN Cluster

Before you get started, verify specific models of hardware devices, and specific versions of drivers and firmware in the VMware Compatibility Guide website at http://www.vmware.com/resources/compatibility/search.php. The following table lists the key software and hardware requirements supported by vSAN.

Caution Using uncertified software and hardware components, drivers, controllers, and firmware might cause unexpected data loss and performance issues.

Table 5-1. vSAN Cluster Requirements

Requirements	Description
ESXi hosts	■ Verify that you are using the latest version of ESXi on your hosts. ■ Verify that there are at least three ESXi hosts with supported storage configurations available to be assigned to the vSAN cluster. For best results, configure the vSAN cluster with four or more hosts.
Memory	■ Verify that each host has a minimum of 32 GB of memory. ■ For larger configurations and better performance, you must have a minimum of 32 GB of memory in the cluster. See Designing and Sizing vSAN Hosts.

Table 5-1. vSAN Cluster Requirements (continued)

Requirements	Description
Storage I/O controllers, drivers, firmware	■ Verify that the storage I/O controllers, drivers, and firmware versions are certified and listed in the VCG website at http://www.vmware.com/resources/compatibility/search.php. ■ Verify that the controller is configured for passthrough or RAID 0 mode. ■ Verify that the controller cache and advanced features are disabled. If you cannot disable the cache, you must set the read cache to 100 percent. ■ Verify that you are using controllers with higher queue depths. Using controllers with queue depths less than 256 can significantly impact the performance of your virtual machines during maintenance and failure.
Cache and capacity	■ Verify that vSAN hosts contributing storage to the cluster must have at least one cache and one capacity device. vSAN requires exclusive access to the local cache and capacity devices of the hosts in the vSAN cluster. They cannot share these devices with other uses, such as Virtual Flash File System (VFFS), VMFS partitions, or an ESXi boot partition. ■ For best results, create a vSAN cluster with uniformly configured hosts.
Network connectivity	■ Verify that each host is configured with at least one network adapter. ■ For hybrid configurations, verify that vSAN hosts have a minimum dedicated bandwidth of 1 GbE. ■ For all-flash configurations, verify that vSAN hosts have a minimum bandwidth of 10 GbE. For best practices and considerations about designing the vSAN network, see Designing the vSAN Network and Networking Requirements for vSAN.
vSAN and vCenter Server compatibility	Verify that you are using the latest version of the vCenter Server.
License key	■ Verify that you have a valid vSAN license key. ■ To use the all-flash feature, your license must support that capability. ■ To use advanced features, such as stretched clusters or deduplication and compression, your license must support those features. ■ Verify that the amount of license capacity that you plan on using equals the total number of CPUs in the hosts participating in the vSAN cluster. Do not provide license capacity only for hosts providing capacity to the cluster. For information about licensing for vSAN, see the *vCenter Server and Host Management* documentation.

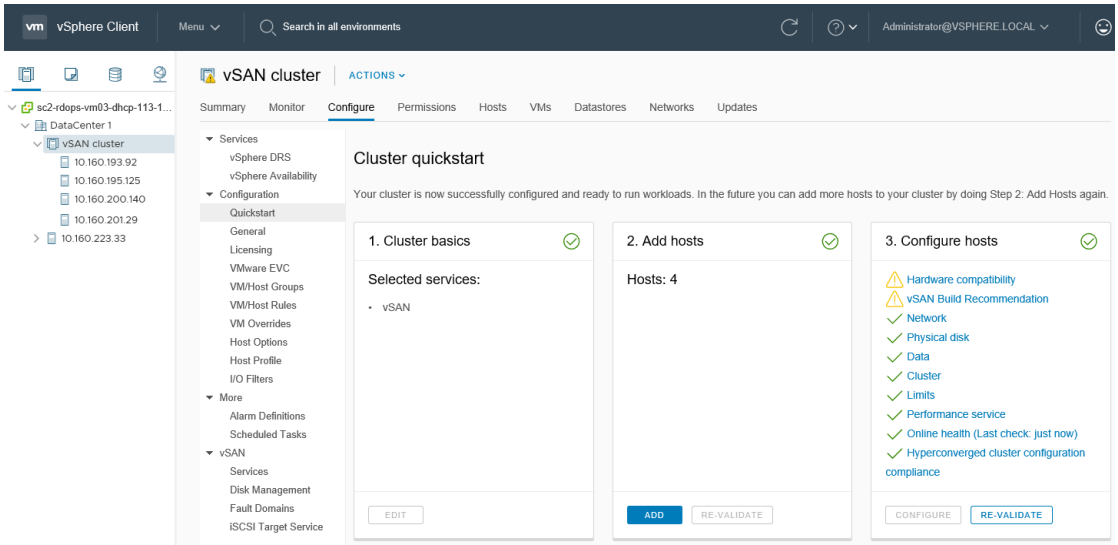
For detailed information about vSAN cluster requirements, see Chapter 2 Requirements for Enabling vSAN.

For in-depth information about designing and sizing the vSAN cluster, see the *VMware vSAN Design and Sizing Guide*.

Using Quickstart to Configure and Expand a vSAN Cluster

You can use the Quickstart workflow to quickly create, configure, and expand a vSAN cluster.

Quickstart consolidates the workflow to enable you to quickly configure a new vSAN cluster that uses recommended default settings for common functions such as networking, storage, and services. Quickstart groups common tasks and uses configuration wizards that guide you through the process. Once you enter the required information on each wizard, Quickstart configures the cluster based on your input.



Quickstart uses the vSAN health service to validate the configuration and help you correct configuration issues. Each Quickstart card displays a configuration checklist. You can click a green message, yellow warning, or red failure to display details.

Hosts added to a Quickstart cluster are automatically configured to match the cluster settings. The ESXi software and patch levels of new hosts must match those in the cluster. Hosts cannot have any networking or vSAN configuration when added to a cluster using the Quickstart workflow. For more information about adding hosts, see "Expanding a vSAN Cluster" in *Administering VMware vSAN*.

Characteristics of a Quickstart Cluster

A vSAN cluster configured using Quickstart has the following characteristics.

- Hosts must have ESXi 6.0 Update 2 or later.
- Host all have similar configuration, including network settings. Quickstart modifies network settings on each host to match the cluster requirements.
- Cluster configuration is based on recommended default settings for networking and services.
- Licenses are not assigned through the Quickstart workflow. You must manually assign a license to your cluster.

Managing and Expanding a Quickstart Cluster

Once you complete the Quickstart workflow, you can manage the cluster through vCenter Server, using the vSphere Client or command-line interface.

You can use the Quickstart workflow to add hosts to the cluster and claim additional disks. But once the cluster is configured through Quickstart, you cannot use Quickstart to modify the cluster configuration.

The Quickstart workflow is available only through the HTML5-based vSphere Client.

Skipping Quickstart

You can use the **Skip Quickstart** button to exit the Quickstart workflow, and continue configuring the cluster and its hosts manually. You can add new hosts individually, and manually configure those hosts. Once skipped, you cannot restore the Quickstart workflow for the cluster.

The Quickstart workflow is designed for new clusters. When you upgrade an existing vSAN cluster to 6.7 Update 1 or later, the Quickstart workflow appears. Skip the Quickstart workflow and continue to manage the cluster through vCenter Server.

Use Quickstart to Configure a vSAN Cluster

You can use the Quickstart workflow to quickly configure a vSAN cluster.

Prerequisites

- Verify that hosts are running ESXi 6.0 Update 2 or later.
- Verify that ESXi hosts in the cluster do not have any existing vSAN or networking configuration.

Note If you perform network configuration through Quickstart, then modify those parameters from outside of Quickstart, you cannot use Quickstart to add or configure additional hosts.

Procedure

1. Navigate to the cluster in the vSphere Client.
2. Click the Configure tab, and select **Configuration > Quickstart**.
3. (optional) On the Cluster basics card, click **Edit** to open the Cluster basics wizard.
   a. (Optional) Enter a cluster name.
   b. Select basic services, such as DRS, vSphere HA, and vSAN.
   c. Click **OK** or **Finish**.

4 On the Add hosts card, click **Add** to open the Add hosts wizard.

   a On the Add hosts page, enter information for new hosts, or click Existing hosts and select from hosts listed in the inventory.

   b On the Host summary page, verify the host settings.

   c On the Ready to complete page, click **Finish**.

Note If you are running vCenter Server on a host, the host cannot be placed into maintenance mode as you add it to a cluster using the Quickstart workflow. The same host also can be running a Platform Services Controller. All other VMs on the host must be powered off.

5 On the Cluster configuration card, click **Configure** to open the Cluster configuration wizard.

   a On the Configure the distributed switches page, enter networking settings, including distributed switches, port groups, and physical adapters.

      - In the **Distributed switches** section, enter the number of distributed switches to configure from the drop-down menu. Enter a name for each distributed switch. Click **Use Existing** to select an existing distributed switch.

        If the host has a standard virtual switch with the same name as the selected distributed switch, the standard switch is migrated to the corresponding distributed switch.

        Network resource control is enabled and set to version 3. Distributed switches with network resource control version 2 cannot be used.

      - In the **Port Groups** section, select a distributed switch to use for vMotion and a distributed switch to use for the vSAN network.

      - In the **Physical adapters** section, select a distributed switch for each physical network adapter. You must assign each distributed switch to at least one physical adapter.

        If the physical adapters chosen are attached to a standard virtual switch with the same name across hosts, the standard switch is migrated to the distributed switch. If the physical adapters chosen are unused, there is no migration from standard switch to distributed switch.

        Network resource control is enabled and set to version 3. Distributed switches with network resource control version 2 cannot be used.

   b On the vMotion traffic page, enter IP address information for vMotion traffic.

   c On the Storage traffic page, enter IP address information for storage traffic.

   d On the Advanced options page, enter information for cluster settings, including DRS, HA, vSAN, host options, and EVC.

e On the Claim disks page, select disks on each host for cache and capacity.

Note Only the vSAN Data Persistence platform can consume vSAN Direct storage. The vSAN Data Persistence platform provides a framework for software technology partners to integrate with VMware infrastructure. Each partner must develop their own plug-in for VMware customers to receive the benefits of the vSAN Data Persistence platform. The platform is not operational until the partner solution running on top is operational. For more information, see *vSphere with Tanzu Configuration and Management*.

f (Optional) On the Create fault domains page, define fault domains for hosts that can fail together.

For more information about fault domains, see "Managing Fault Domains in vSAN Clusters" in *Administering VMware vSAN*.

g (Optional) On the Proxy setting page, configure the proxy server if your system uses one.

h On the Review page, verify the cluster settings, and click **Finish**.

What to do next

You can manage the cluster through vCenter Server.

You can add hosts to the cluster through Quickstart. For more information. see "Expanding a vSAN Cluster" in *Administering VMware vSAN*.

Manually Enabling vSAN

To create a vSAN cluster, you create a vSphere host cluster and enable vSAN on the cluster.

A vSAN cluster can include hosts with capacity and hosts without capacity. Follow these guidelines when you create a vSAN cluster.

- A vSAN cluster must include a minimum of three ESXi hosts. For a vSAN cluster to tolerate host and device failures, at least three hosts that join the vSAN cluster must contribute capacity to the cluster. For best results, consider adding four or more hosts contributing capacity to the cluster.
- Only ESXi 5.5 Update 1 or later hosts can join the vSAN cluster.
- Before you move a host from a vSAN cluster to another cluster, make sure that the destination cluster is vSAN enabled.
- To be able to access the vSAN datastore, an ESXi host must be a member of the vSAN cluster.

After you enable vSAN, the vSAN storage provider is automatically registered with vCenter Server and the vSAN datastore is created. For information about storage providers, see the *vSphere Storage* documentation.

Set Up a VMkernel Network for vSAN

To enable the exchange of data in the vSAN cluster, you must provide a VMkernel network adapter for vSAN traffic on each ESXi host.

Procedure

1. Navigate to the host.
2. Click the **Configure** tab.
3. Under **Networking**, select **VMkernel adapters**.
4. Click **Add Networking** to open the Add Networking wizard.
5. On the **Select connection type** page, select **VMkernel Network Adapter** and click **Next**.
6. On the **Select target device** page, configure the target switching device.
7. On the **Port properties** page, select **vSAN** service.
8. Complete the VMkernel adapter configuration.
9. On the **Ready to complete** page, verify that vSAN is Enabled in the status for the VMkernel adapter, and click **Finish**.

Results

vSAN network is enabled for the host.

What to do next

You can enable vSAN on the host cluster.

Create a vSAN Cluster

You can create a cluster, and then configure the cluster for vSAN.

Procedure

1. Right-click a data center and select **New Cluster**.
2. Type a name for the cluster in the **Name** text box.
3. Turn on DRS, vSphere HA, and vSAN for the cluster.
4. Click **OK**.

   The cluster appears in the inventory.
5. Add hosts to the vSAN cluster.

   vSAN clusters can include hosts with or without capacity devices. For best results, add hosts with capacity.

What to do next

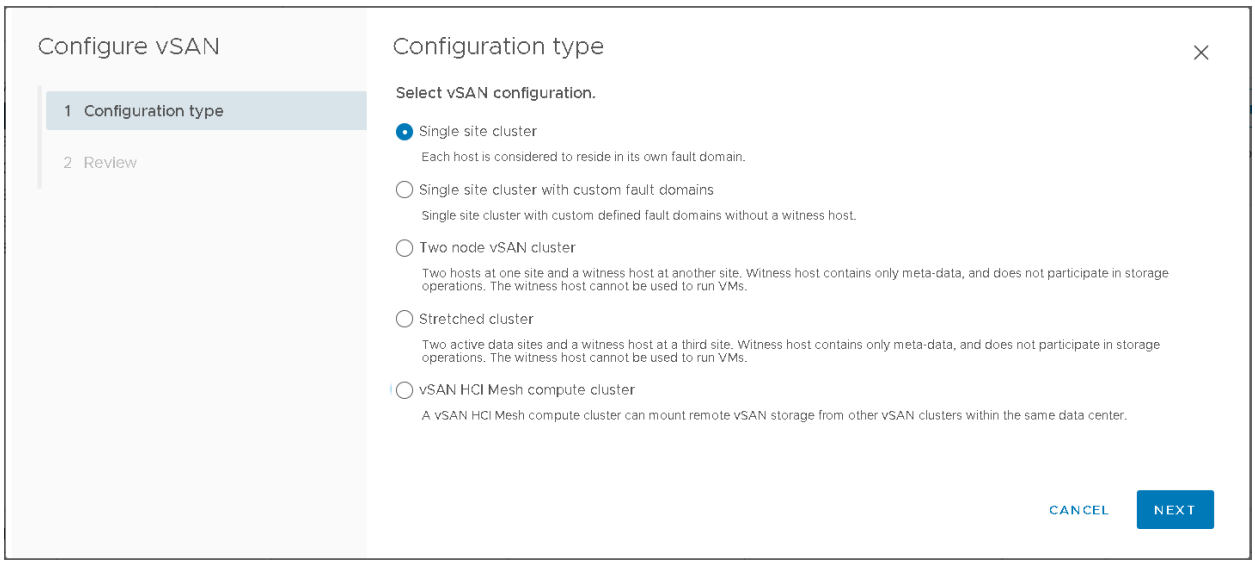
Configure services for the vSAN cluster. See Configure a Cluster for vSAN Using the vSphere Client .

Configure a Cluster for vSAN Using the vSphere Client

You can use the HTML5-based vSphere Client to configure your vSAN cluster.

Note You can use Quickstart to quickly create and configure a vSAN cluster. For more information, see "Using Quickstart to Configure and Expand a vSAN Cluster" in *vSAN Planning and Deployment* .

Note vSAN HCI Mesh compute clusters have limited configuration options.



Prerequisites

Verify that your environment meets all requirements. See "Requirements for Enabling vSAN" in *vSAN Planning and Deployment*.

Create a cluster and add hosts to the cluster before enabling and configuring vSAN.

Procedure

1. Navigate to an existing host cluster.
2. Click the **Configure** tab.
3. Under vSAN, select **Services**.
4. Click **Configure vSAN** to open the Configure vSAN wizard.
5. Select the type of vSAN cluster to configure, and click **Next**.
   - Single site cluster. For more information, see "vSAN Deployment Options" in *vSAN Planning and Deployment*.

- Single site cluster with custom fault domains.
- Two node vSAN cluster.
- Stretched cluster.
- vSAN HCI Mesh compute cluster. For more information, see "Sharing Remote Datastores with HCI Mesh" in *Administering VMware vSAN*.

6 Configure the vSAN services to use, and click **Next**.

Configure data management features, including deduplication and compression, data-at-rest encryption, and data-in-transit encryption. For more details, see Edit vSAN Settings.

7 Claim disks for the vSAN cluster, and click **Next**.

Each host requires at least one flash device in the cache tier, and one or more devices in the capacity tier. For more details, see "Managing Disk Groups and Devices" in *Administering VMware vSAN*.

8 Review the configuration, and click **Finish**.

Results

Enabling vSAN creates a vSAN datastore and registers the vSAN storage provider. vSAN storage providers are built-in software components that communicate the storage capabilities of the datastore to vCenter Server.

What to do next

Claim disks or create disk groups. See "Managing Disk Groups and Devices" in *Administering VMware vSAN*.

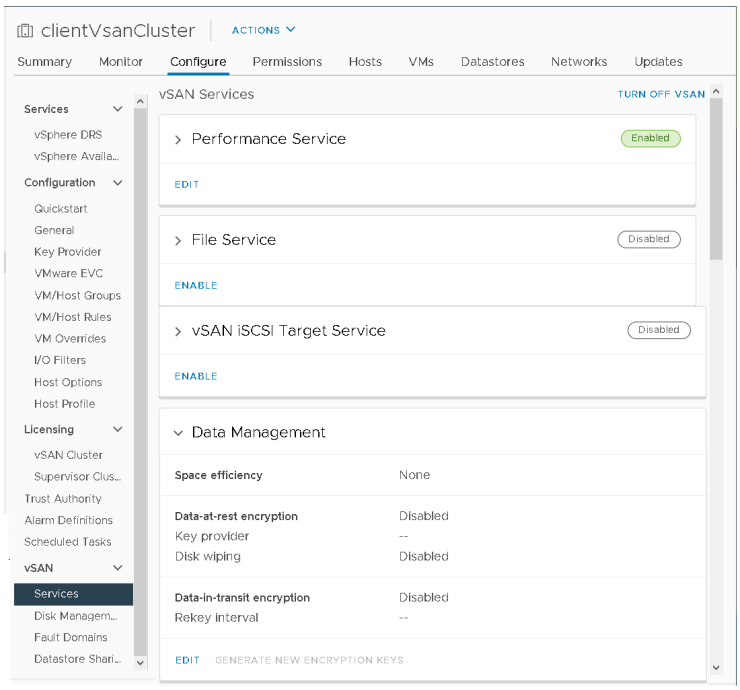
Verify that the vSAN datastore has been created. See View vSAN Datastore.

Verify that the vSAN storage provider is registered.

Edit vSAN Settings

You can edit the settings of your vSAN cluster to configure data management features and enable services provided by the cluster.

Edit the settings of an existing vSAN cluster if you want to enable deduplication and compression, or to enable encryption. If you enable deduplication and compression, or if you enable encryption, the on-disk format of the cluster is automatically upgraded to the latest version.



Procedure

1 Navigate to the vSAN host cluster.

2 Click the **Configure** tab.

   a Under vSAN, select **Services**.

   b Click the **Edit** or **Enable** button for the service you want to configure.

   - Configure vSAN performance service. For more information, see Monitoring vSAN Performance in *vSAN Monitoring and Troubleshooting*.
   - Enable File Service. For more information, see "vSAN File Service" in *Administering VMware vSAN* .
   - Configure vSAN Network options. For more information, see Configuring vSAN Network in *vSAN Planning and Deployment*.
   - Configure vSAN historical health service.
   - Configure iSCSI target service. For more information, see Using the "vSAN iSCSI Target Service" in *Administering VMware vSAN*.
   - Configure Data Management options, including deduplication and compression, data-at-rest encryption, and data-in-transit encryption.
   - Configure capacity reservations and alerts. For more information, see "About Reserved Capacity" in *vSAN Monitoring and Troubleshooting*.
   - Configure advanced options:
     - Object Repair Timer
     - Site Read Locality for stretched clusters
     - Thin Swap provisioning
     - Large Cluster Support for up to 64 hosts
     - Automatic Rebalance

   c Modify the settings to match your requirements.

3 Click **Apply** to confirm your selections.

Enable vSAN on an Existing Cluster

You can edit cluster properties to enable vSAN on an existing cluster.

Prerequisites

Verify that your environment meets all requirements. See "Requirements for Enabling vSAN" in *vSAN Planning and Deployment*.

Note vSAN HCI Mesh compute clusters have limited configuration options.

Procedure

1 Navigate to an existing host cluster.

2 Click the **Configure** tab.

3 Under vSAN, select **Services**.

4 Click **Configure vSAN**.

5 Select the type of vSAN cluster to configure, and click **Next**.

- Single site cluster.
- Single site cluster with custom fault domains.
- Two node vSAN cluster.
- Stretched cluster.
- vSAN HCI Mesh compute cluster. For more information, see "Sharing Remote Datastores with HCI Mesh" in *Administering VMware vSAN*.

6 Configure the vSAN services to use, and click **Next**.

- Configure vSAN performance service. For more information, see "Monitoring vSAN Performance" in *vSAN Monitoring and Troubleshooting*.
- Enable File Service. For more information, see "vSAN File Service" in *Administering VMware vSAN*.
- Configure vSAN Network options. For more information, see "Designing the vSAN Network" in *vSAN Planning and Deployment* .
- Configure vSAN historical health service.
- Configure iSCSI target service. For more information, see "Using the vSAN iSCSI Target Service" in *Administering VMware vSAN*.
- Configure Data Management options, including deduplication and compression, data-at-rest encryption, and data-in-transit encryption.
- Configure capacity reservations and alerts. For more information, see "About Reserved Capacity" in *vSAN Monitoring and Troubleshooting*.
- Configure advanced options:
  - Object Repair Timer
  - Site Read Locality for stretched clusters
  - Thin Swap provisioning
  - Large Cluster Support for up to 64 hosts
  - Automatic Rebalance

7 Claim disks for the vSAN cluster, and click **Next**.

Each host requires at least one flash device in the cache tier, and one or more devices in the capacity tier. For more information, see "Managing Disk Groups and Devices" in *Administering VMware vSAN*.

8 Review the configuration, and click **Finish**.

Configure License Settings for a vSAN Cluster

You must assign a license to a vSAN cluster before its evaluation period expires or its currently assigned license expires.

If you upgrade, combine, or divide vSAN licenses, you must assign the new licenses to vSAN clusters. When you assign a vSAN license to a cluster, the amount of license capacity used equals the total number of CPUs in the hosts participating in the cluster. The license use of the vSAN cluster is recalculated and updated every time you add or remove a host from the cluster. For information about managing licenses and licensing terminology and definitions, see the *vCenter Server and Host Management* documentation.

When you enable vSAN on a cluster, you can use vSAN in evaluation mode to explore its features. The evaluation period starts when vSAN is enabled, and expires after 60 days. To use vSAN, you must license the cluster before the evaluation period expires. Just like vSphere licenses, vSAN licenses have per CPU capacity. Some advanced features, such as all-flash configuration and stretched clusters, require a license that supports the feature.

Prerequisites

- To view and manage vSAN licenses, you must have the **Global.Licenses** privilege on the vCenter Server systems.

Procedure

1 Navigate to your vSAN cluster.

2 Click the **Configure** tab.

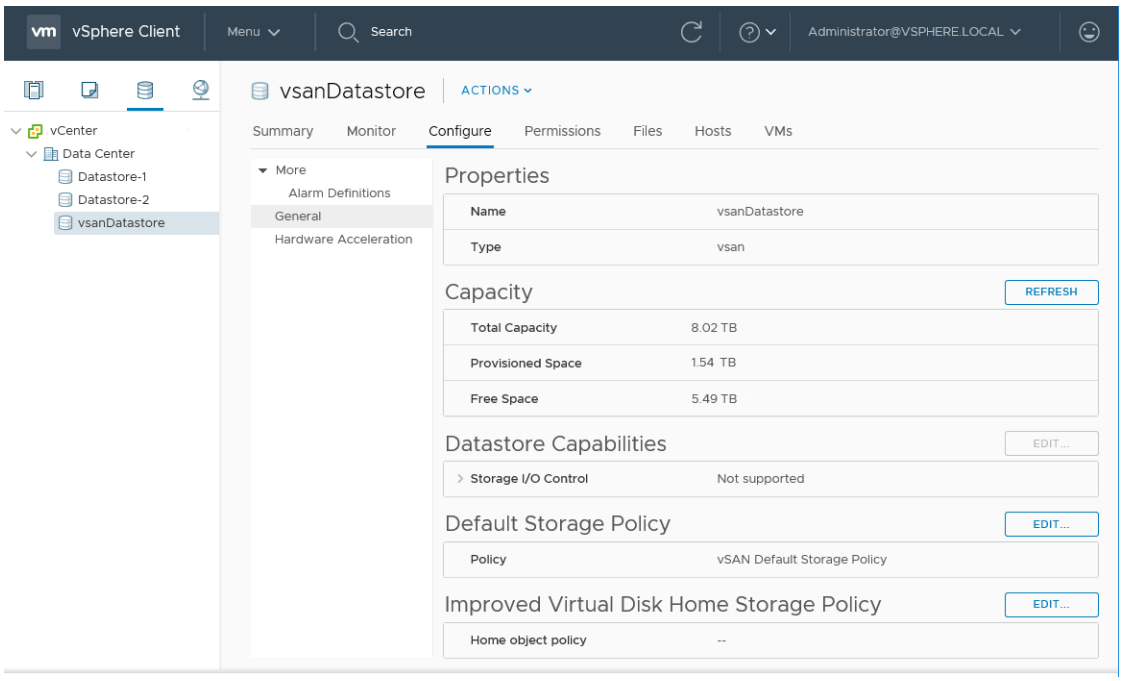
3 Under Licensing, select **vSAN Cluster**.

4 Click Assign License.

5 Select an existing license and click **OK**.

View vSAN Datastore

After you enable vSAN, a single datastore is created. You can review the capacity of the vSAN datastore.



Prerequisites

Activate vSAN and configure disk groups.

Procedure

1. Navigate to Storage.
2. Select the vSAN datastore.
3. Click the **Configure** tab.
4. Review the vSAN datastore capacity.

   The size of the vSAN datastore depends on the number of capacity devices per ESXi host and the number of ESXi hosts in the cluster. For example, if a host has seven 2 TB for capacity devices, and the cluster includes eight hosts, the approximate storage capacity is 7 x 2 TB x 8 = 112 TB. When using the all-flash configuration, flash devices are used for capacity. For hybrid configuration, magnetic disks are used for capacity.

   Some capacity is allocated for metadata.

   - On-disk format version 1.0 adds approximately 1 GB per capacity device.
   - On-disk format version 2.0 adds capacity overhead, typically no more than 1-2 percent capacity per device.
   - On-disk format version 3.0 and later adds capacity overhead, typically no more than 1-2 percent capacity per device. Deduplication and compression with software checksum enabled require additional overhead of approximately 6.2 percent capacity per device.

What to do next

Create a storage policy for virtual machines using the storage capabilities of the vSAN datastore. For information, see the *vSphere Storage* documentation.

Using vSAN and vSphere HA

You can enable vSphere HA and vSAN on the same cluster. vSphere HA provides the same level of protection for virtual machines on vSAN datastores as it does on traditional datastores. This level of protection imposes specific restrictions when vSphere HA and vSAN interact.

ESXi Host Requirements

You can use vSAN with a vSphere HA cluster only if the following conditions are met:

- The cluster's ESXi hosts all must be version 5.5 Update 1 or later.
- The cluster must have a minimum of three ESXi hosts. For best results, configure the vSAN cluster with four or more hosts.

Note vSAN 7.0 Update 2 and later supports Proactive HA. Select the following remediation method: **Maintenance mode for all failures**. Quarantine mode is supported, but it does not protect against data loss if the host in quarantine mode fails, and there are objects with FTT=0 or objects with FTT=1 that are degraded.

Networking Differences

vSAN uses its own logical network. When vSAN and vSphere HA are enabled for the same cluster, the HA interagent traffic flows over this storage network rather than the management network. vSphere HA uses the management network only when vSAN is disabled. vCenter Server chooses the appropriate network when vSphere HA is configured on a host.

Note You must disable vSphere HA before you enable vSAN on the cluster. Then you can re-enable vSphere HA.

When a virtual machine is only partially accessible in all network partitions, you cannot power on the virtual machine or fully access it in any partition. For example, if you partition a cluster into P1 and P2, the VM namespace object is accessible to the partition named P1 and not to P2. The VMDK is accessible to the partition named P2 and not to P1. In such cases, the virtual machine cannot be powered on and it is not fully accessible in any partition .

The following table shows the differences in vSphere HA networking whether or not vSAN is used.

Table 5-2. vSphere HA Networking Differences

	vSAN Enabled	vSAN Disabled
Network used by vSphere HA	vSAN storage network	Management network
Heartbeat datastores	Any datastore mounted to more than one host, but not vSAN datastores	Any datastore mounted to more than one host
Host declared isolated	Isolation addresses not pingable and vSAN storage network inaccessible	Isolation addresses not pingable and management network inaccessible

If you change the vSAN network configuration, the vSphere HA agents do not automatically acquire the new network settings. To change the vSAN network, you must re-enable host monitoring for the vSphere HA cluster:

1 Disable Host Monitoring for the vSphere HA cluster.

2 Make the vSAN network changes.

3 Right-click all hosts in the cluster and select **Reconfigure HA**.

4 Re-enable Host Monitoring for the vSphere HA cluster.

Capacity Reservation Settings

When you reserve capacity for your vSphere HA cluster with an admission control policy, this setting must be coordinated with the corresponding **Primary level of failures to tolerate** policy setting in the vSAN rule set. It must not be lower than the capacity reserved by the vSphere HA admission control setting. For example, if the vSAN rule set allows for only two failures, the vSphere HA admission control policy must reserve capacity that is equivalent to only one or two host failures. If you are using the Percentage of Cluster Resources Reserved policy for a cluster that has eight hosts, you must not reserve more than 25 percent of the cluster resources. In the same cluster, with the **Primary level of failures to tolerate** policy, the setting must not be higher than two hosts. If vSphere HA reserves less capacity, failover activity might be unpredictable. Reserving too much capacity overly constrains the powering on of virtual machines and intercluster vSphere vMotion migrations. For information about the Percentage of Cluster Resources Reserved policy, see the *vSphere Availability* documentation.

vSAN and vSphere HA Behavior in a Multiple Host Failure

After a vSAN cluster fails with a loss of failover quorum for a virtual machine object, vSphere HA might not be able to restart the virtual machine even when the cluster quorum has been restored. vSphere HA guarantees the restart only when it has a cluster quorum and can access the most recent copy of the virtual machine object. The most recent copy is the last copy to be written.

Consider an example where a vSAN virtual machine is provisioned to tolerate one host failure. The virtual machine runs on a vSAN cluster that includes three hosts, H1, H2, and H3. All three hosts fail in a sequence, with H3 being the last host to fail.

After H1 and H2 recover, the cluster has a quorum (one host failure tolerated). Despite this quorum, vSphere HA is unable to restart the virtual machine because the last host that failed (H3) contains the most recent copy of the virtual machine object and is still inaccessible.

In this example, either all three hosts must recover at the same time, or the two-host quorum must include H3. If neither condition is met, HA attempts to restart the virtual machine when host H3 is online again.

Deploying vSAN with vCenter Server

You can create a vSAN cluster as you deploy vCenter Server, and host the vCenter Server on that cluster.

The vCenter Server is a preconfigured virtual machine used to administer ESXi hosts in a cluster. You can host the vCenter Server on a vSAN cluster.

When you use the vCenter Server Installer to deploy a vCenter Server, you can create a single-host vSAN cluster, and host the vCenter Server on the cluster. During Stage 1 of the deployment, when you select a datastore, click **Install on a new vSAN cluster containing the target host**. Follow the steps in the Installer wizard to complete the deployment.

The vCenter Server Installer creates a one-host vSAN cluster, with disks claimed from the host. vCenter Server is deployed on the vSAN cluster.

After you complete the deployment, you can manage the single-host vSAN cluster with the vCenter Server. You must complete the configuration of the vSAN cluster.

You can deploy a Platform Services Controller and vCenter Server on the same vSAN cluster or on separate clusters.

- You can deploy a Platform Services Controller and vCenter Server to the same vSAN cluster. Deploy the PSC and vCenter Server to the same single-host vSAN datastore. After you complete the deployment, the Platform Services Controller and vCenter Server both run on the same cluster.
- You can deploy a Platform Services Controller and vCenter Server to different vSAN clusters. Deploy the Platform Services Controller and vCenter Server to separate single-host vSAN clusters. After you complete the deployment, you must complete the configuration of each vSAN cluster separately.

Disable vSAN

You can turn off vSAN for a host cluster.

When you disable the vSAN cluster, all virtual machines and data services located on the vSAN datastore become inaccessible. If you have consumed storage on the vSAN cluster using vSAN Direct, then the vSAN Direct monitoring services, such as health checks, space reporting, and performance monitoring, are also disabled. If you intend to use virtual machines while vSAN is disabled, make sure you migrate virtual machines from vSAN datastore to another datastore before disabling the vSAN cluster.

Prerequisites

Verify that the hosts are in maintenance mode.

Procedure

1. Navigate to the vSAN cluster.
2. Click the **Configure** tab.
3. Under vSAN, select **Services**.
4. Click **Turn Off vSAN**.
5. On the Turn Off vSAN dialog, confirm your selection.

Extending a Datastore Across Two Sites with Stretched Clusters

6

You can create a stretched cluster that spans two geographic locations (or sites). Stretched clusters enable you to extend the vSAN datastore across two sites to use it as stretched storage. The stretched cluster continues to function if a failure or scheduled maintenance occurs at one site.

This chapter includes the following topics:

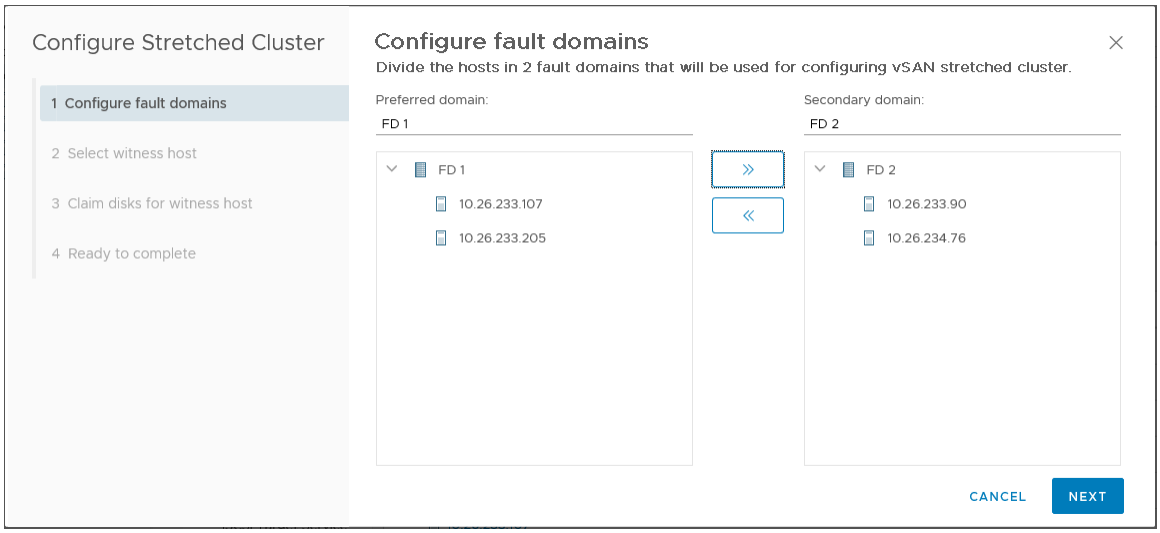
Introduction to Stretched Clusters

Stretched clusters extend the vSAN cluster from a single data site to two sites for a faster level of availability and intersite load balancing. Stretched clusters are typically deployed in environments where the distance between data centers is limited, such as metropolitan or campus environments.

You can use stretched clusters to manage planned maintenance and avoid disaster scenarios, because maintenance or loss of one site does not affect the overall operation of the cluster. In a stretched cluster configuration, both data sites are active sites. If either site fails, vSAN uses the storage on the other site. vSphere HA restarts any VM that must be restarted on the remaining active site.

You must designate one site as the preferred site. The other site becomes a secondary or nonpreferred site. If the network connection between the two active sites is lost, vSAN continues operation with the preferred site. The site designated as preferred typically is the one that remains in operation, unless it is resyncing or has another issue. The site that leads to maximum data availability is the one that remains in operation.

A vSAN stretched cluster can tolerate one link failure at a time without data becoming unavailable. A link failure is a loss of network connection between the two sites or between one site and the witness host. During a site failure or loss of network connection, vSAN automatically switches to fully functional sites.



vSAN 7.0 Update 3 and later stretched clusters can tolerate a witness host failure when one site is unavailable. Configure the storage policy Site disaster tolerance rule to Site mirroring - stretched cluster. If one site is down due to maintenance or failure and the witness host fails, objects become non-compliant but remain accessible.

For more information about working with stretched clusters, see the *vSAN Stretched Cluster Guide*.

Witness Host

Each stretched cluster consists of two data sites and one witness host. The witness host resides at a third site and contains the witness components of virtual machine objects. The witness host does not store customer data, only metadata, such as the size and UUID of vSAN object and components.

The witness host serves as a tiebreaker when a decision must be made regarding availability of datastore components when the network connection between the two sites is lost. In this case, the witness host typically forms a vSAN cluster with the preferred site. But if the preferred site becomes isolated from the secondary site and the witness, the witness host forms a cluster using the secondary site. When the preferred site is online again, data is resynchronized to ensure that both sites have the latest copies of all data.

If the witness host fails, all corresponding objects become noncompliant but are fully accessible.

The witness host has the following characteristics:

- The witness host can use low bandwidth/high latency links.
- The witness host cannot run VMs.
- A single witness host can support only one vSAN stretched cluster. Two-node vSAN clusters can share a single witness host.
- The witness host must have one VMkernel adapter with vSAN traffic enabled, with connections to all hosts in the cluster. The witness host uses one VMkernel adapter for management and one VMkernel adapter for vSAN data traffic. The witness host can have only one VMkernel adapter dedicated to vSAN.
- The witness host must be a standalone host dedicated to the stretched cluster. It cannot be added to any other cluster or moved in inventory through vCenter Server.

The witness host can be a physical host or an ESXi host running inside a VM. The VM witness host does not provide other types of functionality, such as storing or running VMs. Multiple witness hosts can run as VMs on a single physical server. For patching and basic networking and monitoring configuration, the VM witness host works in the same way as a typical ESXi host. You can manage it with vCenter Server, patch it and update it by using `esxcli` or vSphere Lifecycle Manager, and monitor it with standard tools that interact with ESXi hosts.

You can use a witness virtual appliance as the witness host in a stretched cluster. The witness virtual appliance is an ESXi host in a VM, packaged as an OVF or OVA. The appliance is available in different options, based on the size of the deployment.

Stretched Clusters and Fault Domains

Stretched clusters use fault domains to provide redundancy and failure protection across sites. Each site in a stretched cluster resides in a separate fault domain.

A stretched cluster requires three fault domains: the preferred site, the secondary site, and a witness host. Each fault domain represents a separate site. When the witness host fails or enters maintenance mode, vSAN considers it a site failure.

In vSAN 6.6 and later releases, you can provide an extra level of local fault protection for virtual machine objects in stretched clusters. When you configure a stretched cluster, the following policy rules are available for objects in the cluster:

- **Failures to tolerate (FTT)**. For stretched clusters, **FTT** defines the number of site failures that a virtual machine object can tolerate. For a stretched cluster, only a value of 0 or 1 is supported.
- **Site disaster tolerance**. For stretched clusters, this rule defines the number of additional host failures that the object can tolerate after the number of site failures defined by **FTT** is reached.

  The default value is 0, and the maximum value is 3.
- **Data Locality**. This rule is available only if **FTT** = 0. You can set the Data Locality rule to None, Preferred, or Secondary. This rule enables you to restrict virtual machine objects to a selected site in the stretched cluster. The default value is None.

In a stretched cluster with local fault protection, even when one site is unavailable, the cluster can perform repairs on missing or broken components in the available site.

vSAN 7.0 and later continue to serve I/O if any disks or disks on one site reach 96 percent full or 5 GB free capacity (whichever is less) while disks on the other site have free space available. Components on the affected site are marked absent, and vSAN continues to perform I/O to healthy object copies on the other site. When disks on the affected site disk reach 94 percent capacity or 10 GB (whichever is less), the absent components become available. vSAN resyncs the available components and all objects become policy compliant.

Stretched Cluster Design Considerations

Consider these guidelines when working with a vSAN stretched cluster.

- Configure DRS settings for the stretched cluster.
  - DRS must be enabled on the cluster. If you place DRS in partially automated mode, you can control which VMs to migrate to each site. vSAN 7.0 Update 2 enables you to operate DRS in automatic mode, and recover gracefully from network partitions.
  - Create two host groups, one for the preferred site and one for the secondary site.
  - Create two VM groups, one to hold the VMs on the preferred site and one to hold the VMs on the secondary site.
  - Create two VM-Host affinity rules that map VMs-to-host groups, and specify which VMs and hosts reside in the preferred site and which VMs and hosts reside in the secondary site.
  - Configure VM-Host affinity rules to perform the initial placement of VMs in the cluster.
- Configure HA settings for the stretched cluster.
  - HA must be enabled on the cluster.

  - HA rule settings should respect VM-Host affinity rules during failover.
  - Disable HA datastore heartbeats.
- Stretched clusters require on-disk format 2.0 or later. If necessary, upgrade the on-disk format before configuring a stretched cluster. See "Upgrade vSAN Disk Format" in *Administering VMware vSAN*.
- Configure the **FTT** to 1 for stretched clusters.
- vSAN stretched clusters support enabling Symmetric Multiprocessing Fault Tolerance (SMP-FT) VMs when **FTT** is set to 0 and **Data Locality** is set to Preferred or Secondary. vSAN does not support SMP-FT VMs on a stretched cluster with **FTT** set to 1 or more.
- When a host is disconnected or not responding, you cannot add or remove the witness host. This limitation ensures that vSAN collects enough information from all hosts before initiating reconfiguration operations.
- Using `esxcli` to add or remove hosts is not supported for stretched clusters.
- Do not create snapshots of the witness host or backup the witness host. If the witness host fails, Change the Witness Host.

Best Practices for Working with Stretched Clusters

When working with vSAN stretched clusters, follow these recommendations for proper performance.

- If one of the sites (fault domains) in a stretched cluster is inaccessible, new VMs can still be provisioned in the subcluster containing the other two sites. These new VMs are implicitly force provisioned and are non-compliant until the partitioned site rejoins the cluster. This implicit force provisioning is performed only when two of the three sites are available. A site here refers to either a data site or the witness host.
- If an entire site goes offline due to a power outage or loss of network connection, restart the site immediately, without much delay. Instead of restarting vSAN hosts one by one, bring all hosts online approximately at the same time, ideally within a span of 10 minutes. By following this process, you avoid resynchronizing a large amount of data across the sites.
- If a host is permanently unavailable, remove the host from the cluster before you perform any reconfiguration tasks.
- If you want to clone a VM witness host to support multiple stretched clusters, do not configure the VM as a witness host before cloning it. First deploy the VM from OVF, then clone the VM, and configure each clone as a witness host for a different cluster. Or you can deploy as many VMs as you need from the OVF, and configure each one as a witness host for a different cluster.

Stretched Clusters Network Design

All three sites in a stretched cluster communicate across the management network and across the vSAN network. The VMs in both data sites communicate across a common virtual machine network.

A vSAN stretched cluster must meet certain basic networking requirements.

- Management network requires connectivity across all three sites, using a Layer 2 stretched network or a Layer 3 network.
- The vSAN network requires connectivity across all three sites. It must have independent routing and connectivity between the data sites and the witness host. vSAN supports both Layer 2 and Layer 3 between the two data sites, and Layer 3 between the data sites and the witness host.
- VM network requires connectivity between the data sites, but not the witness host. Use a Layer 2 stretched network or Layer 3 network between the data sites. In the event of a failure, the VMs do not require a new IP address to work on the remote site.
- vMotion network requires connectivity between the data sites, but not the witness host. Use a Layer 2 stretched or a Layer 3 network between data sites.

Note vSAN over RDMA is not supported for stretched clusters.

Using Static Routes on ESXi Hosts

If you use a single default gateway on ESXi hosts, each ESXi host contains a default TCP/IP stack that has a single default gateway. The default route is typically associated with the management network TCP/IP stack.

Note vSAN 7.0 and later enables you to override the default gateway for the vSAN VMkernel adapter on each host, and configure a gateway address for the vSAN network.

The management network and the vSAN network might be isolated from one another. For example, the management network might use vmk0 on physical NIC 0, while the vSAN network uses vmk2 on physical NIC 1 (separate network adapters for two distinct TCP/IP stacks). This configuration implies that the vSAN network has no default gateway.

Consider a vSAN network that is stretched over two data sites on a Layer 2 broadcast domain (for example, 172.10.0.0) and the witness host is on another broadcast domain (for example, 172.30.0.0). If the VMkernel adapters on a data site try to connect to the vSAN network on the witness host, the connection fails because the default gateway on the ESXi host is associated with the management network. There is no route from the management network to the vSAN network.

You can use static routes to resolve this issue. Define a new routing entry that indicates which path to follow to reach a particular network. For a vSAN network on a stretched cluster, you can add static routes to ensure proper communication across all hosts.

For example, you can add a static route to the hosts on each data site, so requests to reach the 172.30.0.0 witness network are routed through the 172.10.0.0 interface. Also add a static route to the witness host so that requests to reach the 172.10.0.0 network for the data sites are routed through the 172.30.0.0 interface.

Note If you use static routes, you must manually add the static routes for new ESXi hosts added to either site before those hosts can communicate across the cluster. If you replace the witness host, you must update the static route configuration.

Use the `esxcli network ip route` command to add static routes.

Two-Node vSAN Clusters

A two-node vSAN cluster has two hosts at the same location. The witness function is performed at a second site on a dedicated virtual appliance.

Two-node vSAN clusters are often used for remote office/branch office environments, typically running a small number of workloads that require high availability. A two-node vSAN cluster consists of two hosts at the same location, connected to the same network switch or directly connected. A third host acts as a witness host, which can be located remotely from the branch office. Usually the witness host resides at the main site, with the vCenter Server.

A single witness host can support up to 64 two-node clusters. The number of clusters supported by a shared witness host is based on the host memory.

Use Quickstart to Configure a Stretched Cluster or Two-Node Cluster

You can use the Quickstart workflow to quickly configure a stretched cluster or two-node cluster.

When you create a cluster in the vSphere Client, the Quickstart workflow appears. You can use Quickstart to perform basic configuration tasks, such as adding hosts and claiming disks.

Prerequisites

- Deploy a host outside of any cluster to use as a witness host.
- Verify that hosts are running ESXi 6.0 Update 2 or later. For a two-node cluster, verify that hosts are running ESXi 6.1 or later.
- Verify that ESXi hosts in the cluster do not have any existing vSAN or networking configuration.

Procedure

1. Navigate to the cluster in the vSphere Client.
2. Click the Configure tab, and select **Configuration > Quickstart**.

3 On the Cluster basics card, click **Edit** to open the Cluster basics wizard.

   a Enter the cluster name.

   b Enable the vSAN slider.

      You also can enable other features, such as DRS or vSphere HA.

   c Click **Finish**.

4 On the Add hosts card, click **Add** to open the Add hosts wizard.

   a On the Add hosts page, enter information for new hosts, or click Existing hosts and select from hosts listed in the inventory.

   b On the Host summary page, verify the host settings.

   c On the Ready to complete page, click **Finish**.

5 On the Cluster configuration card, click **Configure** to open the Cluster configuration wizard.

   a On the Configure the distributed switches page, enter networking settings, including distributed switches, port groups, and physical adapters.

      - In the **Distributed switches** section, enter the number of distributed switches to configure from the drop-down menu. Enter a name for each distributed switch. Click **Use Existing** to select an existing distributed switch.

        If the physical adapters chosen are attached to a standard virtual switch with the same name across hosts, the standard switch is migrated to the distributed switch. If the physical adapters chosen are unused, the standard switch is migrated to the distributed switch.

        Network resource control is enabled and set to version 3. Distributed switches with network resource control version 2 cannot be used.

      - In the **Port Groups** section, select a distributed switch to use for vMotion and a distributed switch to use for the vSAN network.

      - In the **Physical adapters** section, select a distributed switch for each physical network adapter. You must assign each distributed switch to at least one physical adapter.

        This mapping of physical NICs to the distributed switches is applied to all hosts in the cluster. If you are using an existing distributed switch, the physical adapter selection can match the mapping of the distributed switch.

   b On the vMotion traffic page, enter IP address information for vMotion traffic.

   c On the Storage traffic page, enter IP address information for storage traffic.

   d On the Advanced options page, enter information for cluster settings, including DRS, HA, vSAN, host options, and EVC.

      In the **vSAN options** section, select Stretched cluster or Two node vSAN cluster as the **Deployment type**.

   e On the Claim disks page, select disks on each host for cache and capacity.

f (Optional) On the Proxy settings, page, configure the proxy server if your system uses one.

g On the Configure fault domains page, define fault domains for the hosts in the Preferred site and the Secondary site.

For more information about fault domains, see "Managing Fault Domains in vSAN Clusters" in *Administering VMware vSAN*.

h On the Select witness host page, select a host to use as a witness host. The witness host cannot be not part of the stretched cluster, and it can have only one VMkernel adapter configured for vSAN data traffic.

Before you configure the witness host, verify that it is empty and does not contain any components. A two-node cluster can share a witness with other two-node clusters.

i On the Claim disks for witness host page, select disks on the witness host for cache and capacity.

j On the Review page, verify the cluster settings, and click **Finish**.

What to do next

You can manage the cluster through vCenter Server.

You can add hosts to the cluster and modify the configuration through Quickstart. You also can modify the configuration manually with the vSphere Client.

Manually Configure vSAN Stretched Cluster

Configure a vSAN cluster that stretches across two geographic locations or sites.

Prerequisites

- Verify that you have a minimum of three hosts: one for the preferred site, one for the secondary site, and one host to act as a witness.
- Verify that you have configured one host to serve as the witness host for the stretched cluster. Verify that the witness host is not part of the vSAN cluster, and that it has only one VMkernel adapter configured for vSAN data traffic.
- Verify that the witness host is empty and does not contain any components. To configure an existing vSAN host as a witness host, first evacuate all data from the host and delete the disk group.

Procedure

1. Navigate to the vSAN cluster.
2. Click the **Configure** tab.
3. Under vSAN, click **Fault Domains**.
4. Click **Configure Stretched Cluster** to open the stretched cluster configuration wizard.

5 Select the hosts that you want to assign to the secondary fault domain and click **>>**.

The hosts that are listed under the Preferred fault domain are in the preferred site.

6 Click **Next**.

7 Select a witness host that is not a member of the vSAN stretched cluster and click **Next**.

8 Claim storage devices on the witness host and click **Next**.

Claim storage devices on the witness host. Select one flash device for the cache tier, and one or more devices for the capacity tier.

9 On the **Ready to complete** page, review the configuration and click **Finish**.

Change the Preferred Fault Domain

You can configure the Secondary site as the Preferred site. The current Preferred site becomes the Secondary site.

Note Objects with **Data locality=Preferred** policy setting always move to the Preferred fault domain. Objects with **Data locality=Secondary** always move to the Secondary fault domain. If you change the Preferred domain to Secondary, and the Secondary domain to Preferred, these objects move from one site to the other. This action might cause an increase in resynchronization activity. To avoid unnecessary resynchronization, you can change the Data locality setting to **None** before you swap the Preferred and Secondary domains. Once you swap the domains back, you can reset the Data locality.

Procedure

1 Navigate to the vSAN cluster.

2 Click the **Configure** tab.

3 Under vSAN, click **Fault Domains**.

4 Select the secondary fault domain and click the **Change Preferred Fault Domain** icon.

5 Click **Yes** or **Apply** to confirm.

The selected fault domain is marked as the preferred fault domain.

Change the Witness Host

You can replace or change the witness host for a vSAN stretched cluster.

Change the ESXi host used as a witness host for your vSAN stretched cluster.

Prerequisites

Verify that the witness host is not in use by another cluster, has a VMkernel configured for vSAN traffic, and has no vSAN partitions on its disks.

Procedure

1 Navigate to the vSAN cluster.

2 Click the **Configure** tab.

3 Under vSAN, click **Fault Domains**.

4 Click the **Change** button. The Change Witness Host wizard opens.

5 Select a new host to use as a witness host, and click **Next**.

6 Claim disks on the new witness host, and click **Next**.

7 On the Ready to complete page, review the configuration, and click **Finish**.

Deploying a vSAN Witness Appliance

Specific vSAN configurations, such as a stretched cluster, require a witness host. Instead of using a dedicated physical ESXi host as a witness host, you can deploy the vSAN witness appliance. The appliance is a preconfigured virtual machine that runs ESXi and is distributed as an OVA file.

Unlike a general purpose ESXi host, the witness appliance does not run virtual machines. Its only purpose is to serve as a vSAN witness.

The workflow to deploy and configure the vSAN witness appliance includes this process.

When you deploy the vSAN witness appliance, you must configure the size of the witness supported by the vSAN stretched cluster. Choose one of the following options:

- Tiny supports up to 750 components (10 VMs or fewer).
- Medium supports up to 21,833 components (500 VMs). As a shared witness, the Medium witness appliance supports up to 21,000 components and up to 21 vSAN two-node clusters.
- Large supports up to 64,000 components (more than 500 VMs). As a shared witness, the Large witness appliance supports up to 24,000 components and up to 24 vSAN two-node clusters.
- Extra Large supports up to 64,000 components (more than 500 VMs). As a shared witness, the Extra Large witness appliance supports up to 64,000 components and up to 64 vSAN two-node clusters.

Note These estimates are based on standard VM configurations. The number of components that make up a VM can vary, depending on the number of virtual disks, policy settings, snapshot requirements, and so on. For more information about witness appliance sizing for vSAN two-node clusters, refer to the *vSAN 2 Node Guide*.

You also must select a datastore for the vSAN witness appliance. The witness appliance must use a different datastore than the vSAN stretched cluster datastore.

1 Download the appliance from the VMware website.

2 Deploy the appliance to a vSAN host or cluster. For more information, see Deploying OVF Templates in the *vSphere Virtual Machine Administration* documentation.

3 Configure the vSAN network on the witness appliance.

4 Configure the management network on the witness appliance.

5 Add the appliance to vCenter Server as a witness ESXi host. Make sure to configure the vSAN VMkernel interface on the host.

Set Up the vSAN Network on the Witness Appliance

The vSAN witness appliance includes two preconfigured network adapters. You must change the configuration of the second adapter so that the appliance can connect to the vSAN network.

Procedure

1 Navigate to the virtual appliance that contains the witness host.

2 Right-click the appliance and select **Edit Settings**.

3 On the **Virtual Hardware** tab, expand the second Network adapter.

4 From the drop-down menu, select the vSAN port group and click **OK**.

Configure Management Network on the Witness Appliance

Configure the witness appliance, so that it is reachable on the network.

By default, the appliance can automatically obtain networking parameters if your network includes a DHCP server. If not, you must configure appropriate settings.

Procedure

1 Power on your witness appliance and open its console.

Because your appliance is an ESXi host, you see the Direct Console User Interface (DCUI).

2 Press F2 and navigate to the Network Adapters page.

3 On the Network Adapters page, verify that at least one vmnic is selected for transport.

4 Configure the IPv4 parameters for the management network.

a Navigate to the IPv4 Configuration section and change the default DHCP setting to static.

b Enter the following settings:

- IP address
- Subnet mask
- Default gateway

5 Configure DNS parameters.

- Primary DNS server

- Alternate DNS server
- Hostname

Configure Network Interface for Witness Traffic

You can separate data traffic from witness traffic in two-node vSAN clusters and stretched clusters.

vSAN data traffic requires a low-latency, high-bandwidth link. Witness traffic can use a high-latency, low-bandwidth and routable link. To separate data traffic from witness traffic, you can configure a dedicated VMkernel network adapter for vSAN witness traffic.

You can add support for a direct network cross-connection to carry vSAN data traffic in a vSAN stretched cluster. You can configure a separate network connection for witness traffic. On each data host in the cluster, configure the management VMkernel network adapter to also carry witness traffic. Do not configure the witness traffic type on the witness host.

Note Network Address Translation (NAT) is not supported between vSAN data hosts and the witness host.

Prerequisites

- Verify that the data site to witness traffic connection has a minimum bandwidth of 2 Mbps for every 1,000 vSAN components.
- Verify the latency requirements:
  - Two-node vSAN clusters must have less than 500 ms RTT.
  - Stretched clusters with less than 11 hosts per site must have less than 200 ms RTT.
  - Stretched clusters with 11 or more hosts per site must have less than 100 ms RTT.
- Verify that the vSAN data connection meets the following requirements.
  - For hosts directly connected in a two-node vSAN cluster, use a 10 Gbps direct connection between hosts. Hybrid clusters also can use a 1 Gbps crossover connection between hosts.
  - For hosts connected to a switched infrastructure, use a 10 Gbps shared connection (required for all-flash clusters), or a 1 Gbps dedicated connection.
- Verify that data traffic and witness traffic use the same IP version.

Procedure

1. Open an SSH connection to the ESXi host.

2 Use the `esxcli network ip interface list` command to determine which VMkernel network adapter is used for management traffic.

For example:

```
esxcli network  ip interface list
vmk0
   Name: vmk0
   MAC Address: e4:11:5b:11:8c:16
   Enabled: true
   Portset: vSwitch0
   Portgroup: Management Network
   Netstack Instance: defaultTcpipStack
   VDS Name: N/A
   VDS UUID: N/A
   VDS Port: N/A
   VDS Connection: -1
   Opaque Network ID: N/A
   Opaque Network Type: N/A
   External ID: N/A
   MTU: 1500
   TSO MSS: 65535
   Port ID: 33554437

vmk1
   Name: vmk1
   MAC Address: 00:50:56:6a:3a:74
   Enabled: true
   Portset: vSwitch1
   Portgroup: vsandata
   Netstack Instance: defaultTcpipStack
  VDS Name: N/A
   VDS UUID: N/A
   VDS Port: N/A
   VDS Connection: -1
   Opaque Network ID: N/A
   Opaque Network Type: N/A
   External ID: N/A
   MTU: 9000
   TSO MSS: 65535
   Port ID: 50331660
```

Note Multicast information is included for backward compatibility. vSAN 6.6 and later releases do not require multicast.

3 Use the `esxcli vsan network ip add` command to configure the management VMkernel network adapter to support witness traffic.

```
esxcli vsan network ip add -i vmk5 -T witness
```

4 Use the `esxcli vsan network list` command to verify the new network configuration.

For example:

```
esxcli vsan network list
Interface
   VmkNic Name: vmk0
   IP Protocol: IP
   Interface UUID: 8cf3ec57-c9ea-148b-56e1-a0369f56dcc0
   Agent Group Multicast Address: 224.2.3.4
   Agent Group IPv6 Multicast Address: ff19::2:3:4
   Agent Group Multicast Port: 23451
   Master Group Multicast Address: 224.1.2.3
   Master Group IPv6 Multicast Address: ff19::1:2:3
   Master Group Multicast Port: 12345
   Host Unicast Channel Bound Port: 12321
   Multicast TTL: 5
   Traffic Type: witness

Interface
   VmkNic Name: vmk1
   IP Protocol: IP
   Interface UUID: 6df3ec57-4fb6-5722-da3d-a0369f56dcc0
   Agent Group Multicast Address: 224.2.3.4
   Agent Group IPv6 Multicast Address: ff19::2:3:4
   Agent Group Multicast Port: 23451
   Master Group Multicast Address: 224.1.2.3
   Master Group IPv6 Multicast Address: ff19::1:2:3
   Master Group Multicast Port: 12345
   Host Unicast Channel Bound Port: 12321
   Multicast TTL: 5
   Traffic Type: vsan
```

Results

In the vSphere Client, the management VMkernel network interface is not selected for vSAN traffic. Do not re-enable the interface in the vSphere Client.

Convert a Stretched Cluster to a Standard vSAN Cluster

You can decommission a stretched cluster and convert it to a standard vSAN cluster.

When you disable a stretched cluster, the witness host is removed, but the fault domain configuration remains. Because the witness host is not available, all witness components are missing for your virtual machines. To ensure full availability for your VMs, repair the cluster objects immediately.

Prerequisites

- Back up all running VMs, and verify that all VMs are compliant with their current storage policy.
- Ensure that no health issues exist, and that all resync activities are complete.

- Change the associated storage policy to move all VM objects to one site. Use the Data locality rule to restrict virtual machine objects to the selected site.

Procedure

1. Navigate to the vSAN stretched cluster.
2. Click the **Configure** tab.
3. Under vSAN, click **Fault Domains**.
4. Disable the stretched cluster.
   a. Click **Disable**. The Remove Witness Host dialog opens.
   b. Click **Remove** to confirm.
5. Remove the fault domain configuration.
   a. Select a fault domain, and choose menu **Actions > Delete**. Click **Yes** to confirm.
   b. Select the other fault domain, and choose menu **Actions > Delete**. Click **Yes** to confirm.
6. Remove the witness host from inventory.
7. Repair the objects in the cluster.
   a. Click the **Monitor** tab.
   b. Under vSAN, click **Health** and click **vSAN object health**.
   c. Click **Repair object immediately**.

   vSAN recreates the witness components within the cluster.

Assign Two-Node Clusters to a Shared Witness Host

You can quickly assign two-node vSAN clusters to a shared witness host.

When you configure a two-node cluster, you can select a witness host. The witness host can be shared by up to 64 two-node clusters. You also can assign multiple two-node clusters to a shared witness host.

Procedure

1. Right-click a host in the vSphere Client navigator.
2. Select menu **vSAN > Assign as shared witness host**.
3. In the Assign clusters to witness dialog box, select two-node clusters to assign to this witness host.
4. Click **Add**.

Results

The selected clusters are assigned to use this witness host.

Reassign Shared Witness Host for Two-Node Clusters

You can quickly reassign a new shared witness host for two-node vSAN clusters.

When you configure a two-node cluster, you can select a witness host. The witness host can be shared by up to 64 two-node clusters. You also can reassign the cluster to a different shared witness host.

Procedure

1. Right-click a two-node vSAN cluster in the vSphere Client navigator.
2. Select menu **vSAN > Assign Shared Witness**.
3. On the Reassign to another witness host dialog box, select a witness host from the drop-down menu.
4. (optional) Click Validate Cluster Compatibility.
5. Claim disks for the new witness host.

   If the witness host is already assigned to one or more two-node clusters, this page does not appear.
6. Review the configuration and click **Finish**.

Results

The two-node cluster is reassigned to use the selected witness host.