Administering VMware Virtual SAN

Virtual SAN 6.1

This document supports the version of each product listed and supports all subsequent versions until the document is replaced by a new edition. To check for more recent editions of this document, see http://www.vmware.com/support/pubs.

EN-001876-01

vmware®

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

http://www.vmware.com/support/

The VMware Web site also provides the latest product updates.

If you have comments about this documentation, submit your feedback to:

docfeedback@vmware.com

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

Contents

About VMware Virtual SAN

Administering VMware Virtual SAN describes how to configure, manage, and monitor a VMware Virtual SAN cluster from the VMware vSphere envrionment. In addition, *VMware Virtual SAN* explains how to organize the local physical storage resources that serve as storage capacity devices in a Virtual SAN cluster, define storage policies for virtual machines deployed to Virtual SAN datastores, and manage failures in a Virtual SAN cluster.

Intended Audience

This information is for experienced virtualization administrators who are familiar with virtualization technology, day-to-day data center operations, and Virtual SAN concepts.

Updated Information

This *Administering VMware Virtual SAN* is updated with each release of the product or when necessary.

This table provides the update history of *Administering VMware Virtual SAN*.

Revision	Description
EN-001876-01	■ In the topic "Best Practices for Working with Stretched Clusters," on page 55, information was added to describe how VMs in stretched clusters might be force provisioned and become non-compliant while one of the sites is unavailable. ■ The topic "Stretched Cluster Design Considerations," on page 54 contains an additional note to indicate that stretched clusters require on-disk format 2.0 or later. ■ The topic "Define a Virtual Machine Storage Policy for Virtual SAN," on page 91 now includes the vCenter privileges required to define a Virtual SAN storage policy.
EN-001876-00	Initial release.

Introduction to Virtual SAN 1

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

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

This chapter includes the following topics:

Virtual SAN Concepts

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

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

You can activate Virtual SAN when you create host clusters and enable Virtual SAN on your existing clusters. When enabled, Virtual SAN aggregates all local capacity devices available on the hosts into a single datastore shared by all hosts in the Virtual SAN cluster. You can later expand the datastore by adding capacity devices or hosts with capacity devices to the cluster. VMware recommends that the ESXi hosts in the cluster share similar or identical configurations across all cluster member, including similar or identical storage configurations. This ensures balanced virtual machine storage components across all devices and hosts in the cluster. However, the hosts without any local devices can participate and run their virtual machines on the Virtual SAN datastore.

If a host contributes its local capacity devices to the Virtual SAN datastore, the host must provide at least one capacity device, also called a data disk, and at least one device for flash cache.

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 Virtual SAN cluster, see the *VMware Virtual SAN 6.0 Design and Sizing Guide*.

Characteristics of Virtual SAN

This topic summarizes characteristics that apply to Virtual SAN, as well as its clusters and datastores.

Virtual SAN provides numerous benefits to your environment.

Table 1-1. Virtual SAN Features

Supported Features	Description
Virtual SAN health service	Virtual SAN health service includes preconfigured health check tests to monitor, troubleshoot, diagnose the cause of cluster component problems, and identify any potential risk.
Shared storage support	Virtual SAN 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.
Just a Bunch Of Disks (JBOD)	Virtual SAN supports JBOD for use in a blade server environment. If your cluster contains blade servers, you can extend the capacity of the datastore with JBOD storage that is connected to the blade servers.
On-disk format	Virtual SAN 6.0 supports a new on-disk virtual file format 2.0 based on Virsto technology, which is a log-based filesystem that provides highly scalable snapshot and clone management support per Virtual SAN cluster. For information about the number of virtual machine snapshots and clones supported per Virtual SAN cluster, see the *Configuration Maximums* documentation.
All-flash and hybrid configurations	Virtual SAN can be configured for all-flash or hybrid cluster.
Fault domains	Virtual SAN supports configuring fault domains to protect hosts from rack or chassis failure when the Virtual SAN cluster spans across multiple racks or blade server chassis in a data center.
Stretched cluster	Virtual SAN supports stretched clusters that span across two geographic locations.
Integration with vSphere storage features	Virtual SAN integrates with vSphere data management features traditionally used with VMFS and NFS storage. These features include snapshots, linked clones, vSphere Replication, and vSphere APIs for Data Protection.
Virtual Machine Storage Policies	Virtual SAN works with virtual machine storage policies to support a virtual machine-centric storage approach. When provisioning a virtual machine, if there is no explicit assignment of a storage policy to the virtual machine, a generic system defined storage policy, called the Virtual SAN Default Storage Policy is automatically applied to the virtual machine.
Rapid provisioning	Virtual SAN enables rapid provisioning of storage in the vCenter Server during virtual machine creation and deployment operations.

Virtual SAN Terms and Definitions

Virtual SAN introduces specific terms and definitions to refer to different features.

Before you get started with Virtual SAN, review the following key Virtual SAN 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 Virtual SAN cluster. On each ESXi host that contributes its local devices to a Virtual SAN 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, or for other uses. A single caching device must be dedicated to a single disk group. In hybrid clusters, magnetic disks are used for the storage capacity layer and flash devices for the cache layer. In an all-flash cluster, flash devices are used for both cache and capacity. For information about creating and managing disk groups, see Chapter 8, "Device Management in a Virtual SAN Cluster," on page 67.

Consumed capacity

Amount of physical capacity consumed by one or more virtual machines at any point. Consumed capacity is determined by many factors, 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

Virtual SAN 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 Virtual SAN datastore, Virtual SAN 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, Virtual SAN provisions and manages each object, individually, which might also involve creating of a RAID configuration for every object.

When creating an object for a virtual disk and determining how to distribute the object in the cluster, Virtual SAN considers the following parameters:

- Verifies that the virtual disk requirements are applied according to the specified virtual machine storage policy settings.
- Verifies that the correct cluster resources are utilized at the time of provisioning. For example, based on the protection policy, Virtual SAN 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.
- Virtual SAN continually monitors and reports the policy compliance status of the virtual disk. If you find any incompliant 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. Virtual SAN actively throttles the storage and network throughput used for reconfiguration to minimize the impact of reconfiguration of objects to normal workload execution. When you change VM storage policy settings, Virtual SAN might initiate an object recreation process and subsequent resynchronization of the objects. See "About the Resynchronization Operation in the Virtual SAN Cluster," on page 94.

- 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 failure, Virtual SAN looks for ESXi hosts that satisfy the placement rules where protection components of virtual machine objects must be placed on two different hosts, instead on the same host, or across different fault domains, if specified.

Virtual SAN datastore

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

Following are the characteristics of a Virtual SAN datastore:

- Virtual SAN creates a single Virtual SAN datastore accessible to all hosts in the cluster, whether or not they have devices. All hosts can also mount any other datastores, including Virtual Volume, VMFS, or NFS.
- You can use Storage vMotion to move virtual machines between the Virtual SAN datastores, NFS 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.
- In automatic mode, a Virtual SAN datastore dynamically grows when you add hosts with capacity to a Virtual SAN cluster, or capacity devices to any cluster member.

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, when the Number of failure to tolerate policy configured to one, Virtual SAN ensures that the protection components, such as replicas and witnesses of the object are placed on separate hosts in the Virtual SAN 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, Virtual SAN also stripes the object across multiple capacity devices and each stripe is considered a component of the specified object. When needed, Virtual SAN might also break large objects into multiple components.

A Virtual SAN 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, snapshot delta description files, and so on.
VMDK	A virtual machine disk or `.vmdk` file that stores the contents of 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 not compliant when one of the mirror copies is inaccessible. If your virtual machines are in compliance with the requirement 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 Virtual SAN cluster, see "Handling Failures in Virtual SAN," on page 103.

Component State: Degraded and Absent states

Virtual SAN acknowledges the following failure states for components:

- Degraded. A component is in Degraded state when Virtual SAN detects a permanent component failure and assumes that the failed component will never recover to its original working state. As a result, Virtual SAN starts to rebuild the degraded components immediately. This state might occur when a component is on a failed device.
- Absent. A component is in Absent state when Virtual SAN detects a temporary component failure where components, including all its data, might recover and return Virtual SAN to its original state. This state might occur when you are restarting hosts or if you unplug a device from a Virtual SAN host. Virtual SAN starts to rebuild the components in absent status after waiting for 60 minutes. After waiting for 60 minutes, the component state is automatically changed to degraded.

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 a full mirror as well as more than 50 percent of an object's components (or votes) are still available, the operational status of the object is considered healthy.
- Unhealthy. If no copy of the mirror is available, or fewer than 50 percent of an object's components (or votes) are available, possibly due to multiple failures in the cluster, the operational status of the object is considered unhealthy and it impacts the availability of your virtual machine. For objects to remain accessible in the cluster, an error-free full replica of object components must be available at all times.

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 needs to 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 Virtual SAN datastore when using on-disk format 1.0 and 4 MB for on-disk format for version 2.0.

With 6.0, Virtual SAN supports a quorum-based system where each component might have more than one vote to decide the availability of virtual machines. To be more precise, 50 percent of the votes that make up a virtual machine's storage object must be accessible at all times. When fewer than 50 percent of the votes accessible to all hosts, the object is no longer available to the Virtual SAN datastore. This impacts the availability of your virtual machine. For objects to remain accessible in the cluster, an error-free full replica of object components must be available at all times.

Storage Policy-Based Management (SPBM)

When you use Virtual SAN, you can define virtual machine storage requirements, such as performance and availability, in the form of a policy. Virtual SAN ensures that the virtual machines deployed to Virtual SAN datastores are assigned at least one virtual machine storage policy. When you know the storage requirements of your virtual machines, you can create user-defined storage policies and assign the policies to your virtual machines. If you do not apply a storage policy when deploying virtual machines, Virtual SAN automatically assigns a default Virtual SAN policy with Number of failures to tolerate configured to one, a single disk stripe for each object, and thin provisioned virtual disk. For best results, you should create and use your own virtual machine storage policies, even if the requirements of the policy are the same as those defined in the default storage policy. For information about working with Virtual SAN storage policies, see Chapter 10, "Using Virtual SAN Policies," on page 87.

Ruby vSphere Console (RVC)

The Ruby vSphere Console (RVC) is a command-line interface tool used for managing and troubleshooting the Virtual SAN cluster. The tool provides a cluster-wide view, instead of the host-centric view offered by `esxcli`. Because the RVC tool is bundled with vCenter Server Appliance and vCenter Server for Windows, you do not need to separately install the tool. For information about the RVC commands, see the *RVC Command Reference Guide*.

Virtual SAN Observer

The VMware Virtual SAN Observer is a Web-based tool that runs on RVC and is used for in-depth performance analysis and monitoring of the Virtual SAN cluster. Use Virtual SAN Observer for information about the performance statistics of the capacity layer, detailed statistical information about physical disk groups, current CPU usage, consumption of Virtual SAN memory pools, and physical and in-memory object distribution across Virtual SAN clusters.

For information about setting up, launching, and using the RVC tool and the Virtual SAN Observer, see the *Virtual SAN Troubleshooting Reference Manual*.

Virtual SAN and Traditional Storage

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

Virtual SAN and traditional storage also differ in the following key ways:

- Virtual SAN 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. Virtual SAN 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.
- Virtual SAN has no concept of traditional storage volumes based on LUNs or NFS shares.
- The standard storage protocols, such as iSCSI, FCP, and so on, do not apply to Virtual SAN.
- Virtual SAN is highly integrated with vSphere. You do not need dedicated plug-ins or a storage console for Virtual SAN, compared to traditional storage. You can deploy, manage, and monitor Virtual SAN by using the vSphere Web Client.
- A dedicated storage administrator does not need to manage Virtual SAN. Instead a vSphere administrator can manage a Virtual SAN environment.
- With Virtual SAN usage, VM storage policies are automatically assigned when you deploy new VMs. The storage policies can be changed dynamically as needed.

Building a Virtual SAN Cluster

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

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

Virtual SAN Ready Node

The Virtual SAN Ready Node is a preconfigured solution of the Virtual SAN software for VMware partners, such as Cisco, Dell, Fujitsu, IBM, and Supermicro. This solution includes validated server configuration in a tested, certified hardware form factor for Virtual SAN deployment that is recommended by the server OEM and VMware. For information about the Virtual SAN Ready Node solution for a specific partner, see the VMware Partner web site at http://partnerweb.vmware.com/programs/vsan/Virtual SAN Ready Nodes.pdf .

User-Defined Virtual SAN Cluster

You can build a Virtual SAN cluster by selecting individual software and hardware components, such as drivers, firmware, and storage I/O controllers that are listed in the Virtual SAN Compatibility Guide (VCG) web site 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, and so on that are certified and listed on the VCG Web site. Review the compatibility information on the VCG Web site before choosing software and hardware components, drivers, firmware, and storage I/O controllers that are supported by Virtual SAN. When designing a Virtual SAN cluster, use only devices, firmware, and drivers that are listed on the VCG Web site. 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 Virtual SAN cluster, see Chapter 3, “Designing and Sizing a Virtual SAN Cluster,” on page 23.

VMware EVO:RAIL

The VMware EVO:RAIL software is a turnkey solution that combines compute, networking, and storage resources into a hyper-converged infrastructure appliance to provide an all-inclusive solution offered by VMware partners. EVO:RAIL software includes Virtual SAN, which is built in to the hardware appliances of VMware partners. EVO:RAIL creates a single Virtual SAN datastore from all local capacity devices on each ESXi host in a EVO:RAIL cluster. For information about the EVO:RAIL software, see the product Web site at http://www.vmware.com/files/pdf/products/evorail/vmware-evorail-introduction.pdf.

Integrating with Other VMware Software

After you have Virtual SAN 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 Virtual SAN on the same cluster. As with traditional datastores, vSphere HA provides the same level of protection for virtual machines on Virtual SAN datastores. This level of protection imposes specific restrictions when vSphere HA and Virtual SAN interact. For specific considerations about integrating vSphere HA and Virtual SAN, see “Using Virtual SAN and vSphere HA,” on page 51.

Integrating with VMware Horizon View

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

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

For information about using Virtual SAN with Horizon View 5.3.1 and configuring a Virtual SAN datastore to use Horizon View, see the VMware knowledge base article at kb.vmware.com/kb/2073795.

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

Limitations of Virtual SAN

This topic discusses the limitations of Virtual SAN.

When working with Virtual SAN, consider the following limitations:

- Virtual SAN does not support hosts participating in multiple Virtual SAN clusters. However, a Virtual SAN host can access other external storage resources, but at any point, can participate in one Virtual SAN cluster.
- Virtual SAN does not support vSphere DPM and Storage I/O Control.
- Virtual SAN does not support SCSI reservations.
- Virtual SAN does not support RDM, VMFS, diagnostic partition, and other device access features.

Requirements for Enabling Virtual SAN 2

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

This chapter includes the following topics:

Hardware Requirements for Virtual SAN

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

Storage Device Requirements

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

Table 2-1. Storage Device Requirements for Virtual SAN Hosts

Storage Component	Requirements
Cache	■ One SAS or SATA solid state disk (SSD) or PCIe flash device. ■ Before considering the number of failures to tolerate, make sure that in each disk group the size of the flash cache is at least 10 percent of the anticipated consumed capacity, without the protection copies. ■ vSphere Flash Read Cache must not use any of the flash devices reserved for Virtual SAN 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 at least one SAS, NL-SAS or SATA magnetic disk. ■ 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 or RAID 0 mode.

Memory

The memory requirements for Virtual SAN depend on the number of disk groups and devices that are managed by the ESXi hypervisor. Each host should contain a minimum of 32 GB of memory to accommodate for the maximum number of 5 disk groups and maximum number of 7 capacity devices per disk group.

Flash Boot Devices

When you boot a Virtual SAN 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 is greater than 512 GB, boot the host from a SATADOM or disk device. When you boot a Virtual SAN host from a SATADOM device, you must use single-level cell (SLC) device and the size of the boot device must be at least 16 GB.

When you boot an ESXi 6.0 host from USB device or from SD card, Virtual SAN 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 of handling Virtual SAN traces when booting an ESXi from USB stick or SD card. Note that if a power failure occurs, Virtual SAN trace logs are not preserved.

When you boot an ESXi 6.0 host from a SATADOM device, Virtual SAN 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 Virtual SAN

Verify that a cluster meets the requirements for enabling Virtual SAN.

- All capacity devices, drivers, and firmware versions in your Virtual SAN configuration must be certified and listed in the Virtual SAN section of the *VMware Compatibility Guide*.
- A Virtual SAN cluster must have a minimum of three hosts that contribute capacity to the cluster. For information about the considerations for a three-host cluster, see "Design Considerations for a Virtual SAN Cluster," on page 30.
- A host must not participate in other clusters beside the Virtual SAN cluster.

Software Requirements for Virtual SAN

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

To use the full set of Virtual SAN capabilities, the ESXi hosts that participate in Virtual SAN Clusters must be version 6.0. During the Virtual SAN upgrade from version 5.5 to version 6.0, you can keep the on-disk format version 1.0, but you cannot use many of the new features. Virtual SAN 6.0 supports both the on-disk formats.

Networking Requirements for Virtual SAN

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

Table 2-2. Networking Requirements for Virtual SAN

Networking Component	Requirement
Host Bandwidth	Each host must have minimum bandwidth dedicated to Virtual SAN. ■ Dedicated 1 Gbps for hybrid configurations ■ Dedicated or shared 10 Gbps for all-flash configurations For information about networking considerations in Virtual SAN, see "Designing the Virtual SAN Network," on page 31.
Connection between hosts	Each host in the Virtual SAN cluster, regardless of whether it contributes capacity, must have a VMkernel network adapter for Virtual SAN traffic type. See "Set Up a VMkernel Network for Virtual SAN," on page 47.
Host network	All hosts in your Virtual SAN cluster must be connected to a Virtual SAN Layer 2 network.
Multicast	Multicast must be enabled on the physical switches and routers that handle Virtual SAN traffic along the Layer 2 path and optionally the Layer 3 path.
IPv4 and IPv6 support	The Virtual SAN network must be IPv4-only. Virtual SAN does not support IPv6.

License Requirements

Verify that you have a valid license for Virtual SAN.

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

You can assign different types of licenses to support different Virtual SAN features, such as all-flash configuration and stretched clusters. Make sure your license supports the features you intend to use. For information about assigning licenses, see "Assign a License to a Virtual SAN Cluster," on page 50.

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

Designing and Sizing a Virtual SAN Cluster 3

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

The *Administering VMware Virtual SAN* documentation examines the key points about designing and sizing a Virtual SAN cluster. For detailed instructions on designing and sizing a Virtual SAN cluster, see *VMware Virtual SAN 6.0 Design and Sizing Guide*.

This chapter includes the following topics:

Designing and Sizing Virtual SAN Storage Components

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

- Design Considerations for Storage Controllers in Virtual SAN on page 28

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

Planning Capacity in Virtual SAN

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

Raw Capacity

You determine the capacity of a Virtual SAN datastore by aggregating the disk groups on the hosts in the cluster and the size of the capacity devices in the groups, less the overhead from the Virtual SAN on-disk format.

Number of Failures to Tolerate

When you plan the capacity of the Virtual SAN datastore, the number of virtual machines and the size of their VMDK files aside, you must consider the number of failures to tolerate attribute of the virtual machine storage policies for the cluster.

The number of failures to tolerate has an important role when you plan and size storage capacity for Virtual SAN. 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 number of failures to tolerate is one, virtual machines can use about 50 percent of the raw capacity. If the number of failures to tolerate is two, the usable capacity is about 33 percent. If the number of failures to tolerate is equal to the maximum of three, the usable capacity is about 25 percent.

For information about the attributes in a Virtual SAN storage policy, see Chapter 10, “Using Virtual SAN Policies,” on page 87.

Calculating Required Capacity

You plan the basic capacity for the virtual machines in the cluster based on the following criteria:

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

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

2 Consider the number of failures that is configured in the storage policies for the virtual machines in the cluster. The number of failures to tolerate attribute in a VM storage policy directly impacts the number of replicas of a VMDK file on the hosts in the cluster.

```
datastore capacity = expected overall consumption  * (number of failures to tolerate + 1)
```

3 Calculate the overhead requirement of the Virtual SAN on-disk format.

- If you use the new on-disk format, Virtual SAN adds an overhead of 1 percent capacity per device.
- If you preserve version 1.0 of the on-disk format that is included with Virtual SAN 5.5 after you upgrade to Virtual SAN 6.0, Virtual SAN consumes about 1 GB capacity per physical capacity.

Capacity Sizing Guidelines

- Leave at least 30 percent unused space to prevent Virtual SAN from rebalancing the storage load. Virtual SAN rebalances the components across the cluster if the consumption on a physical capacity device becomes greater than 80 percent. The rebalance operation might impact the performance of applications. To avoid these issues, keep storage consumption to less than 70 percent.

- Plan extra capacity to handle potential failures or replacements of capacity devices, disk groups, and hosts. When a capacity device is not reachable, Virtual SAN recovers the components on another device in the cluster. When a cache flash device fails or is removed, Virtual SAN recovers the components from the entire disk group.
- Reserve extra capacity to make sure that Virtual SAN 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 successfully rebuild after a host failure or during maintenance. This is important when working in an environment that has more than three hosts, where you must leave sufficient free capacity for rebuilding the failed components and 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, Virtual SAN will not perform the rebuild operation if the Number of Failures to Tolerate is configured to one because when one host fails, there will only be two hosts left in the cluster. To tolerate a failure, you must have at least three hosts.
- Provide enough temporary storage space for changes in the Virtual SAN VM storage policy. When you dynamically change a VM storage policy, Virtual SAN might create a layout of the replicas that make up an object. When Virtual SAN instantiates and synchronizes those replicas with the original replica, the cluster must temporarily provide additional space.

Considerations for Virtual Machine Objects

When you plan the storage capacity in the Virtual SAN datastore, consider the space the VM home namespace objects, snapshots, and swap files require in the datastore.

- 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 flash storage, Virtual SAN applies only the number of 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 number of failures to tolerate is greater than zero.
- Snapshots. Delta devices inherit the policy of the base VMDK file. Plan additional space according to the expected size and number of snapshots, and to the settings in the Virtual SAN 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. Virtual SAN uses an individual storage policy for the swap files of virtual machines. The policy tolerates a single failure, defines no striping and read cache reservation, and enables thin provisioning.

Design Considerations for Flash Caching Devices in Virtual SAN

Plan the configuration of flash devices for Virtual SAN 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 PCIe or SSD flash devices according to the requirements for performance, capacity, write endurance, and cost of the Virtual SAN storage.

- Compatibility. The model of the PCIe or SSD devices must be listed in the Virtual SAN section of the *VMware Compatibility Guide*.
- Performance. PCIe devices have generally faster performance than SSD devices.
- Capacity. The maximum capacity that is available for PCIe devices is generally greater than the maximum capacity that is currently listed for SSD devices for Virtual SAN in the *VMware Compatibility Guide*.

- Write endurance. The write endurance of the PCIe or 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 Virtual SAN 6.0 Design and Sizing Guide*. For information about the write endurance class of PCe and SSD devices, see the Virtual SAN section of the *VMware Compatibility Guide*.

- Cost. PCIe devices have generally higher cost than SSD devices.

Flash Devices as Virtual SAN Cache

Model the configuration of flash cache for Virtual SAN for write endurance, performance, and potential growth based on these considerations.

Table 3-1. Sizing Virtual SAN Cache

Storage Configuration	Considerations
All-flash and hybrid configurations	■ The size of the cache must be at least 10 percent of the consumed capacity that virtual machines are expected to consume, without the protection copies. The number of failures to tolerate attribute from the VM storage policy does not impact the size of the cache. ■ 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. The 10 percent of the consumed capacity does not include protection copies. ■ A higher cache-to-capacity ratio eases future capacity growth. Oversizing cache enables you to easily 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. ■ When a flash caching device is at the end of its life, replacing it 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 failure domain if a flash caching device fails because fewer capacity devices are affected ■ Potentially improved performance if you deploy multiple groups that consist of 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, Virtual SAN uses the cache layer for write caching only. The write cache must be able to handle very high write activities. This approach extends the life of capacity flash that might be less expensive and might have lower write endurance.
Hybrid configurations	If the read cache reservation is configured in the active VM storage policy for performance reasons, the hosts in the Virtual SAN 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.

Design Considerations for Flash Capacity Devices in Virtual SAN

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

Choosing Between PCIe or SSD Flash Devices

Choose PCIe or SSD flash devices according to the requirements for performance, capacity, write endurance, and cost of the Virtual SAN storage.

- Compatibility. The model of the PCIe or SSD devices must be listed in the Virtual SAN section of the *VMware Compatibility Guide*.
- Performance. PCIe devices have generally faster performance than SSD devices.
- Capacity. The maximum capacity that is available for PCIe devices is generally greater than the maximum capacity that is currently listed for SSD devices for Virtual SAN in the *VMware Compatibility Guide*.
- Write endurance. The write endurance of the PCIe or 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 Virtual SAN 6.0 Design and Sizing Guide*. For information about the write endurance class of PCe and SSD devices, see the Virtual SAN section of the *VMware Compatibility Guide*.
- Cost. PCIe devices have generally higher cost than SSD devices.

Flash Devices as Virtual SAN Capacity

In all- flash configurations, Virtual SAN 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 Virtual SAN, 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 Virtual SAN

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

SAS, NL-SAS, and SATA Magnetic Devices

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

- Compatibility. The model of the magnetic disk must be certified and listed in the Virtual SAN section of the *VMware Compatibility Guide*.
- Performance. SAS and NL-SAS devices have faster performance than SATA disks.
- Capacity. The capacity of SAS, NL-SAS, and SATA magnetic disks for Virtual SAN is available in the Virtual SAN 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 are more expensive than SATA disks.

Using SATA disks instead of SAS and NL-SAS devices is justifiable in environments where capacity and reduced cost have higher priority than performance.

Magnetic Disks as Virtual SAN Capacity

Plan a magnetic disk configuration by following these guidelines:

- For better performance of Virtual SAN, 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 Virtual SAN 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 Virtual SAN datastore.
- Dedicate a high enough number of magnetic disks to satisfy the value of the number 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 Virtual SAN, see Chapter 10, "Using Virtual SAN Policies," on page 87.

Design Considerations for Storage Controllers in Virtual SAN

Include storage controllers on the hosts of a Virtual SAN 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 Virtual SAN in the *VMware Compatibility Guide*.
- Use multiple storage controllers, if possible, to improve performance and to isolate a potential controller failure only to 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 Virtual SAN is rebuilding components after a failure or when a host enters maintenance mode.
- Consider that storage controllers in RAID 0 mode require higher configuration and maintenance efforts compared to storage controllers in passthrough mode. Use storage controllers in passthrough mode for best performance of Virtual SAN.

Designing and Sizing Virtual SAN Hosts

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

Memory and CPU

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

Table 3-2. Sizing Memory and CPU of Virtual SAN 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 Virtual SAN with 5 disk groups per host and 7 capacity devices per disk group Hosts that have less than 512 GB memory 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 disk.
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 Virtual SAN

Host Networking

Provide more bandwidth for Virtual SAN traffic for improved performance.

- If you plan to use hosts that have 1-GbE adapters, dedicate adapters for Virtual SAN 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 Virtual SAN traffic to isolate the traffic by using Network I/O Control and VLANs.
- Create a team of physical adapters for Virtual SAN traffic for redundancy.

Multiple Disk Groups

A disk group represents a single failure domain in the Virtual SAN datastore. The capacity of the disk group becomes inaccessible if the flash cache or storage controller stops responding. As a result, Virtual SAN rebuilds all components from the disk group on another location in the cluster.

Design multiple disk groups with less capacity for the following benefits and disadvantages:

- Benefits
  - Improved performance because the datastore has more aggregated cache and I/O operations are faster
  - Greater number and size of the failure domains, and improved performance in the case of disk group failure because Virtual SAN rebuilds fewer components
- Disadvantages
  - High cost because you use two caching devices for the same cache size instead of one
  - Requirements for more memory to handle more disk groups
  - Requirements for multiple storage controllers to reduce the failure domains

Drive Bays

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

Blade Servers and External Storage

The capacity of blade servers usually does not scale in a Virtual SAN datastore because they have a limited number of disk slots. To extend the planned capacity of blade servers, use external storage enclosures. For information about the supported models of external storage enclosures, see *VMware Compatibility Guide*.

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 to make the host discover the new drive.

Design Considerations for a Virtual SAN Cluster

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

Sizing the Virtual SAN Cluster for Failures to Tolerate

You configure the Number of failures to tolerate attribute in the VM storage policies to handle host failures. The number of hosts required for the cluster is equal to `2 * number of failures to tolerate + 1`. The more failures the cluster tolerates, the more capacity hosts are required.

You can also organize the hosts from the cluster in fault domains to improve failure management if the hosts are connected in rack servers. See "Designing and Sizing Virtual SAN Fault Domains," on page 33.

Limitations in a 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. Virtual SAN saves each of the two required replicas of virtual machine data on separate hosts. The witness object is on the third host. Because of the small number of hosts in the cluster, the following limitations exist:

- When a host fails, Virtual SAN cannot rebuild data on another host to protect against another failure.
- If a host must enter maintenance mode, Virtual SAN cannot reprotect evacuated data. Data is exposed to a potential failure while the host is in maintenance mode.

  You can use only the ensure accessibility option. The full data migration option is not available because the cluster does not have a spare host that it can use for evacuating data.

  As a result, virtual machines are at risk because they become inaccessible if another failure occurs.

Balanced and Unbalanced Cluster Configuration

Virtual SAN works best on hosts with uniform configurations.

Adding hosts with different configurations to the Virtual SAN cluster has the following disadvantages:

- Reduced predictability of storage performance because Virtual SAN 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 Virtual SAN

If you deploy vCenter Server on the Virtual SAN datastore, you might not be able to use vCenter Server for troubleshooting, if a problem occurs in the Virtual SAN cluster.

Designing the Virtual SAN Network

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

For details about the Virtual SAN network configuration, see the *VMware Virtual SAN 6.0 Design and Sizing Guide* and *Virtual SAN 6.0 Network Design Guide*.

Networking Failover and Load Balancing

Virtual SAN uses the teaming and failover policy that is configured on the backing virtual switch for network redundancy only. Virtual SAN 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 standard switch and LACP port channel for distributed switch
Route based on physical network adapter load	Active/Active with LACP port channel for distributed switch

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

Virtual SAN does not support multiple VMkernel adapters on the same subnet. You can use multiple VMkernel adapters on different subnets, such as another VLAN or separate physical fabric. Providing availability by using several VMkernel adapters has configuration costs including vSphere and the network infrastructure. Network availability by teaming physical network adapters is easier to achieve with less setup.

Multicast Considerations in Virtual SAN Network

Multicast must be enabled on the physical switches to enable heartbeat and exchange of metadata between the hosts in the Virtual SAN cluster. You can configure an IGMP snooping querier on the physical switches for delivery of multicast messages only through the physical switch ports that are connected to the Virtual SAN host network adapters. In the case of multiple Virtual SAN clusters on the same network, before you deploy an additional Virtual SAN cluster in production, change the multicast address for the new cluster so that the member hosts do not receive unrelated multicast messages from another cluster. For information about assigning a multicast address for a Virtual SAN cluster, see "Change the Multicast Address for a Virtual SAN Cluster," on page 42.

Allocating Bandwidth for Virtual SAN by Using Network I/O Control

If Virtual SAN traffic uses 10-GbE physical network adapters that are shared with other system traffic types, such as vSphere vMotion traffic, vSphere HA traffic, virtual machine traffic, and so on, you can use the vSphere Network I/O Control in vSphere Distributed Switch to guarantee the amount of bandwidth that is required for Virtual SAN.

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

- Set a reservation so that Network I/O Control guarantees that minimum bandwidth is available on the physical adapter for Virtual SAN.

- Set shares so that when the physical adapter assigned for Virtual SAN becomes saturated, certain bandwidth is available to Virtual SAN and to prevent Virtual SAN 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 Virtual SAN, 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 Virtual SAN

Traffic Type	Reservation, Gbps	Shares
Virtual SAN	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 Virtual SAN on the physical adapter.

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

Marking Virtual SAN Traffic

Priority tagging is a mechanism to indicate to the connected network devices that Virtual SAN traffic has higher QoS demands. You can assign Virtual SAN traffic to a certain class and accordingly mark the traffic with a Class of Service (CoS) value from 0 to 7 by using the traffic filtering and marking policy of vSphere Distributed Switch. The lower the CoS value is, the higher the priority of the Virtual SAN data is.

Segmenting Virtual SAN Traffic in a VLAN

Consider isolating Virtual SAN 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 Virtual SAN 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.

Best Practices for Virtual SAN Networking

Consider networking best practices for Virtual SAN to improve performance and throughput.

- For hybrid configurations, dedicate at least 1-GbE physical network adapter. Place Virtual SAN 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 Virtual SAN traffic on a distributed switch and configure Network I/O Control to guarantee bandwidth to Virtual SAN.

Designing and Sizing Virtual SAN Fault Domains

The Virtual SAN fault domains feature instructs Virtual SAN to 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

You must define at least three fault domains, each of which might consist of one or more hosts. Fault domain definitions must acknowledge physical hardware constructs that might represent a potential failure domain, for example, an individual computing rack enclosure.

If possible, use at least four fault domains. Using a three domains does not allow the use of certain evacuation modes, nor is Virtual SAN able to reprotect data after a failure. In this case, you need a spare fault domain of a capacity for rebuilding, which you cannot provide with the three-domain configuration.

If fault domains are enabled, Virtual SAN applies the active virtual machine storage policy against the fault domains instead of against the individual hosts.

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

```
number of fault domains = 2 * number of failures to tolerate + 1
```

If a host is not a member of a fault domain, Virtual SAN interprets it as a separate domain.

Using Fault Domains Against Failures of Several Hosts

Consider a cluster that contains four server racks, each with two hosts. If the number of failures to tolerate is equal to one and fault domains are not enabled, Virtual SAN 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 in separate fault domains, Virtual SAN ensures that each protection component (replicas and witnesses) is placed on a separate fault domain.

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

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

- Provide enough fault domains to satisfy the number of failures to tolerate that are configured in the storage policies.

  Define at least three 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 domain of free capacity for rebuilding data after a failure, if possible.

Using Boot Devices and Virtual SAN

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

Use a high-quality USB or SD flash drive of 4 GB or larger.

NOTE If the memory size of the ESXi host is greater than 512 GB, boot it from a SATADOM or disk device. When you boot a Virtual SAN host from a SATADOM device, you must use single-level cell (SLC) device and the size of the boot device must be at least 16 GB. In addition, hosts that boot from a disk have a local VMFS. Hence, you must separate the disk for ESXi boot that is not for Virtual SAN.

Log Information and Boot Devices in Virtual SAN

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

You should not store log information on the Virtual SAN datastore. A failure in the Virtual SAN cluster might impact the accessibility of log information.

Consider the following options for persistent log storage:

- Use a storage device that is not used for Virtual SAN 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 *vSphere Installation and Setup* documentation.

Persistent Logging in a Virtual SAN Cluster

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

If you install ESXi on a USB, SD or SATADOM device and you allocate local storage to Virtual SAN, 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. See the *vSphere Installation and Setup* documentation.

Preparing a New or Existing Cluster for Virtual SAN

4

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

This chapter includes the following topics:

Selecting or Verifying the Compatibility of Storage Devices

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

You can choose from several options for Virtual SAN compatibility.

- Use EVO:RAIL, which represents a hyper converged infrastructure appliance to deploy an all-in-one solution.
- Use a Virtual SAN Ready Node server, a physical server that OEM vendors and VMware validate for Virtual SAN 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.
Virtual SAN	■ Magnetic disk SAS or SATA 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 Virtual SAN. Consider also write endurance and performance class. ■ Storage controller model that supports passthrough. Virtual SAN 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 Virtual SAN and for the virtualized workloads that use the Virtual SAN datastore.

Preparing Storage Devices

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

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

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

- The storage devices are local to the ESXi hosts. Virtual SAN cannot claim remote devices.
- The storage devices do not have any preexisting 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. 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.

Virtual SAN 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 Virtual SAN.

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 number of failures to tolerate value in a virtual machine storage policy. A number of failures to tolerate that is greater than 0 extends the device footprint. If the number of failures to tolerate is equal to 1, the footprint is double. If the number of failures to tolerate is equal to 2, the footprint is triple, and so on.
- Verify whether the Virtual SAN datastore has enough space for an operation by examining the space on the individual hosts rather than on the consolidated Virtual SAN datastore object. For example, when you evacuate a host, all free space in the datastore might be on the host that you are evacuating and 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 Virtual SAN cluster.
- Consider the Virtual SAN overhead to the usable storage space.
  - For the new on-disk format version 2.0, the overhead is 1 percent of the capacity per capacity device.
  - For the on-disk format version 1.0, the overhead is 1 GB per capacity device. Both the versions are supported by Virtual SAN.

For more information about planning the capacity of Virtual SAN datastores, see the *VMware Virtual SAN 6.0 Design and Sizing Guide*.

Virtual SAN Policy Impact on Capacity

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

Table 4-1. Virtual SAN VM Policy and Raw Capacity

Aspects of Policy Influence	Description
Policy changes	■ The number of failures to tolerate influences the physical storage space that you must supply for virtual machines. The greater the number of failures to tolerate is for higher availability, the more space you must provide. When a number of failures to tolerate configured to 1, it imposes two replicas of the VMDK file of a virtual machine. If the number of failures to tolerate is set to 1, a VMDK file that is 50 GB requires 100 GB space on different hosts. If the number of failures to tolerate 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. Virtual SAN recreates the new objects that are affected by the change and 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, although the Virtual SAN datastore indicates that enough space is available, the datastore might not be able to evacuate the virtual machine objects because the free space is on the host that is being placed in maintenance mode.

Mark Flash Devices as Capacity Using ESXCLI

Use the esxcli command to manually mark the flash devices on each host as capacity devices.

Prerequisites

Verify that you are using Virtual SAN 6.0.

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 you verify the device status before tagging.

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

Prerequisites

- Verify that you are using Virtual SAN version 6.0.
- Verify that SSH is enabled on the vCenter Server Appliance.

Procedure

1. Open an SSH connection to the vCenter Server Appliance.
2. Log into the appliance 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 Appliance 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 Virtual SAN on the cluster and claim capacity devices.

Providing Memory for Virtual SAN

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

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

Preparing Your Hosts for Virtual SAN

As a part of the preparation for enabling Virtual SAN, 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 Virtual SAN section of the *VMware Compatibility Guide*.
- Make sure that a minimum of three hosts contribute storage to the Virtual SAN 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 a lot of 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 very 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 enclose that is connected to the blade servers and is listed in the Virtual SAN 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.

Virtual SAN and vCenter Server Compatibility

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

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

Preparing Storage Controllers

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

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

Table 4-3. Examining Storage Controller Configuration for Virtual SAN

Storage Controller Feature	Storage Controller Requirement
Required mode	■ Review the Virtual SAN 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.
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 Virtual SAN. 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 Virtual SAN Network

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

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

For information about network design guidelines, see "Designing the Virtual SAN Network," on page 31.

Placing Hosts in the Same Subnet

Hosts must be connected in the same subnet for best networking performance. In Virtual SAN 6.0, you can also connect hosts in the same Layer 3 network if required.

Enabling IP Multicast on the Physical Switches

Verify that the physical switches are configured for multicast traffic so that the hosts can exchange Virtual SAN metadata. Configure an IGMP snooping querier on the physical switches for delivery of multicast messages only through the physical switch ports that are connected to the Virtual SAN hosts.

If you have several Virtual SAN clusters in the same subnet, change the default multicast address for the added cluster.

Dedicating Network Bandwidth on a Physical Adapter

Allocate at least 1 Gbps bandwidth for Virtual SAN. 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 Virtual SAN 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 Virtual SAN.

Configuring a Port Group on a Virtual Switch

Configure a port group on a virtual switch for Virtual SAN.

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

  In the case of a NIC team for network availability, select a teaming algorithm based on the connection of the physical adapters to the switch.
- If designed, assign Virtual SAN traffic to a VLAN by enabling tagging in the virtual switch.

Examining the Firewall on a Host for Virtual SAN

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

Table 4-4. Ports on the Hosts in Virtual SAN

Virtual SAN Service	Traffic Direction	Communicating Nodes	Transport Protocol	Port
Virtual SAN Vendor Provider (vsanvp)	Incoming and outgoing	vCenter Server and ESXi	TCP	8080
Virtual SAN Clustering Service		ESXi	UDP	12345, 23451
Virtual SAN Transport		ESXi	TCP	2233

Change the Multicast Address for a Virtual SAN Cluster

In an environment that contains several Virtual SAN clusters on the same Layer 2 network, for the hosts in each cluster you must assign a unique multicast address so that they receive traffic only for the hosts in the cluster.

Prerequisites

- If you change the multicast address for an active Virtual SAN cluster, you must disable Virtual SAN on it.
- Verify that SSH is enabled on the host.

Procedure

1 Open an SSH connection to a host in the Virtual SAN cluster.

2 To identify the VMkernel adapters for Virtual SAN, run the `esxcli vsan network list` console command and write down the VMkernel adapter identifiers.

3 To change the multicast address for the Virtual SAN cluster, run the `esxcli vsan network set` console command on each VMkernel adapter for Virtual SAN.

```
esxcli vsan network ipv4 set –i vmkX –d agent_group_multicast_address –u
master_group_multicast_address
```

For example, to set the master group multicast address to 224.2.3.5 and the agent group multicast address to 224.2.3.6 on VMkernel adapter vmk1 on the host, run this command.

```
esxcli vsan network ipv4 set –i vmk1 –d 224.2.3.6 –u 224.2.3.5
```

Considerations about the Virtual SAN License

When you prepare your cluster for Virtual SAN, review the requirements of the Virtual SAN 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 Virtual SAN expires, you can continue to use the current configuration of Virtual SAN 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 cluster is a stretched cluster, verify that the stretched clusters feature is available under your license.

- Consider the CPU capacity of the Virtual SAN license across the cluster when adding and removing hosts to the cluster.

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

Creating a Virtual SAN Cluster 5

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

This chapter includes the following topics:

Characteristics of a Virtual SAN Cluster

Before working on a Virtual SAN environment, you should be aware of the characteristics of a Virtual SAN cluster.

A Virtual SAN cluster includes the following characteristics:

- You can have multiple Virtual SAN clusters for each vCenter Server instance. You can use a single vCenter Server to manage more than one Virtual SAN cluster.
- Virtual SAN consumes all devices, including flash cache and capacity devices, and does not share devices with other features.
- Virtual SAN clusters can include hosts with or without capacity devices. The minimum requirement is three hosts with capacity devices. For best result, create a Virtual SAN 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. Virtual SAN allocates 70 percent of all available cache for read cache and 30 percent of available cache for write buffer. In this configurations, the flash devices serve as a read cache and a write buffer.
- In all-flash cluster, 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 Virtual SAN cluster. Virtual SAN cannot consume other external storage, such as SAN or NAS, attached to cluster.

For best practices about designing and sizing a Virtual SAN cluster, see Chapter 3, "Designing and Sizing a Virtual SAN Cluster," on page 23.

Before Creating a Virtual SAN Cluster

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

Requirements for Virtual SAN Cluster

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

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

Table 5-1. Virtual SAN 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 Virtual SAN cluster. For best results, configure the Virtual SAN cluster with four or more hosts.
Memory	■ Verify that each host has a minimum of 6 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 Virtual SAN Hosts," on page 28.
Storage I/O controllers, drivers, firmware	■ Verify that the storage I/O controllers, drivers, and firmware versions are certified and listed in the VCG web site 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 Virtual SAN hosts contributing storage to the cluster must have at least one cache and one capacity device. Virtual SAN requires exclusive access to the local cache and capacity devices of the hosts that you add to the Virtual SAN cluster and 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 Virtual SAN cluster with uniformly configured hosts.
Network connectivity	■ Verify that each host is configured with at least one network adapter. ■ For hybrid configurations, verify that Virtual SAN hosts have a minimum dedicated bandwidth of 1 GbE. ■ For all-flash configurations, verify that Virtual SAN hosts have a minimum bandwidth of 10 GbE. For best practices and considerations about designing the Virtual SAN network, see "Designing the Virtual SAN Network," on page 31 and "Networking Requirements for Virtual SAN," on page 21.

Table 5-1. Virtual SAN Cluster Requirements (Continued)

Requirements	Description
Virtual SAN and vCenter Server Compatibility	Verify that you are using the latest version of the vCenter Server.
License key	■ Verify that you have a valid Virtual SAN license key. ■ To use the all-flash feature, it must be supported by your license. ■ To use the stretched cluster feature, it must be supported by your license. ■ Verify that the amount of license capacity that you plan on using equals the total number of CPUs in the hosts participating in the Virtual SAN cluster and not only the hosts providing capacity to the cluster. For information about licensing for Virtual SAN, see the *vCenter Server and Host Management* documentation.

For detail information about Virtual SAN Cluster requirements, see Chapter 2, "Requirements for Enabling Virtual SAN," on page 19.

For in-depth information about designing and sizing the Virtual SAN cluster, see the *VMware Virtual SAN 6.0 Design and Sizing Guide*.

Enabling Virtual SAN

To use Virtual SAN, you must create a host cluster and enable Virtual SAN on the cluster.

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

- A Virtual SAN cluster must include a minimum of three ESXi hosts. For a Virtual SAN cluster to tolerate host and device failures, at least three hosts that join the Virtual SAN 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 Virtual SAN cluster.
- Before you move a host from a Virtual SAN cluster to another cluster, make sure that the destination cluster is Virtual SAN enabled.
- To be able to access the Virtual SAN datastore, an ESXi host must be a member of the Virtual SAN cluster.

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

Set Up a VMkernel Network for Virtual SAN

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

Procedure

1. In the vSphere Web Client, navigate to the host.
2. Click the **Manage** tab and click **Networking**.
3. Select **VMkernel adapters** and click the **Add host networking** icon.
4. On the Select connection type page, select **VMkernel Network Adapter** and click **Next**.
5. Configure your target device.
6. On the Port properties page, select **Virtual SAN traffic**.

7 Complete the VMkernel adapter configuration.

8 Verify that the Virtual SAN column shows Enabled as the status for the VMkernel adapter.

Virtual SAN is enabled for networking.

What to do next

You can now activate Virtual SAN on the host cluster.

Create a Virtual SAN Cluster

You can enable Virtual SAN when you create a cluster.

Procedure

1 Right-click a data center in the vSphere Web Client and select **New Cluster**.

2 Type a name for the cluster in the **Name** text box.

This name appears in the vSphere Web Client navigator.

3 Select the Virtual SAN **Turn ON** check box.

4 Select the mode for storage devices to be claimed and click **OK**.

Option	Description
Automatic	Claims all empty devices on the included hosts for Virtual SAN. Virtual SAN in automatic mode claims only local devices on the ESXi hosts in the cluster. You can add any remote non-shared devices manually.
Manual	Requires manual claiming of the devices on the included hosts. New devices on the host are not added to Virtual SAN. With the manual mode, two methods of organizing devices into disk groups exist, semi-automatic and manual. NOTE When you use this mode, a Virtual SAN datastore is created, with the initial size of zero byte. The datastore remains unusable until you manually claim devices.

5 Add hosts to the Virtual SAN cluster.

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

The cluster appears in the inventory.

What to do next

Verify that the Virtual SAN datastore has been created. See "View Virtual SAN Datastore," on page 50.

Verify that the Virtual SAN storage provider is registered. See "View Virtual SAN Storage Providers," on page 89.

If you use the manual mode for Virtual SAN, claim devices or create disk groups. See Chapter 8, "Device Management in a Virtual SAN Cluster," on page 67.

Enable Virtual SAN on Existing Clusters

You can edit cluster properties to enable Virtual SAN for an existing cluster.

After enabling Virtual SAN on your cluster, you cannot move Virtual SAN hosts from a Virtual SAN enabled cluster to a non-Virtual SAN cluster.

Prerequisites

Verify that your environment meets all requirements. See Chapter 2, "Requirements for Enabling Virtual SAN," on page 19.

Procedure

1 Browse to the host cluster in the vSphere Web Client navigator.

2 Click the **Manage** tab and click **Settings**.

3 Under Virtual SAN, select **General** and click **Edit** to edit Virtual SAN settings.

4 Select the mode for storage devices to be claimed.

Option	Description
Automatic	Claims all empty devices on the included hosts for Virtual SAN. Virtual SAN in automatic mode claims only local devices on the ESXi hosts in the cluster. You can add any remote non-shared devices manually.
Manual	Requires manual claiming of the devices on the included hosts. New devices on the host are not added to Virtual SAN. With the manual mode, two methods of organizing devices into disk groups exist, semi-automatic and manual. NOTE When you use this mode, a Virtual SAN datastore is created, with the initial size of zero byte. The datastore remains unusable until you manually claim devices.

5 Click **OK**.

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

What to do next

Verify that the Virtual SAN datastore has been created. See "View Virtual SAN Datastore," on page 50.

Verify that the Virtual SAN storage provider is registered. See "View Virtual SAN Storage Providers," on page 89.

If you use the manual mode for Virtual SAN, claim devices or create disk groups. See Chapter 8, "Device Management in a Virtual SAN Cluster," on page 67.

If you want to disable Virtual SAN on a cluster, see "Disable Virtual SAN," on page 49.

Disable Virtual SAN

You can turn off Virtual SAN for a host cluster.

When you disable the Virtual SAN cluster, all virtual machines located on the shared Virtual SAN datastore become inaccessible. If you intend to use virtual machine while Virtual SAN is disabled, make sure you migrate virtual machines from Virtual SAN datastore to another datastore before disabling the Virtual SAN cluster.

Prerequisites

Verify that the hosts are in maintenance mode.

Procedure

1 Browse to the host cluster in the vSphere Web Clientnavigator.

2 Click the **Manage** tab and click **Settings**.

3 Under Virtual SAN, select **General** and click **Edit** to edit Virtual SAN settings.

4 Deselect the **Turn On Virtual SAN** check box.

Assign a License to a Virtual SAN Cluster

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

If you upgrade, combine, or divide Virtual SAN licenses, you must assign the new licenses to Virtual SAN clusters. When you assign a Virtual SAN license to a cluster, the amount of license capacity that is used equals the total number of CPUs in the hosts participating in the cluster. The license usage of the Virtual SAN 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 Virtual SAN on a cluster, you can use Virtual SAN in evaluation mode to explore its features. The evaluation period starts when Virtual SAN is enabled, and expires after 60 days. To use Virtual SAN, you must license the cluster before the evaluation period expires. Just like vSphere licenses, Virtual SAN licenses have per CPU capacity. Some features, such as all-flash configuration and stretched clusters, require a license that supports the feature.

Prerequisites

- To view and manage Virtual SAN licenses, you must have the **Global.Licenses** privilege on the vCenter Server systems, where the vSphere Web Client runs.

Procedure

1 In the vSphere Web Client, navigate to a cluster where you have enabled Virtual SAN.

2 On the **Manage** tab, click **Settings**.

3 Under **Configuration**, select **Licensing**, and click **Assign License**.

4 Select a licensing option.

- Select an existing license and click **OK**.
- Create a new Virtual SAN license.
  a Click the Create New License (+) icon.
  b In the New Licenses dialog box, type or copy and paste a Virtual SAN license key and click **Next**.
  c On the Edit license names page, rename the new license as appropriate and click **Next**.
  d Click **Finish**.
  e In the Assign License dialog, select the newly created license and click **OK**.

View Virtual SAN Datastore

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

Prerequisites

Activate Virtual SAN and configure disk groups.

Procedure

1 Browse to Datastores in the vSphere Web Client navigator.

2 Select the Virtual SAN datastore.

3 Click the **Manage** tab, and click **Settings**.

4 Review the Virtual SAN datastore capacity.

The size of the Virtual SAN 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 would be 7 x 2 TB x 8 = 112 TB. Note that 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. For metadata overhead, 1 GB per drive is allotted when using the on-disk format 1.0. For the new on-disk format 2.0, 1 percent of the total size of the drive is allotted for metadata.

What to do next

Use the storage capabilities of the Virtual SAN datastore to create a storage policy for virtual machines. For information, see the *vSphere Storage* documentation.

Using Virtual SAN and vSphere HA

You can enable vSphere HA and Virtual SAN on the same cluster. As with traditional datastores, vSphere HA provides the same level of protection for virtual machines on Virtual SAN datastores. This level of protection imposes specific restrictions when vSphere HA and Virtual SAN interact.

ESXi Host Requirements

You can use Virtual SAN 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 Virtual SAN cluster with four or more hosts.

Networking Differences

Virtual SAN uses its own logical network. When Virtual SAN 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 Virtual SAN is disabled.vCenter Server chooses the appropriate network when vSphere HA is configured on a host.

NOTE Virtual SAN can be enabled only when vSphere HA is disabled.

When a virtual machine is partially accessible in all network partitions, you cannot power on the virtual machine or fully accessible 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 or cannot be fully accessible in any partition.

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

Table 5-2. vSphere HA Networking Differences

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

If you change the Virtual SAN network configuration, the vSphere HA agents do not automatically acquire the new network settings. To make changes to the Virtual SAN network, you must reenable host monitoring for the vSphere HA cluster by using vSphere Web Client:

1 Disable Host Monitoring for the vSphere HA cluster.

2 Make the Virtual SAN network changes.

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

4 Reenable 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 Number of failures to tolerate policy setting in the Virtual SAN rule set and must not be lower than the capacity reserved by the vSphere HA admission control setting. For example, if the Virtual SAN 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 Number 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.

Virtual SAN and vSphere HA Behavior in a Multiple Host Failure

After a Virtual SAN 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 Virtual SAN virtual machine is provisioned to tolerate one host failure. The virtual machine runs on a Virtual SAN 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.

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 Virtual SAN 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 Virtual SAN cluster from a single site to two sites for a higher 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 sites are active sites. If either site fails, Virtual SAN 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. The system uses the preferred site only in cases where there is a loss of network connection between the two active sites, so the one designated as preferred is the one that remains operational.

A Virtual SAN 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, Virtual SAN automatically switches to fully functional sites.

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

Witness Host

Each stretched cluster consists of two sites and one witness host. The witness host resides at a third site and contains the witness components of virtual machine objects. It contains only metadata, and does not participate in storage operations.

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 Virtual SAN 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 Virtual SAN stretched cluster.
- The witness host must have at least one VMkernel adapter with Virtual SAN traffic enabled, with connections to all hosts in the cluster.
- 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 Update 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 Cluster Versus Fault Domains

Stretched clusters provide redundancy and failure protection across data centers in two geographical locations. Fault domains provide protection from rack-level failures within the same site. 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.

Stretched Cluster Design Considerations

Consider these guidelines when working with a Virtual SAN 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.
  - 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 Virtual SAN Disk Format," on page 63.
- Configure the Number of Failures To Tolerate to one (FTT=1) for stretched clusters.
- Virtual SAN stretched clusters do not support symmetric multiprocessing fault tolerance (SMP-FT).
- When a host is disconnected or not responding, you cannot add or remove the witness host. This limitation ensures that Virtual SAN collects enough information from all hosts before initiating reconfiguration operations.
- Using `esxcli` to add or remove hosts is not supported for stretched clusters.

Best Practices for Working with Stretched Clusters

When working with Virtual SAN 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 sub-cluster containing the other two sites. These new VMs are implicitly force provisioned and will be 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 Virtual SAN 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.

Network Design for Stretched Clusters

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

A Virtual SAN 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.
- Virtual SAN network requires connectivity across all three sites. VMware recommends using a Layer 2 stretched network between the two data sites and a Layer 3 network between the data sites and the witness host.
- VM network requires connectivity between the data sites, but not the witness host. VMware recommends using a Layer 2 stretched 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. VMware supports using a Layer 2 stretched or a Layer 3 network between data sites.

Using Static Routes on ESXi Hosts

If you use a single default gateway on ESXi hosts, note that 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.

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

Consider a Virtual SAN 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 Virtual SAN network on the witness host, the connection will fail because the default gateway on the ESXi host is associated with the management network and there is no route from the management network to the Virtual SAN 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 Virtual SAN 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.

Configure Virtual SAN Stretched Cluster

Configure a Virtual SAN 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 Virtual SAN cluster.
- Verify that the witness host is empty and does not contain any components. To configure an existing Virtual SAN host as a witness host, first evacuate all data from the host and delete the disk group.

Procedure

1 Browse to the Virtual SAN cluster in the vSphere Web Client.

2 Click the **Manage** tab and click **Settings**.

3 Under Virtual SAN, click **Fault Domains**.

4 Click the **Configure Stretched Cluster** icon.

5 Select the fault domain that you want to assign to the secondary site 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 and click **Next**.

All disk groups on the witness host are automatically claimed.

8 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.

Procedure

1 Browse to the Virtual SAN cluster in the vSphere Web Client.

2 Click the **Manage** tab and click **Settings**.

3 Under Virtual SAN, click **Fault Domains**.

4 Select the secondary fault domain and click the **Mark Fault Domain as preferred for Stretched Cluster** icon.

5 Click **Yes** to confirm.

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

Replace the Witness Host

You can replace the witness host for a Virtual SAN stretched cluster.

Remove the existing witness host and add a new witness host.

Prerequisites

Verify that the witness host is not in use.

Procedure

1 Browse to the Virtual SAN cluster in the vSphere Web Client.

2 Click the **Manage** tab and click **Settings**.

3 Under Virtual SAN, click **Fault Domains**.

4 Select the witness host and click the **Remove Witness Host** icon.

5 Click **Yes** to confirm.

The witness host no longer appears in the list of hosts under the fault domain.

6 Select a new witness host and click **Next**.

All disk groups on the new witness host are automatically claimed.

7 Click **Finish**.

Upgrading the Virtual SAN Cluster 7

Upgrading Virtual SAN is a multistage process, in which you must perform the upgrade procedures in the order described here.

Before you attempt to upgrade, make sure you understand the complete upgrade process clearly to ensure a smooth and uninterrupted upgrade. If you are not familiar with the general vSphere upgrade procedure, you should first review the *vSphere Upgrade* documentation.

NOTE Failure to follow the sequence of upgrade tasks described here will lead to data loss and cluster failure.

The Virtual SAN cluster upgrade proceeds in the following sequence of tasks.

1 Upgrade the vCenter Server. See the *vSphere Upgrade* documentation.

2 Upgrade the ESXi hosts. See "Upgrade the ESXi Hosts," on page 61. For information about migrating and preparing your ESXi hosts for upgrade, see the *vSphere Upgrade* documentation.

3 Upgrade the Virtual SAN disk format. Upgrading the disk format is optional and you can run Virtual SAN with hosts for 6.0 Update 1 and the vCenter Server for 6.0 Update 1, while the disk format remains at version 1.0. However, for best results, consider upgrading the objects to use the latest version. The on-disk format exposes your environment to the complete feature set of Virtual SAN. See "Upgrade Virtual SAN Disk Format Using the RVC tool," on page 64.

This chapter includes the following topics:

Before You Upgrade Virtual SAN

Plan and design your upgrade to be fail-safe. Before you attempt to upgrade Virtual SAN, verify that your environment meets the vSphere hardware and software requirements.

Upgrade Prerequisite

Consider the aspects that could delay the overall upgrade process. For guidelines and best practices, see the *vSphere Upgrade* documentation.

Review the key requirements before you upgrade your cluster to Virtual SAN 6.1.

Table 7-1. Upgrade Prerequisite

Upgrade Prerequisites	Description
Software, hardware, drivers, firmware, and storage I/O controllers	Verify that the software and hardware components, drivers, firmware, and storage I/O controllers that you plan on using are supported by Virtual SAN for 6.0 and listed on the VMware Compatibility Guide Web site at http://www.vmware.com/resources/compatibility/search.php.
Virtual SAN version	Verify that you are using the latest version of Virtual SAN. If you are currently running a beta version and plan on upgrading Virtual SAN to 6.1, your upgrade will fail. When you upgrade from a beta version, you must perform a fresh deployment of Virtual SAN.
Disk space	Verify that you have enough space available to complete the software version upgrade. The amount of disk storage needed for the vCenter Server installation depends on your vCenter Server configuration. For guidelines about the disk space required for upgrading vSphere, see the *vSphere Upgrade* documentation.
Virtual SAN disk format	Verify that you have enough capacity available to upgrade the disk format. To upgrade to the new on-disk format 2.0, you must have free space equal to the consumed capacity of the largest disk group. This space must be available on disk groups other than the disk groups that are currently being converted. For example, the largest disk group in a cluster has 10 TB of physical capacity, but only 5 TB is being consumed. An additional 5 TB of spare capacity will be needed elsewhere in the cluster, excluding the disk groups that are being migrated. When upgrading the Virtual SAN disk format, verify that the hosts are not in maintenance mode. When any member host of a Virtual SAN cluster enters maintenance mode, the cluster capacity is automatically reduced, because the member host no longer contributes storage to the cluster and the capacity on the host is unavailable for data. For information about various evacuation modes, see the "Place a Member of Virtual SAN Cluster in Maintenance Mode," on page 81.
Virtual SAN hosts	Verify that you have placed the Virtual SAN hosts in maintenance mode and selected the Ensure Accessibility or Full data migration mode. **You can use the vSphere Update Manager for automating and testing the upgrade process.** However, when you use vSphere Update Manager to upgrade Virtual SAN, the default evacuation mode is Ensure Accessibility. When you use the Ensure Accessibility mode, your data is not completely protected, and if you encounter a failure while upgrading Virtual SAN, you might experience unexpected data loss. However, the Ensure Accessibility mode is faster than the Full data migration mode, because you do not need to move all data to another host in the cluster. For information about various evacuation modes, see the "Place a Member of Virtual SAN Cluster in Maintenance Mode," on page 81.
Virtual Machines	Verify that you have backed up your virtual machines.

Recommendations

Consider the following recommendations when deploying ESXi hosts for use with Virtual SAN:

- Use SATADOM, SD, USB, or hard disk devices as the installation media whenever ESXi hosts are configured with memory capacity of approximately 512 GB or less.
- Use a separate magnetic disk or flash device as the installation device whenever ESXi hosts are configured with more than 512 GB memory. If you are using a separate device, verify that Virtual SAN is not claiming the device.
- When you boot a Virtual SAN host from a SATADOM device, you must use a single-level cell (SLC) device and the size of the boot device must be at least 16 GB.

Upgrade the vCenter Server

This first task to perform during the Virtual SAN upgrade is a general vSphere upgrade, which includes upgrading vCenter Server and ESXi hosts.

VMware supports in-place upgrades on 64-bit systems from vCenter Server 4.x, vCenter Server 5.0.x, vCenter Server 5.1.x, and vCenter Server 5.5 to vCenter Server 6.0 and later. The vCenter Server upgrade includes a database schema upgrade and an upgrade of the vCenter Server. Instead of performing an in-place upgrade to vCenter Server, you can use a different machine for the upgrade. For detailed instructions and various upgrade options, see the *vSphere Upgrade* documentation.

Upgrade the ESXi Hosts

After you upgrade the vCenter Server, the next task for the Virtual SAN cluster upgrade is upgrading the ESXi hosts to use the current version.

If you have multiple hosts in the Virtual SAN cluster, and you use vSphere Update Manager to upgrade the hosts, the default evacuation mode is Ensure Accessibility. If you use this mode, and while upgrading Virtual SAN you encounter a failure, your data will be at risk. For information about working with evacuation modes, see "Place a Member of Virtual SAN Cluster in Maintenance Mode," on page 81

For information about using vSphere Update Manager, see the documentation Web site at https://www.vmware.com/support/pubs/vum_pubs.html .

Before you attempt to upgrade the ESXi hosts, review the best practices discussed in the *vSphere Upgrade* documentation. VMware provides several ESXi upgrade options. Choose the upgrade option that works best with the type of host that you are upgrading. For more information about various upgrade options, see the *vSphere Upgrade* documentation.

Prerequisites

- Verify that you have sufficient disk space for upgrading the ESXi hosts. For guidelines about the disk space requirement, see the *vSphere Upgrade* documentation.
- Verify that you are using the latest version of ESXi. You can download the latest ESXi installer from the VMware product download Web site at https://my.vmware.com/web/vmware/downloads.
- Verify that you are using the latest version of vCenter Server.
- Verify the compatibility of the network configuration, storage I/O controller, storage device, and backup software.
- Verify that you have backed up the virtual machines.
- Use Distributed Resource Scheduler (DRS) to prevent virtual machine downtime during the upgrade. Verify that the automation level for each virtual machine is set to Fully Automated mode to help DRS migrate virtual machines when hosts are entering maintenance mode. Alternatively, you can also power off all virtual machines or perform manual migration.

Procedure

1 Place the host that you intend to upgrade in maintenance mode.

  You must begin your upgrade path with ESXi 5.5 or later hosts in the Virtual SAN cluster.

2 Select the **Ensure Accessibility** or **Full data migration** evacuation mode, depending on your requirement, and wait for the host to enter maintenance mode.

   If you are using vSphere Update Manager to upgrade the host, or if you are working with a three-host cluster, the default evacuation option available is Ensure Accessibility. This option is faster than the Full data migration evacuation mode. However, the Ensure Accessibility mode does not fully protect your data. During failure your data might be at risk and you might experience downtime, and unexpected data loss.

3 Upload the software to the datastore of your ESXi host and verify that the file is available in the directory inside the datastore. For example, you can upload the software to `/vmfs/volumes/<datastore>/VMware-ESXi-6.0.0-1921158-depot.zip`.

4 Run the `esxcli` command `install -d /vmfs/volumes/53b536fd-34123144-8531-00505682e44d/depot/VMware-ESXi-6.0.0-1921158-depot.zip --no-sig-check`. Use the `esxcli` software VIB to run this command.

   After the ESXi host for Virtual SAN 6.0 has installed successfully, you see the following message:

   ```
   The update completed successfully, but the system needs to be rebooted for the changes to be effective.
   ```

5 You must manually restart your ESXi host from the vSphere Web Client.

   a Navigate to the ESXi host in the vSphere Web Client inventory.

   b Right-click the host, select **Power > Reboot**, click **Yes** to confirm, and then wait for the host to restart.

   c Right-click the host, select **Connection > Disconnect**, and then select **Connection > Connect** to reconnect to the host.

      To upgrade the remaining hosts in the cluster, repeat this procedure for each host.

      If you have multiple hosts in your Virtual SAN cluster, you can use vSphere Update Manager to upgrade the remaining hosts.

What to do next

1 (Optional) Upgrade the Virtual SAN disk format. See the "Upgrade Virtual SAN Disk Format Using the RVC tool," on page 64 topic.

2 Verify the host license. In most cases, you must reapply your host license. You can use vSphere Web Client and vCenter Server for applying host licenses. For more information about applying host licenses, see the *vCenter Server and Host Management* documentation.

3 (Optional) Upgrade the virtual machines on the hosts by using the vSphere Web Client or vSphere Update Manager.

Before You Upgrade Virtual SAN Disk Format

Upgrading the disk format 1.0 to the new on-disk format 2.0 is optional and a Virtual SAN cluster will continue to run smoothly if you choose to use disk format version 1.0.

For best results, upgrade the objects to use the new on-disk format 2.0. The new on-disk format 2.0 provides the complete feature set of Virtual SAN.

Depending on the size of disk groups, the disk format upgrade can be time-consuming because RVC upgrades one disk group at a time. For each disk group upgrade, all data from each device in a disk group is evacuated and the disk group is removed from the Virtual SAN cluster. The disk group is then added back to Virtual SAN with the new on-disk format 2.0.

During the upgrade, you can monitor the performance of the applications by using the RVC tool or from the vSphere Web Client when you navigate to the Resyncing components page. See "Monitor the Resynchronization Tasks in the Virtual SAN Cluster," on page 95.

You can monitor other upgrade tasks, such as device removal and upgrade, from the vSphere Web Client in the Recent Tasks pane of the status bar.

The following considerations apply when upgrading the disk format:

- If you are upgrading the Virtual SAN cluster that contains three hosts where each host contains a disk group, and you want to perform a full evacuation to protect against a potential failure that could cause data loss, the evacuation will fail for objects that are configured with Number of Failures to Tolerate greater than zero. The reason is that a three-host cluster cannot reprotect a disk group that is being fully evacuated using the resources of only two hosts, For example, when the Number of Failures to Tolerate=1, Virtual SAN requires three protection components (two mirrors and a witness), where each protection component is placed on separate hosts.

  For a three-host cluster, you must choose the Ensure Accessibility evacuation mode. When in this mode, any hardware failure might result in data loss.

  You also must ensure that enough free space is available. The space must be equal to the logical consumed capacity of the largest disk group. This capacity must be available on hosts separate from the one that is being migrated. Use the `vsan.whatif_host_failures` RVC command to determine if enough capacity is available.

- When working with a three-host cluster or when upgrading Virtual SAN with limited resources, run the RVC command with the option, `vsan.v2_ondisk_upgrade --allow-reduced-redundancy`, to allow the virtual machines to operate in a reduced redundancy mode during upgrade.

- Using the `--allow-reduced-redundancy` command option means certain virtual machines might be unable to tolerate failures during the migration. This lowered tolerance for failure also can cause data loss. Virtual SAN restores full compliance and redundancy after the upgrade is completed. During the upgrade, the compliance status of virtual machines and their redundancies are temporarily noncompliant. After you complete the upgrade and finish all rebuild tasks, the virtual machines will become compliant.

  For information about the RVC commands and command options, see the *RVC Command Reference Guide*.

Upgrade Virtual SAN Disk Format

After you have finished upgrading the Virtual SAN hosts, you can continue with the disk format upgrade.

Prerequisites

- Verify that you are using the updated version of vCenter Server.
- Verify that you are using the latest version of ESXi hosts.
- Verify that the disks are in a healthy state. Navigate to the Disk Management page in the vSphere Web Client to verify the object status.
- Verify that the hardware and software that you plan on using are certified and listed in the VMware Compatibility Guide Web site at http://www.vmware.com/resources/compatibility/search.php.
- Verify that you have enough free space to perform the disk format upgrade. Run the RVC command, `vsan.whatif_host_failures`, to determine whether you have enough capacity to successfully finish the upgrade or perform a component rebuild, in case you encounter any failure during the upgrade.
- Verify that your hosts are not in maintenance mode. When upgrading the disk format, you should not place the hosts in maintenance mode. When any member host of a Virtual SAN cluster enters maintenance mode, the available resource capacity in the cluster is reduced because the member host no longer contributes capacity to the cluster and the cluster upgrade might fail.

- Verify that there are no component rebuilding tasks currently in progress in the Virtual SAN cluster. See "Monitor the Resynchronization Tasks in the Virtual SAN Cluster," on page 95.

Procedure

1 Browse to the Virtual SAN cluster in the vSphere Web Client navigator.

2 Click the **Manage** tab and click **Settings**.

3 Under Virtual SAN, select **General**.

4 Under On-disk Format Version, click **Upgrade**.

The Disk Format Version column displays the updated disk format version (version 2.0).

The Disks with outdated version column indicates the number of devices using the new format.

What to do next

Verify the disk format upgrade. See "Upgrade Virtual SAN Disk Format," on page 63.

Upgrade Virtual SAN Disk Format Using the RVC tool

After you have finished upgrading the Virtual SAN hosts, you can use the RVC tool to continue with the disk format upgrade.

Prerequisites

- Verify that you are using the updated version of vCenter Server.
- Verify that the version of the ESXi hosts running in the Virtual SAN cluster is 6.0.
- Verify that the disks are in a healthy state from the Disk Management page in the vSphere Web Client. You can also run the `vsan.disk_stats` RVC command to verify disk status.
- Verify that the hardware and software that you plan on using are certified and listed in the VMware Compatibility Guide Web site at http://www.vmware.com/resources/compatibility/search.php.
- Verify that you have enough free space to perform the disk format upgrade. Run the RVC `vsan.whatif_host_failures` command to determine that you have enough capacity to successfully finish the upgrade or perform a component rebuild in case you encounter failure during the upgrade.
- Verify that you have PuTTY or similar SSH client installed for accessing the RVC tool.

  For detailed information about downloading the RVC tool and using the RVC commands, see the *RVC Command Reference Guide*.
- Verify that your hosts are not in maintenance mode. When upgrading the disk format from version 1.0 to the new on-disk format, version 2.0, do not place your hosts in maintenance mode. When any member host of a Virtual SAN cluster enters maintenance mode, the available resource capacity in the cluster is reduced because the member host no longer contributes capacity to the cluster and the cluster upgrade might fail.
- Verify that there are no component rebuilding tasks currently in progress in the Virtual SAN cluster by running the RVC `vsan.resync_dashboard` command.

Procedure

1 Log in to your vCenter Server using the RVC tool.

2 Run the `vsan.disks_stats /< vCenter IP address or hostname>/<data center name>/computers/<cluster name>` command to view the disk status.

   For example: `vsan.disks_stats /192.168.0.1/BetaDC/computers/VSANCluster`

   The command lists the names of all devices and hosts in the Virtual SAN cluster. The command also displays the current disk format and its health status. You can also check the current health of the devices in the **Health Status** column from the Disk Management page. For example, the device status appears as Unhealthy in the **Health Status** column for the hosts or disk groups that have failed devices.

3 Run the `vsan.v2_ondisk_upgrade <path to vsan cluster>` command .

   For example: `vsan.v2_ondisk_upgrade /192.168.0.1/BetaDC/computers/VSANCluster`

4 Monitor the progress in RVC.

   RVC upgrades one disk group at a time.

   After the disk format upgrade has completed successfully, the following message appears.

   ```
   Done with disk format upgrade phase

   There are n v1 objects that require upgrade Object upgrade progress: n upgraded, 0 left

   Object upgrade completed: n upgraded

   Done VSAN upgrade
   ```

5 Run the `vsan.obj_status_report` command to verify that the object versions are upgraded from version 1.0 to version 2.0.

Verify the Virtual SAN Disk Format Upgrade

After you finish upgrading the disk format, you must verify whether the Virtual SAN cluster is using the new on-disk format.

Procedure

1 Browse to the Virtual SAN cluster in the vSphere Web Client navigator.

2 Click the **Manage** tab and click **Settings** > **Virtual SAN** > **Disk Management**.

   The disk format version that you are using appears in the Disk Format Version column. For example, if you are using disk format 1.0, it appears as version 1 in the Disk Format Version column. For on-disk format 2.0, the disk format version appears as version 2.

Verify the Virtual SAN Cluster Upgrade

The Virtual SAN cluster upgrade is not complete until you have verified that you are using the latest version of vSphere and Virtual SAN is available for use.

Procedure

1 Browse to the Virtual SAN cluster in the vSphere Web Client navigator.

2 Click the **Settings** tab, expand the Configuration section, and verify whether Virtual SAN is listed.

   - You can also browse to your ESXi host and select **Summary** > **Configuration**, and verify that you are using the latest version of the ESXi host.

Using the RVC Upgrade Command Options

The `vsan.v2_ondisk_upgrade` command provides various command options that you can use to control and manage the Virtual SAN cluster upgrade. For example, you can downgrade disk formats and eliminate objects version upgrade.

Run the `vsan.v2_ondisk_upgrade --help` command on your RVC tool to display the list of command options.

Use these command options with the `vsan.v2_ondisk_upgrade` command.

Table 7-2. Upgrade Command Options

Options	Description
`-- hosts_and_clusters`	Use to specify paths to all host systems in the cluster or cluster's compute resources.
`--ignore-objects, -i`	Use to skip Virtual SAN object upgrade. You can also use this command option to eliminate the object version upgrade. When you use this command option, objects continue to use version 1.0.
`--downgrade-format, -d:`	Use to downgrade disk format and file system. Use only if there is no version 2.0 object in the Virtual SAN cluster. This option also disables the Virtual SAN file system version 2.0 on the selected hosts and restricts creating disk groups based on version 2.0.
`--allow-reduced-redundancy, -a`	Use to remove the requirement of having a free space equal to one disk group during disk upgrade. With this option, virtual machines operate in a reduced redundancy mode during upgrade, which means certain virtual machines might be unable to tolerate failures temporarily and that inability might cause data loss. Virtual SAN restores full compliance and redundancy after the upgrade is completed.
`--force, -f`	Use to enable force-proceed and automatically answer all confirmation questions.
`--help, -h`	Use to display the help options.

For information about using the RVC commands, see the *RVC Command Reference Guide*.

Device Management in a Virtual SAN Cluster 8

You can perform various device management tasks in a Virtual SAN cluster, such as, create hybrid or all-flash disk groups, enable Virtual SAN to automatically claim devices for capacity and cache, enable or disable LEDs indicator on devices, mark devices as flash, mark remote devices as local, and so on.

This chapter includes the following topics:

Managing Disk Groups And Devices

Depending on the mode you select when enabling Virtual SAN on a cluster, you can use different ways to organize devices into groups.

Automatic Mode

In automatic mode, Virtual SAN automatically discovers and claims all local, empty, and usable devices on each host and organizes them into default disk groups with one cache and one or multiple capacity devices on each host in the cluster. If you add more capacity to the hosts or add new hosts with capacity to the Virtual SAN cluster, the local storage on the host is automatically claimed by Virtual SAN and the total capacity of the Virtual SAN datastore is adjusted. Virtual SAN in automatic mode claims only local devices on the Virtual SAN hosts in the cluster. In an all-flash cluster, you have to manually mark flash devices for capacity. In an environment without any HDD disks, if there are no flash disks marked for capacity flash, Virtual SAN in automatic mode won't claim any disks or create any disk groups. See "Mark Flash Devices as Capacity Using ESXCLI," on page 37 or "Mark Flash Devices as Capacity using RVC," on page 39.

NOTE When hosts are using SAS controllers, Virtual SAN might identify certain devices as remote and unable to automatically claim devices as local and the devices are displayed as remote. In such situations, you have to manually create disk groups, even though the cluster is set up as automatic. You can also manually add any remote non-shared devices. For information about manually marking storage devices as local, see the "Mark Devices as Local," on page 73 topic.

After the automatic claiming of devices is completed, the Virtual SAN shared datastore is created. The total size of the datastore reflects the capacity of all capacity devices in disk groups across all hosts in the cluster, except for some allotted for metadata overhead. For metadata overhead, 1 GB per drive is allotted when using on-disk format for version 1. For version 2 disk format, 1 percent of the size of the drive is allotted for metadata.

Manual Mode

In manual mode, you must manually specify hosts and devices on the hosts to be used for the Virtual SAN datastore. You can organize devices into disk groups using the semi-automatic or manual method.

When you use the semi-automatic method, Virtual SAN claims and organizes the devices that you specify into default disk groups.

When you use the manual method, you create user-defined disk groups and individually select devices for each disk group. When creating a disk group manually, your main consideration should be the ratio of flash cache to consumed capacity. Although the ratio should depend on the requirement and workload of the cluster, however, for best results for both hybrid and all-flash configurations, you must consider at least a 10 percent of flash cache to consumed capacity ratio, without counting the protection copies.

When you select the manual mode, the Virtual SAN cluster is created with a single Virtual SAN datastore with zero bytes consumed, initially.

As you create disk groups on each host and add at least one cache and one or more capacity devices to each of the disk groups, the size of the datastore grows according to the amount of physical capacity added to the cluster. Virtual SAN creates a single distributed Virtual SAN datastore using the local empty capacity available from the hosts added to the cluster.

When working with manual mode, if the cluster requires multiple flash cache devices, then you need to create multiple disk groups manually as there can only be a maximum of one flash cache device per disk group.

NOTE If a new ESXi host is added to the Virtual SAN cluster, which has been set up in manual mode, the local storage from that host is not added to the Virtual SAN Datastore automatically. You have to manually create disk groups and add the devices to the disk group in order to use the new storage from the new ESXi host.

The manual method of claiming devices differs from the semi-automatic method. If you use the semi-automatic method, you can select multiple devices to be consumed by Virtual SAN, but it does not allow you to organize them. Virtual SAN creates default disk groups for you. But if you manually create a disk group, you can be more specific and organize selected devices into disk groups. You first select a cache device and then add one or more capacity devices of your choice to form a disk group.

Use Semi-Automatic Method to Claim Devices

When Virtual SAN is enabled in manual mode for the host cluster, you must select cache and capacity devices to support the cluster. After you select the devices, Virtual SAN organizes them into default disk groups.

Procedure

1 Browse to the Virtual SAN cluster in the vSphere Web Client navigator.

2 Click the **Manage** tab and click **Settings**.

3 Under Virtual SAN, click **Disk Management**.

4 Click the **Claim Disks** icon.

5 Select devices to be added to the disk group.

- For hybrid disk group, you must select at least one flash cache and one or multiple capacity devices on each host. If you select multiple flash cache devices, the number of capacity devices must be equal to or greater than the number of flash cache devices. You can add only one flash cache device per disk group.
  - From the list of devices, select an HDD device to be used as capacity and click the **Claim for capacity tier** icon.
  - Select a flash device to be used as cache and click the **Claim for cache tier** icon.
  - Click **OK**.
- For all-flash disk group, select flash devices for both capacity and cache.
  - Select a flash device to be used for capacity and click the **Claim for capacity tier** icon.
  - Select a flash device to be used as cache and click the **Claim for cache tier** icon.
  - Click **OK**.

To verify the role of each device added to the all-flash disk group, navigate to the Disk Role column at the bottom of the Disk Management page. The column shows the list of devices and their purpose in a disk group.

Virtual SAN claims the devices that you selected and organizes them into default disk groups that backs the Virtual SAN datastore.

Use Manual Method to Claim Devices for Virtual SAN

When Virtual SAN is enabled in manual mode for the host cluster, you can manually combine specific cache devices with specific capacity devices to create user-defined disk groups on a particular host.

In this method, you manually select devices to create disk groups for each host and add at least one capacity device and one cache device to each disk group.

Procedure

1 Browse to the Virtual SAN cluster in the vSphere Web Client navigator.

2 Click the **Manage** tab and click **Settings**.

3 Under Virtual SAN, click **Disk Management**.

4 Select the host and click the **Create a new disk group** icon.

- Select the flash device to be used for cache.
- From the **Capacity type** drop-down menu, select the type of capacity disks to use, depending on the type of disk group you want to create (HDD for hybrid or Flash for all-flash).
  - Select the devices you want to use for capacity.

5 Click **OK**.

The new disk group appears on the list.

Working with Individual Devices

This topic discusses the various device management tasks that you can perform in the Virtual SAN cluster.

Add Devices to the Disk Group

When you use Virtual SAN in manual mode, you can add additional local devices to existing disk groups.

The devices that you intend to add must be of the same type as the ones existing in the disk groups, for example SSD or magnetic disks.

Procedure

1 Browse to the Virtual SAN cluster in the vSphere Web Client navigator.

2 Click the **Manage** tab and click **Settings**.

3 Under Virtual SAN, click **Disk Management**.

4 Select the disk group, and click the **Add a disk to the selected disk group** icon.

5 Select the device that you want to add and click **OK**.

If you added devices to the disk group that have been used and contains residual data or partition information, you must clean the device. For information about removing partition information from devices, see "Remove Partition From Devices," on page 74. You can also run the `host_wipe_vsan_disks` RVC command to format the device. For detail information about the RVC commands, see the *RVC Command Reference Guide*.

Remove Disk Groups or Devices from Virtual SAN

You can remove an entire disk group or selected devices from the disk group.

Because removing unprotected devices might be disruptive for the Virtual SAN datastore and virtual machines on the datastore, avoid removing devices or disk groups.

Typically, you delete devices or disk groups from Virtual SAN when you are upgrading a device or a device failure is detected in the cluster, or when you must remove a cache device. Other vSphere storage features can use any flash-based device that you remove from the Virtual SAN cluster.

Deleting a disk group permanently deletes the disk membership as well as the data stored on the devices.

NOTE Removing one flash cache device or all capacity devices from a disk group removes the entire disk group.

Evacuating data from devices or disk groups might result in the temporary noncompliance of virtual machine storage policies.

Prerequisites

- You can either place the Virtual SAN host in maintenance mode by selecting the **Full data migration** option or select the **Evacuate data** check box when deleting a device or a disk group. If you select the **Evacuate data** check box, your data might be at risk, if a failure occurs during evacuation.
- You can remove devices or disk groups only when the Virtual SAN cluster is set up in manual mode. For the automatic device claim mode, the remove action is not supported.

Procedure

1 Browse to the Virtual SAN cluster in the vSphere Web Client navigator.

2 Click the **Manage** tab and click **Settings**.

3 Under Virtual SAN, click **Disk Management**.

4 Remove a disk group or selected devices.

Option	Description
Remove the Disk Group	a Under Disk Groups, select the disk group to remove, and click the **Remove the Disk Group** icon. b Select the **Evacuate data** check box to evacuate all data from the selected disk group. NOTE The **Remove the Disk Group** icon does not appear when the Virtual SAN cluster is set up in automatic mode. The icon appears only when the cluster is in manual mode.
Remove the Selected Disk	a Under Disk Groups, select the disk group that contains the device that you are removing. b Under Disks, select the device to remove, and click the **Remove the Selected Disks** icon. c Select the **Evacuate data** check box to evacuate all data from the selected devices.

You can move the evacuated data to another disk or disk group on the same host.

5 Click **Yes** to confirm.

The data is evacuated from the selected devices or a disk group and is no longer available to Virtual SAN services on the selected host.

Using Locator LEDs

Locator LEDs enable you to identify the location of storage devices during failure.

Enabling locator LEDs, allows Virtual SAN to light LED on a failed device so that you can easily identify the device. This is particularly useful when you are working with multiple hot plug and host swap scenarios.

When you want to use locator LEDs, you should consider using I/O storage controllers with pass-through mode, because controllers with RAID 0 mode require additional steps to enable the controllers to recognize locator LEDs.

For information about configuring storage controllers with RAID 0 mode, see your vendor documentation.

Enable and Disable Locator LEDs

You can turn on or off locator LEDs on Virtual SAN storage devices to identify the location of specific storage device.

When you no longer need a visual alert on the status of your Virtual SAN devices, you can disable locator LEDs on the selected Virtual SAN devices.

Prerequisites

- Verify that you have installed the supported drivers for storage I/O controllers that enable this feature. For information about the drivers that are certified by VMware, see the VCG Web site at http://www.vmware.com/resources/compatibility/search.php.
- In some cases, you might need to use third-party utilities to configure the Locator LED feature on your storage I/O controllers. Such as, when you are using HP you should verify that the HP SSA CLI is installed.

For information about installing third-party VIBs, see the *vSphere Upgrade* documentation.

Procedure

1 Browse to the Virtual SAN cluster in the vSphere Web Client navigator.

2 Click the **Manage** tab and click **Settings**.

3 Under Virtual SAN, click **Disk Management**.

4 Select a host to view the list of devices.

5 From the bottom of the page, select one or more storage devices from the list of device, and enable or disable the locator LEDs on the selected storage device.

Option	Action
Turns on the locator LED of the selected disk(s) icon	Enables locator LED on the selected storage device. You can enable locator LEDs from the **Manage** tab and click **Storage > Storage Devices**.
Turns off the locator LED of the selected disk(s) icon	Disables locator LED on the selected storage device. You can disable locator LEDs from the **Manage** tab and click **Storage > Storage Devices**.

Mark Flash Devices as Cache

When flash devices are not automatically identified as flash by ESXi hosts, you can manually mark them as local flash cache or flash capacity devices.

This can also happen when they are enabled for RAID 0 mode rather than passthrough mode. When devices are not recognized as local flash, they are excluded from the list of devices offered for Virtual SAN and you cannot use them in the Virtual SAN cluster. Marking these devices as local flash makes them available to Virtual SAN.

Prerequisites

- Verify that the device is local to your host.
- Verify that the device is not in use.
- Make sure that the virtual machines accessing the device are powered off and the datastore is unmounted.

Procedure

1 Browse to the Virtual SAN cluster in the vSphere Web Client navigator.

2 Click the **Manage** tab and click **Settings**.

3 Under Virtual SAN, click **Disk Management**.

4 Select the host to view the list of available devices that you want to mark as local flash device.

5 From the **Show** drop-down menu at the bottom of the page, select **Not in Use**.

6 Select one or more flash devices from the list and click the **Mark the selected disks as flash disks** icon.

7 Click **Yes** to save your changes.

The Drive type for the selected devices appear as Flash.

Mark HDD Devices as Capacity

In a hybrid cluster, you can manually mark local magnetic disks that are not automatically detected by ESXi.

In case, if you marked a magnetic disk as a flash device, you can change the disk type of the device manually by marking it as a magnetic disk.

Prerequisites

- Verify that the magnetic disk is local to your host.
- Verify that the magnetic disk is not in use and is empty.

- Verify that the virtual machines accessing the device are powered off.

Procedure

1. Browse to the Virtual SAN cluster in the vSphere Web Client navigator.
2. Click the **Manage** tab and click **Settings**.
3. Under **Virtual SAN**, click **Disk Management**.
4. Select the host to view the list of available magnetic disks that you want to use as a capacity device.
5. From the **Show** drop-down menu at the bottom of the page, select **Not in Use**.
6. Select one or more magnetic disks from the list and click **Mark the selected disks as HDD disks** icon.
7. Click **Yes** to save.

   The Drive Type for the selected magnetic disks appear as HDD.

Mark Devices as Local

When hosts are using external SAS enclosures, Virtual SAN might recongnize certain devices as remote and unable to automatically claim them as local.

In such cases, you should manually mark devices as local.

Prerequisites

Make sure that the storage device is not shared.

Procedure

1. Browse to the Virtual SAN cluster in the vSphere Web Client navigator.
2. Click the **Manage** tab and click **Settings**.
3. Under **Virtual SAN**, click **Disk Management**.
4. Select a host to view the list of devices.
5. From the **Show** drop-down menu at the bottom of the page, select **Not in Use**.
6. From the list of devices, select one or more remote devices that you want to mark as local and click the **Mark the selected disks as local for the host** icon.
7. Click **Yes** to save your changes.

Mark Devices as Remote

Devices can be shared between hosts, when they are using external SAS enclosures. You can manually mark those shared devices as remote, so that Virtual SAN does not automatically claim the devices, when creating disk groups.

In Virtual SAN, you cannot add shared devices to a disk group.

If you want to replace hardware components, see "Replacing Existing Hardware Components," on page 113.

Procedure

1. Browse to the Virtual SAN cluster in the vSphere Web Client navigator.
2. Click the **Manage** tab and click **Settings**.
3. Under **Virtual SAN**, click **Disk Management**.
4. Select a host to view the list of devices.

5 From the **Show** drop-down menu at the bottom of the page, select **Not in Use**.

6 Select one or more devices that you want to mark as remote and click the **Marks the selected disk(s) as remote for the host** icon.

7 Click **Yes** to confirm.

Add a Capacity Device

You can add a capacity device to an existing Virtual SAN disk group.

You cannot add a shared device to a disk group.

Prerequisites

Verify that the device is formatted and is not in use.

Procedure

1 Browse to the Virtual SAN cluster in the vSphere Web Client navigator.

2 Click the **Manage** tab and click **Settings**.

3 Under Virtual SAN, click **Disk Management**.

4 From the **All Actions** menu at the bottom of the page, select **Add a disk**.

5 Select the capacity device that you want to add to the disk group.

6 Click **OK**.

The device is added to the disk group and is added to the list of the devices.

Remove Partition From Devices

You can remove partition information from devices using the vSphere Web Client.

If you have added a device that contains residual data or partition information, you must remove all preexisting partition information from the device before you can claim it for Virtual SAN use. Virtual SAN recommends adding clean devices to disk groups.

When you remove partition information from a device, Virtual SAN deletes primary partition that includes disk format information and logical partitions from the device.

Prerequisites

Verify that the device is not in use by ESXi as boot disk, VMFS datastore, or Virtual SAN.

Procedure

1 Browse to the Virtual SAN cluster in the vSphere Web Client navigator.

2 Click the **Manage** tab and click **Settings**.

3 Under **Virtual SAN**, click **Disk Management**.

4 Select the disk group and select the disks to format from the ineligible group.

5 Click the **Erase partitions on the selected disks** icon.

6 Click **OK** to confirm.

The device is clean and does not include any partition information.

Unclaim Devices

You can unplug devices from a disk group that are no longer in use.

Unclaimed devices are automatically unplugged from the disk group and become available to other Virtual SAN disk groups.

Prerequisites

Verify that the device is not in use.

Procedure

1 Browse to the Virtual SAN cluster in the vSphere Web Client navigator.

2 Click the **Manage** tab and click **Settings**.

3 Under Virtual SAN, click **Disk Management**.

4 Select the disk group and click the **Claim disks** icon.

5 Select the device that you want to unclaim and click the **Do not claim** icon.

You can also unclaim a device by navigating to the Claim For column, selecting a device and choosing the **Do not claim** option.

The selected device is no longer part of the disk group.

Expanding and Managing a Virtual SAN Cluster

9

After you have set up your Virtual SAN cluster, you can use the vSphere Web Client to add hosts and capacity devices, remove hosts and devices, and manage failure scenarios.

This chapter includes the following topics:

Expanding a Virtual SAN Cluster

You can expand an existing Virtual SAN cluster by adding hosts or devices to the hosts without disrupting any ongoing operations.

Use one of the following methods to expand your Virtual SAN cluster.

- Add new ESXi hosts to the cluster that are configured using the supported cache and capacity devices. See "Add a Host to the Virtual SAN Cluster," on page 78. When you add a new device or a add a host with capacity, Virtual SAN does not automatically distribute data to the newly added device. To enable Virtual SAN to distribute data to the devices that are recently added, you must perform manual rebalance operation in the cluster by using the Ruby vSphere Console (RVC) tool RVC. See "Manual Rebalance," on page 96.
- Move existing ESXi hosts to the Virtual SAN cluster by using host profile. See "Configuring Hosts Using Host Profile," on page 79. New cluster members add storage and compute capacity. If you use Virtual SAN in automatic mode, the local capacity devices on the new cluster member will be automatically aggregated into a disk group and claimed by the Virtual SAN datastore. If Virtual SAN is set to manual, you must manually create a subset of disk groups from the local capacity devices on the newly added host. See "Use Manual Method to Claim Devices for Virtual SAN," on page 69.

  Make sure that the hardware components, drivers, firmware, and storage I/O controllers that you plan on using are certified and listed in the VMware Compatibility Guide web site at http://www.vmware.com/resources/compatibility/search.php. When adding capacity devices, make sure that the devices are unformatted and not partitioned. Otherwise, Virtual SAN will not to recognize the devices.
- Add new capacity devices to ESXi hosts that are cluster members. If you use Virtual SAN in automatic mode, the new devices that you add joins an existing disk group. If Virtual SAN is set to manual, you must manually add the device to the disk group on the host. See "Add Devices to the Disk Group," on page 70.

Expanding Virtual SAN Cluster Capacity and Performance

If your Virtual SAN cluster is running out of storage capacity or when you notice reduced performance of the cluster, you can expand the cluster for capacity and performance.

- Expand the storage capacity of your cluster either by adding storage devices to existing disk groups or by creating a disk group. Creating disk groups also requires new flash devices for cache. For information about adding devices to disk groups, see "Add Devices to the Disk Group," on page 70. Adding capacity devices without increasing the cache might reduce your cache-to-capacity ratio to an unsupported level. See "Design Considerations for Flash Caching Devices in Virtual SAN," on page 25.
- Improve the cluster performance by adding at least one flash cache and one capacity devices to an existing storage I/O controller or to a new server. Adding one or more servers with additional disk groups has the same impact after Virtual SAN completes a proactive rebalance in the Virtual SAN cluster.

Although compute-only hosts can exist in a Virtual SAN environment and consume capacity from other hosts in the cluster, add uniformly configured hosts.

For best results, add hosts configured with cache and capacity devices. For information about adding devices to disk groups, see"Add Devices to the Disk Group," on page 70.

Add a Host to the Virtual SAN Cluster

You can add an ESXi host to a running Virtual SAN cluster without disrupting any ongoing operations. The host's resources become associated with the cluster.

Prerequisites

- Verify that the resources, including drivers, firmware, and storage I/O controllers, are listed in the VMware Compatibility Guide web site at http://www.vmware.com/resources/compatibility/search.php.
- VMware recommends creating uniformly configured hosts in the Virtual SAN cluster for an even distribution of components and objects across devices in the cluster. However, there might be situations where the cluster becomes unevenly balanced, particularly during maintenance or if you overcommit the capacity of the Virtual SAN datastore with excessive virtual machine deployments.

Procedure

1. Browse to the Virtual SAN cluster in the vSphere Web Client navigator.
2. Right-click the cluster and select **Add Host**.
3. Enter the host name, user name, and password, and click **Next**.
4. View the summary information and click **Next**.
5. Assign an existing or a new license key and click **Next**.
6. (Optional) Enable lockdown mode to prevent remote users from logging directly into the host.

   You can configure this option later by editing the Security Profile in host settings.
7. Select what to do with the host's virtual machines and resource pools.
   - **Put this host's virtual machines in the cluster's root resource pool**

     vCenter Server removes all existing resource pools of the host. The virtual machines in the host's hierarchy are all attached to the root. Because share allocations are relative to a resource pool, you might have to manually change a virtual machine's shares, which destroys the resource pool hierarchy.
   - **Create a resource pool for this host's virtual machines and resource pools**

vCenter Server creates a top-level resource pool that becomes a direct child of the cluster and adds all children of the host to that new resource pool. You can type a name for that new top-level resource pool. The default is **Grafted from <host_name>**.

8 Review the settings and click **Finish**.

The host is added to the cluster.

Configuring Hosts Using Host Profile

When you have multiple hosts in the Virtual SAN cluster, you can reuse the profile of an existing Virtual SAN host and apply its profile settings on the rest of the hosts in the Virtual SAN cluster.

The host profile include information about storage configuration, network configuration, or other characteristics of the host. Typically, if you planing on preparing a cluster with a large number of hosts, such as, 8, 16, 32, or 64 hosts, you should use the host profile feature for adding more than one host at a time to the Virtual SAN cluster.

Prerequisites

- Verify that the host is in maintenance mode.
- Verify that the hardware components, drivers, firmware, and storage I/O controllers are listed in the VMware Compatibility Guide web site at http://www.vmware.com/resources/compatibility/search.php.

Procedure

1 Create a host profile.

a Navigate to the Host profiles view.

b Click the **Extract Profile from a Host** icon ( ).

c Select the host that you intend to use as the reference host and click **Next**.

The selected host must be an active host.

d Type a name and description for the new profile and click **Next**.

e Review the summary information for the new host profile and click **Finish**.

The new profile appears in the Profile list.

2 Attach the host to the intended host profile.

a From the Profile list in the Host Profiles view, select the host profile to be applied to the Virtual SAN host.

b Click the **Attach/Detach Hosts and clusters to a host profile** icon ( ).

c Select the host from the expanded list and click **Attach** to attach the host to the profile.

The host is added to the Attached Entities list.

d Click **Next**.

e Click **Finish** to complete the attachment of the host to the profile.

3 Detach the referenced Virtual SAN host from the host profile.

When a host profile is attached to a cluster, the host or hosts within that cluster are also attached to the host profile. However, when the host profile is detached from the cluster, the association between the host or hosts in the cluster and that of the host profile remains intact.

a From the Profile List in the Host Profiles view, select the host profile to be detached from a host or cluster.

b Click the **Attach/Detach Hosts and clusters to a host profile** icon ( ).

c Select the host or cluster from the expanded list and click **Detach**.

The selected host or cluster is added to the Attached Entities list.

d Click **Detach All** to detach all the listed hosts and clusters from the profile.

e Click **Next**.

f Click **Finish** to complete the attachment of the host to the profile.

4 Verify the compliance of the Virtual SAN host to its attached host profile and determine if any configuration parameters on the host are different from those specified in the host profile.

a Navigate to a host profile.

The **Objects** tab lists all host profiles, the number of hosts attached to that host profile, and the summarized results of the last compliance check.

b Click the **Check Host Profile Compliance** icon ( ).

To see specific details about which parameters differ between the host that failed compliance and the host profile, click the **Monitor** tab and select the Compliance view. Then, expand the object hierarchy and select the failing host. The parameters that differ are displayed in the Compliance window, below the hierarchy.

If compliance fails, use the Remediate action to apply the host profile settings to the host. This action changes all host profile-managed parameters to the values that are contained in the host profile attached to the host.

c To see specific details about which parameters differ between the host that failed compliance and the host profile, click the **Monitor** tab and select the Compliance view.

d Expand the object hierarchy and select the failing host.

The parameters that differ are displayed in the Compliance window, below the hierarchy.

5 Remediate the host to fix compliance errors on the host.

a Select the **Monitor** tab and click **Compliance**.

b Right-click the host or hosts to remediate and select **All vCenter Actions > Host Profiles > Remediate**.

You can update or change the user input parameters for the host profiles policies by customizing the host.

c Click **Next**.

d Review the tasks that are necessary to remediate the host profile and click **Finish**.

The host is part of the Virtual SAN cluster and its resources are accessible to the Virtual SAN cluster. The host can also access all existing Virtual SAN storage I/O policies in the Virtual SAN cluster.

Working with Maintenance Mode

Before you shut down, reboot, or disconnect a host that is a member of a Virtual SAN cluster, you must put the host in maintenance mode.

When working with maintenance mode, consider the following guidelines:

- When you place an ESXi host in maintenance mode, you must select a specific evacuation mechanism.
- When any member host of a Virtual SAN cluster enters maintenance mode, the cluster capacity automatically reduces as the member host no longer contributes storage to the cluster.
- Although a virtual machine's compute resources might not be on the host that is being placed in maintenance mode and the storage for virtual machines might be located anywhere in the cluster.
- The Ensure accessibility mode is faster than the Full data migration mode because the Ensure accessibility migrates only the components from the hosts that are essential for running the virtual machines. When in this mode, if you encounter a failure, the availability of your virtual machine is affected. Selecting the Ensure accessibility mode does not reprotect your data during failure and you might experience unexpected data loss.
- When you select the Full data migration mode, your data is automatically reprotected against a failure, if the resources are available and the Number of failures to tolerate configured to one or more. When in this mode, all components from the host are migrated and depending on the amount of data you have on the host, the migration could take longer. With Full data migration mode, your virtual machines can tolerate failures, even during planned maintenance.
- When working with a three-host cluster, you cannot use the Full data migration mode on a server requiring maintenance. You should consider designing a cluster with four or more hosts for maximum availability.

Before you place a host in maintenance mode, you must verify the following:

- Make sure you have enough hosts and capacity available in the cluster, if are using the Full data migration mode to meet the number of Failures To Tolerate policy requirements.
- Verify that you have enough flash capacity on the remaining hosts to handle any flash read cache reservations. You can run the `vsan.whatif_host_failures` RVC command to analyze the current capacity utilization per host and whether a single host failure could make the cluster run out of space, and impact the cluster capacity, cache reservation, and cluster components. For information about the RVC commands, see the *RVC Command Reference Guide*.
- Make sure that you have enough capacity devices in the remaining hosts to handle stripe width policy requirements, if selected.
- Make sure you have enough free capacity on the remaining hosts to handle the amount of data that must be migrated from the host entering maintenance mode.

Place a Member of Virtual SAN Cluster in Maintenance Mode

Before you shut down, reboot, or disconnect a host that is a member of a Virtual SAN cluster, you must place the host in maintenance mode. When you place a host in maintenance mode, you can select a specific evacuation mechanism.

When any member host of a Virtual SAN cluster enters maintenance mode, the cluster capacity is automatically reduced as the member host no longer contributes capacity to the cluster.

Prerequisites

Verify that your environment has the capabilities required for the option you select.

Procedure

1 Right-click the host and select **Enter Maintenance Mode**.

2 Select one of the evacuation modes and click **OK**.

Option	Description
Ensure accessibility	This is the default option. When you power off or remove the host from the cluster, Virtual SAN ensures that all accessible virtual machines on this host remain accessible. Select this option if you want to take the host out of the cluster temporarily, for example, to install upgrades, and plan to have the host back in the cluster. This option is not appropriate if you want to permanently remove the host from the cluster. Typically, only partial data evacuation is required. However, the virtual machine might no longer be fully compliant to a VM storage policy during evacuation. That means, it might not have access to all its replicas. In case of a failure while the host is in maintenance mode and the Number of Failure to Tolerate is set to one, you might experience data loss in the cluster. NOTE This is the only evacuation mode available if you are working with a three-host cluster or a Virtual SAN cluster configured with three fault domains.
Full data migration	Virtual SAN evacuates all data to other hosts in the cluster, maintains or fixes availability compliance for the affected components, and protects data when sufficient resources exists in the cluster. Select this option if you plan to migrate the host permanently. When evacuating data from the last host in the cluster, make sure you migrate the virtual machines to another datastore and then place the host in maintenance mode. This evacuation mode results in the largest amount of data transfer and consumes the most time and resources. All of the components on the local storage of the selected host will be migrated elsewhere in the cluster so that when the host enters maintenance mode, all virtual machines will have access to their storage components and will still be compliant to their assigned storage policies. NOTE If a virtual machine object that has data on the host is not accessible and is not fully evacuated, the host will unable to enter the maintenance mode.
No data migration	Virtual SAN does not evacuate any data from this host. If you power off or remove the host from the cluster, some virtual machines might become unaccessible.

A cluster with three fault domains has the same restrictions that a three-host cluster has, such as, inability to use the Full data migration mode, reprotect data after failure.

What to do next

You can track the progress of data migration in the cluster. See "Monitor the Resynchronization Tasks in the Virtual SAN Cluster," on page 95.

Managing Fault Domains in Virtual SAN Clusters

If your Virtual SAN cluster spans across multiple racks or blade server chassis in a data center and you want to make sure that your hosts are protected against rack or chassis failure, you can create fault domains and add one or more hosts to it.

A fault domain consists of one or more Virtual SAN hosts grouped together according to their physical location in the data center. When configured, fault domains enable Virtual SAN to tolerate failures of entire physical rack as well as failures of a single host, capacity device, network link or a network switch dedicated to fault domains.

The number of failures your cluster can tolerate depends on the number of failures a virtual machine is provisioned to tolerate. For example, when a virtual machine is configured with Number of failures to tolerate=1 and using multiple fault domains, Virtual SAN can tolerate a single failure of any kind and of any component in a fault domain, including the failure of an entire rack.

When you configure fault domains on a rack and provision a new virtual machine, Virtual SAN ensures that protection objects, such as replicas and witnesses are placed on different fault domains. If, for example, a virtual machine's storage policy is Number of failures to tolerate=n, Virtual SAN requires a minimum of 2*n +1 fault domains in the cluster. When virtual machines are provisioned in a cluster with fault domains using this policy, the copies of the associated virtual machine objects are stored across separate racks.

A minimum of three fault domains are required. For best results, configure four or more fault domains in the cluster. A cluster with three fault domains has the same restrictions that a three host cluster has, such as, inability to reprotect data after failure, and to use the Full data migration mode. For information about designing and sizing fault domains, see "Designing and Sizing Virtual SAN Fault Domains," on page 33.

Consider a scenario where you have a Virtual SAN cluster with 16 hosts. The hosts are spread across 4 racks, that is, 4 hosts per rack. In order to tolerate an entire rack failure, you should create a fault domain for each rack. A cluster of such capacity can be configured to tolerate the Number of failures to tolerate=1. If you want to configure the cluster to allow for virtual machines with Number of failures to tolerate=2, you should configure 5 fault domains in a cluster.

When a rack fails, all resources including the CPU, memory in the rack become unavailable to the cluster. To reduce the impact of a potential rack failure, you should configure fault domains of smaller sizes. This increases the total amount of resource availability in the cluster after a rack failure.

When working with fault domains, follow these guidelines and best practices.

- Configure a minimum of three or more fault domains in the Virtual SAN cluster. For best results, it is recommended that you configure four or more fault domains.
- A host not added to any fault domain, either because it was moved out or was never included in a fault domain, is considered to be its own single host fault domain.
- You do not need to assign every Virtual SAN host to a fault domain. If you decide to use fault domains to protect the Virtual SAN environment, consider creating equal sized fault domains.
- When moved to another cluster, Virtual SAN hosts retain their fault domain assignments.
- When designing a fault domain, it is recommended that you configure fault domains with uniform number of hosts.

  For guidelines about designing fault domains, see "Designing and Sizing Virtual SAN Fault Domains," on page 33.
- You can add any number of hosts to a fault domain. Each fault domain is considered to be one host.

Create a New Fault Domain in Virtual SAN Cluster

To ensure that the virtual machine objects continue to run smoothly during a rack failure, you can group hosts in different fault domains.

When you provision a virtual machine on the cluster with fault domains, Virtual SAN distributes protection components, such as witnesses and replicas of the virtual machine objects across different fault domains. As a result, the Virtual SAN environment becomes capable of tolerating entire rack failures in addition to a single host, storage disk, or network failures.

Prerequisites

- Choose a unique fault domain name. Virtual SAN does not support duplicate fault domain names in a cluster.
- Verify the version of your ESXi hosts. You can only include hosts that are 6.0 or later in fault domains.

- Verify that your Virtual SAN hosts are online. You cannot assign hosts to a fault domain that is offline or unavailable due to hardware configuration issue .

Procedure

1. Browse to the Virtual SAN cluster in the vSphere Web Client navigator.
2. Click the **Manage** tab and click **Settings**.
3. Under Virtual SAN, click **Fault Domains**.
4. Click the **Create a new fault domain** icon.
5. Type the fault domain name.
6. From the **Show** drop-down menu, select **Hosts not in fault domain** to view the list of hosts that are not assigned to a fault domain or select **Show All Hosts** to view all hosts in the cluster.
7. Select one or more hosts to the fault domain.

   A fault domain cannot be empty. You must select at least one host to include in the fault domain.
8. Click **OK.**

   The selected hosts appear in the fault domain.

Move Hosts into Selected Fault Domain

You can move a host into a selected fault domain in the Virtual SAN cluster.

Procedure

1. Browse to the Virtual SAN cluster in the vSphere Web Client navigator.
2. Click the **Manage** tab and click **Settings**.
3. Under Virtual SAN, click **Fault Domains**.
4. Select the fault domain and click the **Move hosts into selected fault domain** icon.
5. From the **Show** drop-down menu at the bottom of the page, select **Hosts not in fault domain** to view the hosts that are available to be added to fault domains or select **Show All Hosts** to view all hosts in the cluster.
6. Select the host that you want to add to the fault domain.
7. Click **OK**.

Move hosts into an Existing Fault Domain

You can move a host to an existing fault domain in the Virtual SAN cluster.

Procedure

1. Browse to the Virtual SAN cluster in the vSphere Web Client navigator.
2. Click the **Manage** tab and click **Settings**
3. Under Virtual SAN, click **Fault Domains**.
4. Select one or more hosts and click the **Move hosts into or out of fault domain** icon.
5. Select the **Move to another fault domain** to move the host to the selected fault domain and click **OK**.

   You cannot configure a fault domain without adding at least one host to it. If the host that you are moving is the only host in the source fault domain, Virtual SAN warns you that the empty fault domain will be deleted from the cluster.

Move Hosts out of a Fault Domain

Depending on your requirement, you can move hosts out of a fault domain.

Prerequisites

Verify that the host is online. You cannot move offline or unavailable hosts in a fault domain.

Procedure

1 Browse to the Virtual SAN cluster in the vSphere Web Client navigator.

2 Click the **Manage** tab and click **Settings**.

3 Under Virtual SAN, click **Fault Domains**.

4 Select the host that you want to move and click the **Move hosts into or out of fault domain** icon.

5 Select the **Move out of fault domains** and click **OK**.

The selected host is no longer part of any fault domain. Any host that is not part of a fault domain is considered to be its own single host fault domain.

What to do next

You can add hosts to fault domains. See "Move hosts into an Existing Fault Domain," on page 84.

Rename a Fault Domain

You can change the name of an existing fault domain in your Virtual SAN cluster.

Procedure

1 Browse to the Virtual SAN cluster in the vSphere Web Client navigator.

2 Click the **Manage** tab and click **Settings**

3 Under Virtual SAN, click **Fault Domains**.

4 Select the fault domain that you want to rename and click the **Rename selected fault domain** icon.

5 Enter a new fault domain name.

6 Click **OK.**

The new name appears in the list of fault domains.

Remove Selected Fault Domains

When you no longer need a fault domain, you can remove it from the Virtual SAN cluster.

Procedure

1 Browse to the Virtual SAN cluster in the vSphere Web Client navigator.

2 Click the **Manage** tab and click **Settings**

3 Under Virtual SAN, click **Fault Domains**.

4 Select the fault domain that you want to delete and click the **Remove selected fault domains** icon.

5 Click **Yes**.

All hosts in the fault domain are moved out and the selected fault domain is deleted from the Virtual SAN cluster. The host is now available as a single independent host. Any host that is not part of a fault domain is considered to be its own single host fault domain.

Using Virtual SAN Policies 10

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

Once assigned, the storage policy requirements are then pushed down to the Virtual SAN layer when a virtual machine is being created. The virtual device is distributed across the Virtual SAN datastore to meet the performance and availability requirements.

When you enable Virtual SAN on a host cluster, a single Virtual SAN datastore is created and a default storage policy is automatically created. Enabling Virtual SAN configures and registers Virtual SAN storage providers.

Virtual SAN storage providers are built-in software components that communicate datastore capabilities to vCenter Server. A storage capability is typically represented by a key-value pair, where the key is a specific property that the datastore can offer and the value is a metric, or a range, that the datastore can provide for a provisioned object, such as a virtual machine home namespace object or a virtual disk. You can also use tags to create user-defined storage capabilities and reference them when defining a storage policy for a virtual machine. For information on how to use and apply tags to datastores, see the *vSphere Storage* documentation.

When you know storage requirements of your virtual machines, you can create a storage policy referencing capabilities that the datastore advertises. You can create several policies to capture different types or classes of requirements.

You can apply these storage policies when you create or edit virtual machines.

NOTE If you do not apply a storage policy to a virtual machine, it will use a default Virtual SAN policy with a Number of failures to tolerate configured to one, a single disk stripe per object, and thin provisioned virtual disk.

Table 10-1. Storage Policy Attributes

Capability	Description
Number of disk stripes per object	The number of capacity devices across which each replica of a virtual machine object is striped. A value higher than 1 might result in better performance, but also results in higher use of system resources. Default value is 1. Maximum value is 12. VMware recommends that you do not change the default striping value. In a hybrid environment, the disk stripes are spread across magnetic disks. In case of all flash configuration, the striping would be across flash devices that make up the capacity layer. Make sure that your Virtual SAN environment has sufficient capacity devices present to accommodate the request.
Flash read cache reservation	Flash capacity reserved as read cache for the virtual machine object. Specified as a percentage of the logical size of the virtual machine disk (vmdk) object. Reserved flash capacity cannot be used by other objects. Unreserved flash is shared fairly among all objects. This option should be used only to address specific performance issues. You do not have to set a reservation to get cache. Setting read cache reservations could cause problem when you move the virtual machine object because the cache reservation settings are always included with the object. The Flash Read Cache Reservation storage policy attribute is not supported for all-flash cluster and you should not use this attribute when defining a VM storage policy. This attribute is only supported for hybrid configurations. Default value is 0%. Maximum value is 100%. **NOTE** By default, Virtual SAN dynamically allocates read cache to storage objects based on demand. This represents the most flexible and the most optimal use of resources. As a result, typically, you do not need to change the default 0 value for this parameter. To increase the value when solving a performance problem, exercise caution. Over-provisioned cache reservation across several virtual machines can cause flash devices space to be wasted on over-reservations and not being available to service the workloads that need the required space at a given time. This might lead to performance degradation.
Number of failures to tolerate	Defines the number of host and device failures a virtual machine object can tolerate. For `n` failures tolerated, `n+1` copies of the virtual machine object are created and `2*n+1` hosts contributing storage are required. When provisioning a virtual machine, if you do not choose a storage policy, Virtual SAN assigns this policy as the default virtual machine storage policy. Default value is 1. Maximum value is 3. If fault domains are configured, `2n+1` fault domains with hosts contributing capacity are required. A host, which is not part of any fault domain is considered as its own single host fault domain. Default value is 1. Maximum value is 3. **NOTE** If you do not want Virtual SAN to protect a single mirror copy of virtual machine objects, you can specify the Number of failures to tolerate=**0**. However, the host might experience unusual delays when entering maintenance mode. The delay occurs because Virtual SAN has to evacuate the object from the host for the maintenance operation to complete successfully. Setting the Number of failures to tolerate=**0** means that your data is unprotected, and you might lose data when the Virtual SAN cluster encounters a device failure. **NOTE** When creating a new storage policy, if you do not specify any value for Number of failures to tolerate, by default, Virtual SAN creates a single mirror copy of the virtual machine objects and tolerates only one failure. However, in the event of a multiple component failures your data might be at risk.

Table 10-1. Storage Policy Attributes (Continued)

Capability	Description
Force provisioning	If the option is set to **Yes**, the object will be provisioned even if the policy specified in the storage policy is not satisfiable by the datastore. Use this parameter in bootstrapping scenarios and during an outage when standard provisioning is not longer possible. The default **No** is acceptable for most production environments. Virtual SAN fails to provision a virtual machine when the policy requirements are not met, however, successfully creates the user-defined storage policy.
Object space reservation	Percentage of the logical size of the virtual machine disk (vmdk) object that should be reserved, or thick provisioned when deploying virtual machines. Default value is 0%. Maximum value is 100%.

When working with virtual machine storage policies, you must understand how the storage capabilities affect the consumption of storage capacity in the Virtual SAN cluster. For detail information about designing and sizing considerations of storage policies, see Chapter 3, "Designing and Sizing a Virtual SAN Cluster," on page 23.

This chapter includes the following topics:

View Virtual SAN Storage Providers

Virtual SAN automatically configures and registers a storage provider for each host in the Virtual SAN cluster.

The Virtual SAN storage providers report a set of underlying storage capabilities to vCenter Server. They also communicate with the Virtual SAN layer to report the storage requirements of the virtual machines. For more information about storage providers, see the *vSphere Storage* documentation.

Virtual SAN registers a separate storage provider for each host in the Virtual SAN cluster, using the following URL:

http://*host_ip*:8080/version.xml

where *host_ip* is the actual IP of the host.

Verify that the storage providers are registered.

Procedure

1 Browse to vCenter Server in the vSphere Web Client navigator.

2 Click the **Manage** tab, and click **Storage Providers**.

The storage providers for Virtual SAN appear on the list. All hosts have a storage provider, but only one is active. Storage providers that belong to other hosts are in standby. If the host that currently has the online storage provider fails, another host will bring its provider online.

NOTE You cannot manually unregister storage providers used by Virtual SAN. If for troubleshooting purposes you need to remove or unregister the Virtual SAN storage providers, remove corresponding hosts form the Virtual SAN cluster and then add the hosts back. Make sure at least one storage provider is in active state.

About the Virtual SAN Default Storage Policy

Virtual SAN requires that the virtual machines deployed on the Virtual SAN datastores are assigned at least one storage policy. When provisioning a virtual machine, if you do not explicitly assign a storage policy to the virtual machine, a generic system defined storage policy, called the Virtual SAN Default Storage Policy, is applied to the virtual machine.

The default policy contains Virtual SAN rulesets and a set of basic storage capabilities, typically used for the placement of virtual machines deployed on Virtual SAN datastores.

Following are the specifications for the Virtual SAN Default Storage Policy.

- Number of failures to tolerate is set to 1.
- Number of disk stripes per object is set to 1.
- Flash read cache reservation or the flash capacity used for read cache is set to zero.
- Object space reservation is set to zero. Setting the Object space reservation to zero means that the virtual disk will be thin provisioned, by default.
- Force Provisioning is set to no.

You can review the configuration settings for the default virtual machine storage policy from the vSphere Web Client client when you navigate to the **VM Storage Policies** > **Virtual SAN Default Storage Policy** > **Manage** > **Rule-Set 1: VSAN**.

For best results, consider creating and using your own VM storage policies, even if the requirements of the policy are same as those defined in the default storage policy. For information about creating a user-defined VM storage policy, see the "Define a Virtual Machine Storage Policy for Virtual SAN," on page 91.

When you assign a user-defined storage policy as the default policy to a datastore, Virtual SAN automatically removes the association to the default storage policy and applies the settings for the user-defined policy on the specified datastore. At any point, you can assign only one virtual machine storage policy as the default policy to the Virtual SAN datastore.

Characteristics

The following characteristics apply to the Virtual SAN Default Storage Policy.

- The default storage policy that VMware provides is applied to all virtual machine objects, if you do not select any other Virtual SAN policy when you have provisioned a virtual machine, that is when the **VM Storage Policy** field is set to **Datastore default** on the Select Storage page. For information about using storage policies, see the *vSphere Storage* documentation.
- The Virtual SAN default policy only applies to Virtual SAN datastores. You cannot apply the default storage policy to non-Virtual SAN datastores, such as, NFS or a VMFS datastore.
- Because the default virtual machine storage policy is compatible with any Virtual SAN datastore in the vCenter Server, you can move your virtual machine objects provisioned with the default policy to any Virtual SAN datastore in the vCenter Server.
- You can clone the default policy and use it as a template to create a user-defined storage policy.
- You can edit the default policy, if you have the StorageProfile.View privilege. You must have at least one Virtual SAN enabled cluster that contains at least one host. VMware highly recommends that you do not edit the settings of the default storage policy.
- You cannot edit the name and description of the default policy, or the Virtual SAN storage provider specification. All other parameters including the policy rules are editable.
- You cannot delete the default policy.

- The default storage policy is applied when the policy that you assign during virtual machine provisioning does not include rules specific to Virtual SAN.

Assign a Default Storage Policy to Virtual SAN Datastores

When you want to reuse a storage policy that matches your requirement, you can assign a user-defined storage policy as the default policy to a datastore.

Prerequisites

Verify that the VM storage policy that you want to assign as the default policy to the Virtual SAN datastore meets the requirements of your virtual machines in the Virtual SAN cluster.

Procedure

1. Browse to the Virtual SAN cluster in the vSphere Web Client navigator.
2. Select one or more Virtual SAN Datastore.
3. Click the **Manage** tab, and click **Settings**.
4. Click **Edit** to change the storage policy associated with the datastore and select the storage policy that you want to assign as the default policy to the Virtual SAN datastore.

   A list of storage policies that are compatible with the Virtual SAN datastore, such as the Virtual SAN Default Storage Policy and user-defined storage policies that have Virtual SAN rulesets defined, appear.
5. Click **OK**.

   The storage policy will be applied as the default policy when you provision new virtual machine without explicitly specifying a storage policy for a datastore.

What to do next

You can define a new storage policy for virtual machines. See "Define a Virtual Machine Storage Policy for Virtual SAN," on page 91 for information.

Define a Virtual Machine Storage Policy for Virtual SAN

When you need to define storage requirements and a type of storage service for a virtual machine and its virtual disks, you create a storage policy. In this policy, you reference storage capabilities that the Virtual SAN datastore supports.

Prerequisites

- Verify that the Virtual SAN storage provider is available. See "View Virtual SAN Storage Providers," on page 89.
- Ensure that the virtual machine storage policies are enabled. For information about storage policies, see the *vSphere Storage* documentation.
- Required privileges: **Profile-driven storage.Profile-driven storage view** and **Profile-driven storage.Profile-driven storage update**

Procedure

1. From the vSphere Web Client Home, click **Policies and Profiles > VM Storage Policies**.
2. Click the **Create a New VM Storage Policy** icon.
3. Select the vCenter Server instance.
4. Type a name and a description for the storage policy.

5 On the Rule-Set 1 window, define the first rule set.

   a Select **VSAN** as the vendor from the Rules Based on Vendor Specific Capabilities drop-box.

     The page expands to show capabilities reported by the Virtual SAN datastore.

   b Add a capability and specify its value.

     Make sure that the values you provide are within the range of values advertised by storage capabilities of the Virtual SAN datastore.

     From the Storage Consumption model, you can review the virtual disk size available for use and the corresponding flash cache and storage capacity including the reserved storage space your virtual machines would potentially consume when you apply the specified storage policy.

   c (Optional) Add tag-based capabilities.

6 (Optional) Add another rule set.

7 Review the list of datastores that match this policy and click **Finish**.

   To be eligible, a datastore does not need to satisfy all rule sets within the policy. The datastore must satisfy at least one rule set and all rules within this set. Make sure that the Virtual SAN datastore meets the requirements set in the storage policy and appears on the list of compatible datastores.

The new policy is added to the list.

What to do next

Apply this policy to a virtual machine and its virtual disks. Virtual SAN will place the virtual machine objects in accordance with the requirements specified in the policy. For information about applying the storage policies to virtual machine objects, see the *vSphere Storage* documentation.

Monitoring Virtual SAN 11

Virtual SAN supports extensive monitoring facilities. You can monitor a Virtual SAN environment from the vSphere Web Client.

You can monitor different objects in a Virtual SAN environment. You can monitor hosts that participate in a Virtual SAN cluster just like any other hosts that are managed by vCenter Server. In addition, you can monitor the Virtual SAN datastore. For information about monitoring objects and storage resources in a Virtual SAN cluster, see the *vSphere Monitoring and Performance* documentation.

This chapter includes the following topics:

Monitor the Virtual SAN Cluster

You can monitor the Virtual SAN cluster and all the objects related to it.

Procedure

1 Browse to the Virtual SAN cluster in the vSphere Web Client navigator.

2 Select the **Monitor** tab and click **Virtual SAN**.

3 Select **Physical Disks** to review all hosts and capacity devices that belong to the cluster.

Information appears about capacity devices, such as total capacity, used capacity, reserved capacity, functional status, physical location, and so on. The physical location information is based on the hardware location of capacity and flash cache devices on Virtual SAN hosts.

4 Select a capacity device and click **Virtual Disks** to review the virtual machines that use the selected device.

You can monitor many aspects of all virtual machine objects, including their current state and whether they are compliant with the storage policies assigned to them.

5 Select the **Manage** tab and click **General** to verify the networking status of the Virtual SAN cluster, verify whether all hosts have joined the cluster, review information about all capacity and flash devices that are claimed by Virtual SAN, and review the total capacity of the Virtual SAN datastore.

In case of any networking issue, the status will be reported on this page. By default, the networking status is normal.

Monitor Virtual Devices in the Virtual SAN Cluster

You can view the status of virtual disks in the Virtual SAN cluster.

When one or more hosts are unable to communicate with the Virtual SAN datastore, the information about virtual devices is not displayed.

Procedure

1 Browse to the Virtual SAN cluster in the vSphere Web Client navigator.

2 Select the **Monitor** tab and click **Virtual SAN**.

3 Select **Virtual Disks** to view all hosts and the corresponding virtual disks that belong to the Virtual SAN cluster, including which hosts, flash cache and capacity devices their components are currently consuming.

4 Select a **VM home** folder from one of the virtual machines and click the **Physical Disk Placement** tab to view device information, such as name, identifier or UUID, and so on.

5 Select a **hard disk** from one of the virtual machines and click the **Physical Disk Placement** tab to view the device information, such as name, identifier or UUID, number of devices used for each virtual machine, and how they are mirrored across hosts.

6 Click the **Compliance Failures** tab to check the compliance status of your virtual machines.

About the Resynchronization Operation in the Virtual SAN Cluster

You can monitor the status of virtual machine objects that are being resynchronized in the Virtual SAN cluster.

When a hardware device, host, or network fails, or if a host is placed into maintenance mode, Virtual SAN initiates resynchronization in the Virtual SAN cluster. However, Virtual SAN might briefly wait for the failed components to come back online before initiating resynchronization tasks in the cluster.

The following events trigger a resynchronization operation in the cluster:

- Editing a virtual machine (VM) storage policy. When you change VM storage policy settings, Virtual SAN might initiate an object recreation process and subsequent resynchronization of the objects.

  Certain policy changes might cause Virtual SAN to create another version of an object and synchronize it with the previous version. When the synchronization is complete, the original object is discarded.

  Virtual SAN ensures that VMs continue to run and are not interrupted by this process. This process might require additional temporary capacity.

- Restarting a host after a failure.
- Recovering hosts from a permanent or long-term failure. If a host is unavailable for more than 60 minutes (by default), Virtual SAN creates copies of data to recover the full policy compliance.
- Evacuating data by using the Full data migration mode before you place a host in maintenance mode.
- Exceeding the utilization of a capacity device. Capacity device utilization in the Virtual SAN cluster that is approaching or has already exceeded the threshold level of 80 percent triggers a resynchronization operation.

Monitor the Resynchronization Tasks in the Virtual SAN Cluster

To evaluate the status of the objects that are being resynchronized, you can monitor the resynchronization tasks in the Virtual SAN cluster that are currently in progress.

Prerequisites

Verify that your Virtual SAN cluster is running ESXi host 6.0 or later.

Procedure

1 Browse to the Virtual SAN cluster in the vSphere Web Client navigator.

2 Select the **Monitor** tab and click **Virtual SAN**.

3 Select **Resyncing Components** to track the progress of resynchronization of virtual machine objects and the number of bytes that are remaining before the resynchronization is complete.

You can also view information about the number of objects that are currently being synchronized in the cluster, the estimated time to finish the resynchronization, the time remaining for the storage objects to fully comply with the assigned storage policy, and so on.

If your cluster is currently facing connectivity issues, the data on the Resyncing Components page might not get refreshed as expected and the fields will reflect inaccurate information.

Monitor Devices that Participate in Virtual SAN Datastores

Verify the status of the devices that back up the Virtual SAN datastore. You can check whether the devices experience any problems.

Procedure

1 Browse to Datastores in the vSphere Web Client navigator.

2 Select the Virtual SAN datastore.

3 Click the **Manage** tab, and click **Settings**.

4 Click **Device Backing**.

5 Select the disk group and review the local devices listed in the Device Details table.

6 If the columns for the devices are not visible, right-click the column heading and select **Show/Hide Columns**.

7 Select the colums that you want to display and click **OK**.

The selected columns are displayed in the Device Details table.

About the Rebalance Operation in the Virtual SAN Cluster

When any capacity device in your cluster reaches above 80 percent, Virtual SAN automatically performs the rebalance operation in the cluster, until the capacity utilization is below the default threshold level of 80 percent.

The purpose of the rebalance operation is to evenly distribute resources across the cluster to maintain consistent cluster performance and availability.

Other operations can initiate the automatic rebalance operation in the cluster:

- Hardware failures that are detected in the Virtual SAN cluster

- Virtual SAN hosts that are placed in maintenance mode with either the **Ensure accessibility** or **Full data migration** option set to migrate data

NOTE To provide enough space for maintenance and reprotection, and to minimize automatic rebalancing events in the Virtual SAN cluster, consider having 30-percent free capacity at all times.

You can also manually initiate the rebalance operation in your Virtual SAN cluster by using the Ruby vSphere Console (RVC) tool. See "Manual Rebalance," on page 96.

Automatic Rebalance

By default, Virtual SAN automatically activates rebalance operation across the Virtual SAN cluster when a capacity device utilization increases more than 80 percent. The automatic rebalance is also enabled when you place a Virtual SAN host in maintenance mode.

Run the following RVC commands to monitor the rebalance operation in the cluster:

- `vsan.check_limits`. Verifies whether the disk space utilization is balanced in the cluster.
- `vsan.whatif_host_failures`. Analyzes the current capacity utilization per host, interprets whether a single host failure can force the cluster to run out of space for reprotection, and analyzes how a host failure might impact cluster capacity, cache reservation, and cluster components.

  The physical capacity usage shown as the command output is the average usage of all devices in the Virtual SAN cluster.
- `vsan.resync_dashboard`. Monitors any rebuild tasks in the cluster.

For information about the RVC command options, see the *RVC Command Reference Guide.*

Manual Rebalance

You can manually start the rebalance operation in the cluster by using the RVC tool.

- `vsan.check_limits`. Verifies whether any capacity device use in the Virtual SAN cluster is approaching the 80 percent threshold limit.
- `vsan.proactive_rebalance [opts]<Path to ClusterComputeResource> --start`. Manually starts the rebalance operation. When you run the command, Virtual SAN scans the cluster for the current distribution of components, and begins to balance the distribution of components in the cluster. Use the command options to specify how long to run the rebalance operation in the cluster, and how much data to move per hour for each Virtual SAN host. For more information about the command options for managing the rebalance operation in the Virtual SAN cluster, see the *RVC Command Reference Guide*.

  Because the rebalance operation generates substantial I/O operations, it can be time-consuming and can affect the performance of virtual machines.

You can set up an alarm that notifies you when the provisioned space reaches a certain threshold. See "Creating a vCenter Server Alarm for a Virtual SAN Event," on page 99.

Using the Virtual SAN Default Alarms

You can use the default Virtual SAN alarms for monitoring hosts, cluster, and existing Virtual SAN licenses.

The default alarms are automatically triggered when the events corresponding to the alarms are activated or if one or all of the conditions specified in the alarms are met. You cannot edit the conditions or delete the default alarms. To configure alarms that are specific to your requirements, you should create custom alarms for Virtual SAN. See"Creating a vCenter Server Alarm for a Virtual SAN Event," on page 99.

The table lists the default Virtual SAN alarms.

Table 11-1. Default Virtual SAN Alarms

Virtual SAN Alarms	Description
Expired Virtual SAN time-limited license	Use to monitor Virtual SAN trial licenses.
Registeration/unregistration of a VASA Vendor provider on a Virtual SAN hosts failed	Use to register or unregister VASA providers on the failed Virtual SAN hosts.
Expired Virtual SAN license	Use to monitor expired Virtual SAN licenses.
Erros occurred on the disk(s) of a Virtual SAN host	Use to monitor errors on Virtual SAN devices.
Virtual SAN Health Service Alarm for Group Test 'Cluster health'	Use to monitor Virtual SAN cluster health.
Virtual SAN Health Service Alarm for Group Test 'Data health'	Use to monitor Virtual SAN cluster data health.
Virtual SAN Health Service Alarm for Group Test 'Limits health'	Use to monitor Virtual SAN cluster limits.
Virtual SAN Health Service Alarm for Group Test 'Network health'	Use to monitor Virtual SAN network health.
Virtual SAN Health Service Alarm for Group Test 'Physical disk health'	Use to monitor the health of physical devices in the cluster.
Virtual SAN Health Service Alarm for Group Test 'Virtual SAN HCL health'	Use to monitor the cluster components to ensure they are using supported hardware, software and drivers.
Virtual SAN Health Service Alarm for Group Test 'software state health'	Use to monitor the health of the software currently used in the cluster.
Virtual SAN Health Service Alarm for Group Test 'Unexpected Virtual SAN health'	Use to monitor any unexpected cluster health issues.
Virtual SAN Health Service Alarm for Group Test 'Virtual SAN CLOMD liveness'	Use to monitor that the CLOMD (Cluster Level Object Manager Daemon), which runs on ESXi hosts and is responsible for data moves and evacuations, is alive or not.
Virtual SAN Health Service Alarm for Group Test 'Virtual SAN cluster partition'	Use to monitor Virtual SAN cluster partition.
Virtual SAN Health Service Alarm for Group Test 'Virtual SAN HCL DB up-to-date'	Use to monitor whether the cluster is using the outdated HCL database.
Virtual SAN Health Service Alarm for Group Test 'Virtual SAN object health'	Use to monitor Virtual SAN object health.
Virtual SAN Health Service Alarm for Group Test 'Virtual SAN Service up-to-date'	Use to monitor Virtual SAN service components to ensure they are up-to-date.

For information about monitoring alarms, events, and editing existing alarm settings, see the *vSphere Monitoring and Performance* documentation.

View Virtual SAN Default Alarms

Use the default Virtual SAN alarms to monitor your cluster, hosts, analyze any new events, and assess the overall cluster health.

Procedure

1 Browse to the Virtual SAN cluster, click **Manage** and then click **Alarm Definitions**.

2 In the search box, type **Virtual SAN** as the search term to filter out the alarms that are specific to Virtual SAN.

Type Virtual SAN Health Service Alarm to search for Virtual SAN health service alarms.

The default Virtual SAN alarms are displayed.

3 From the list of alarms, click on each alarm to view the alarm definition.

Using the VMkernel Observations for Creating Alarms

VMkernel Observations (VOBs) are system events that you can use to set up Virtual SAN alarms for monitoring and troubleshooting performance and networking issues in the Virtual SAN cluster. In Virtual SAN, these events are known as observations.

VMware ESXi Observation IDs for Virtual SAN

Each VOB event is associated with an ID. Before you create a Virtual SAN alarm in the vCenter Server, you must identify an appropriate VOB ID for the Virtual SAN event that you want to create an alert on. You can create alerts in the VMware ESXi Observation Log file, `vobd.log`. For example, you should use the following VOB IDs for creating alerts for any device failures in the cluster.

- `esx.problem.vob.vsan.lsom.diskerror`
- `esx.problem.vob.vsan.pdl.offline`

To review the list of VOB IDs for Virtual SAN, open the `vobd.log` file located on your ESXi host in the `/var/log` directory. The log file contains the following VOB IDs that you can use for creating Virtual SAN alarms.

Table 11-2. VOB IDs for Virtual SAN

VOB ID	Description
esx.audit.vsan.clustering.enabled	Alerts when the Virtual SAN clustering service is enabled.
esx.clear.vob.vsan.pdl.online	Alerts when the Virtual SAN device has come online.
esx.clear.vsan.clustering.enabled	Alerts when the Virtual SAN clustering services is enabled.
esx.clear.vsan.vsan.network.available	Alerts when Virtual SAN has one active network configuration.
esx.clear.vsan.vsan.vmknic.ready	Alerts when a previously reported vmknic has acquired a valid IP.
esx.problem.vob.vsan.lsom.componentthreshold	Alerts when Virtual SAN reaches the near node component count limit.
esx.problem.vob.vsan.lsom.diskerror	Alerts when a Virtual SAN device is in a permanent error state.
esx.problem.vob.vsan.lsom.diskgrouplimit	Alerts when Virtual SAN fails to create a new disk group.
esx.problem.vob.vsan.lsom.disklimit	Alerts when Virtual SAN fails to add devices to a disk group.
esx.problem.vob.vsan.pdl.offline	Alerts when a Virtual SAN device is offline.
esx.problem.vsan.clustering.disabled	Alerts when Virtual SAN clustering services are disabled.
esx.problem.vsan.lsom.congestionthreshold	Alerts when Virtual SAN device memory or SSD congestion has been updated.
esx.problem.vsan.net.not.ready	Alerts when a vmknic is added to Virtual SAN network configuration without a valid IP address. This happens when the Virtual SAN network is not ready.
esx.problem.vsan.net.redundancy.lost	Alerts when the Virtual SAN network configuration does not have the required redundancy.

Table 11-2. VOB IDs for Virtual SAN (Continued)

VOB ID	Description
esx.problem.vsan.no.network.connectivity	Alerts when Virtual SAN does not have existing networking configuration, which is in use.
esx.problem.vsan.vmknic.not.ready	Alerts when a vmknic is added to the Virtual SAN network configuration without a valid IP address.

Creating a vCenter Server Alarm for a Virtual SAN Event

You can create alarms to monitor events on the selected Virtual SAN object, including the cluster, hosts, datastores, networks, and virtual machines.

Prerequisites

Required Privilege: Alarms.Create Alarm or Alarm.Modify Alarm

Procedure

1 Select the vCenter Server object in the inventory that you want to monitor.

2 Select the **Manage** tab > **Alarm Definitions** > click the + icon.

3 Type a name and description for the new alarm.

4 From the **Monitor** drop-down menu, select the object where you want to set up your alarm.

5 Click the **specific event occurring on this object for example VM Power On** and click **Next**.

6 Click **Triggers** to add a Virtual SAN event that will trigger the alarm. The options on the Triggers page change depending on the type of activity you plan to monitor.

7 Click the + icon.

8 Click in the **Event** column, and select an option from the drop-down menu.

9 Click in the **Status** column, and select an option from the drop-down menu.

10 (Optional) Configure additional conditions to be met before the alarm triggers.

   a Click the **Add** icon to add an argument.

   b Click in the **Argument** column, and select an option from the drop-down menu.

   c Click in the **Operator** column, and select an option from the drop-down menu.

   d Click in the **Value** column, and enter a value into the text field.

   You can add more than one argument.

11 Click **Next**.

You selected and configured alarm triggers.

Handling Failures and Troubleshooting Virtual SAN 12

If you encounter problems when using Virtual SAN, you can use troubleshooting topics. The topics help you understand the problem and offer you a workaround, when it is available.

This chapter includes the following topics:

Using esxcli Commands with Virtual SAN

Use esxcli commands to obtain information about Virtual SAN and to troubleshoot your Virtual SAN environment.

The following commands are available:

Command	Description
esxcli vsan network list	Verify which VMkernel adapters are used for Virtual SAN communication.
esxcli vsan storage list	List storage disks that were claimed by Virtual SAN.
esxcli vsan cluster get	Get Virtual SAN cluster information.

Virtual SAN Configuration on an ESXi Host Might Fail

In certain circumstances, the task of configuring Virtual SAN on a particular host might fail.

Problem

An ESXi host that joins a Virtual SAN cluster fails to have Virtual SAN configured.

Cause

If a host does not meet hardware requirements or experiences other problems, Virtual SAN might fail to configure the host. For example, insufficient memory on the host might prevent Virtual SAN from being configured.

Solution

1 Place the host that causes the failure in Maintenance Mode.

2 Move the host out of the Virtual SAN cluster.

3 Resolve the problem that prevent the host to have Virtual SAN configured.

4 Exit Maintenance Mode.

5 Move the host back to the Virtual SAN cluster.

Not Compliant Virtual Machine Objects Do Not Become Compliant Instantly

When you use the **Check Compliance** button, a virtual machine object does not change its status from Not Compliant to Compliant even though Virtual SAN resources have become available and satisfy the virtual machine profile.

Problem

When you use a force provisioning option, you can provision a virtual machine object even when the policy specified in the virtual machine profile is not satisfiable with the resources currently available in the Virtual SAN cluster. The object is created, but remains in the non-compliant status.

Virtual SAN is expected to bring the object into compliance when storage resources in the cluster become available, for example, when you add a host. However, the object's status does not change to compliant immediately after you add resources.

Cause

This occurs because Virtual SAN regulates the pace of the reconfiguration to avoid overloading the system. The amount of time it takes for compliance to be achieved depends on the number of objects in the cluster, the IO load on the cluster and the size of the object in question. In most cases, compliance will be achieved within the reasonable time.

Virtual SAN Cluster Configuration Issues

After you make any changes to Virtual SAN configuration, vCenter Server performs validation checks for Virtual SAN configuration. Validation checks are also performed as a part of a host synchronization process. If vCenter Server detects any configuration problems, it displays error messages.

Problem

A number of error messages indicate that vCenter Server has detected a problem with Virtual SAN configuration.

Solution

Use the following methods to fix Virtual SAN configuration problems.

Table 12-1. Virtual SAN Configuration Errors and Solutions

Virtual SAN Configuration Error	Solution
Host with the VSAN service enabled is not in the vCenter cluster	Add the host to the Virtual SAN cluster. 1 Right-click the host, and select **Move To**. 2 Select the Virtual SAN cluster and click **OK**.
Host is in a VSAN enabled cluster but does not have VSAN service enabled	Verify whether Virtual SAN network is properly configured and enabled on the host. See "Configuring Virtual SAN Network," on page 41.
VSAN network is not configured	Configure Virtual SAN network. See "Configuring Virtual SAN Network," on page 41.

Table 12-1. Virtual SAN Configuration Errors and Solutions (Continued)

Virtual SAN Configuration Error	Solution
Host cannot communicate with all other nodes in the VSAN enabled cluster	Might be caused by network isolation. See "Networking Requirements for Virtual SAN," on page 21 documentation.
Found another host participating in the VSAN service which is not a member of this host's vCenter cluster.	Make sure that the Virtual SAN cluster configuration is correct and all Virtual SAN hosts are in the same subnet. See "Designing the Virtual SAN Network," on page 31.

Handling Failures in Virtual SAN

Virtual SAN handles failures of the storage devices, hosts and network in the cluster according to the severity of the failure. You can diagnose problems in Virtual SAN by observing the performance of the Virtual SAN datastore and network.

Failure Handling in Virtual SAN

Virtual SAN implements mechanisms for indicating failures and rebuilding unavailable data for data protection.

Failure States of Virtual SAN Components

In Virtual SAN, components that have failed can be in absent or degraded state. According to the component state, Virtual SAN uses different approaches for recovering virtual machine data.

Virtual SAN also provides alerts about the type of component failure. See "Using the VMkernel Observations for Creating Alarms," on page 98 and "Using the Virtual SAN Default Alarms," on page 96.

Virtual SAN supports two types of failure states for components:

Table 12-2. Failure States of Components in Virtual SAN

Component Failure State	Description	Recovery	Cause
Degraded	A component is in degraded state if Virtual SAN detects a permanent component failure and assumes that the component is not going to recover to working state.	Virtual SAN starts rebuilding the affected components immediately.	■ Failure of a flash caching device ■ Magnetic or flash capacity device failure ■ Storage controller failure
Absent	A component is in absent state if Virtual SAN detects a temporary component failure where the component might recover and restore its working state.	Virtual SAN starts rebuilding absent components if the they are not available within certain timeout. By default, Virtual SAN starts rebuilding absent components after 60 minutes.	■ Lost network connectivity ■ Failure of a physical network adapter ■ ESXi host failure ■ Unplugged flash caching device ■ Unplugged magnetic disk or flash capacity device

Examine the Failure State of a Component

Use the vSphere Web Client to examine whether a component is in the absent or degraded failure state.

If a failure occurs in the cluster, Virtual SAN marks the components for an object as absent or degraded based on the failure severity.

Procedure

1 In the vSphere Web Client, navigate to the Virtual SAN cluster.

2 On the **Monitor** tab, click **Virtual SAN** and select **Virtual Disks**.

The home directories and virtual disks of the virtual machines in the cluster appear.

3 Select a virtual machine object.

4 On the **Physical Disk Placement** tab, examine the Component State property of the components for the selected object.

If a failure has occurred in the Virtual SAN cluster, the Component State property is equal to Absent or Degraded.

Object States That Indicate Problems in Virtual SAN

Examine the compliance status and the operational state of a virtual machine object to determine how a failure in the cluster affects the virtual machine.

Table 12-3. Object State

Object State Type	Description
Compliance Status	The compliance status of a virtual machine object indicates whether it meets the requirements of the assigned VM storage policy.
Operational State	The operational state of an object can be healthy or unhealthy. It indicates the type and number of failures in the cluster. An object is healthy if an intact replica is available and more than 50 percent of the object's votes are still available. An object is unhealthy if an entire replica is not available or less than 50 percent of the object's votes are unavailable. For example, an object might become unhealthy if a network failure occurs in the cluster and a host becomes isolated.

To determine the overall influence of a failure on a virtual machine, examine the compliance status and the operational state. If the operational state remains healthy although the object is noncompliant, the virtual machine can continue using the Virtual SAN datastore. If the operational state is unhealthy, the virtual machine cannot use the datastore.

Examine the Health of an Object in Virtual SAN

Use the vSphere Web Client to examine whether a virtual machine is healthy. A virtual machine is considered as healthy when a replica of the VM object and more than 50 percent of the votes for an object are available.

Procedure

1 In the vSphere Web Client, navigate to the Virtual SAN cluster.

2 On the **Monitor** tab, click **Virtual SAN** and select **Virtual Disks**.

The home directories and virtual disks of the virtual machines in the cluster appear.

3 For a virtual machine object, examine the value of the Operational State property.

If the Operational State is Unhealthy, the vSphere Web Client indicates the reason for the unhealthy state in brackets.

Examine the Compliance of a Virtual Machine in Virtual SAN

Use the vSphere Web Client to examine whether a virtual machine object is compliant with the assigned VM storage policy.

Procedure

1 Examine the compliance status of a virtual machine.

   a Browse to the virtual machine in the vSphere Web Client navigator.

   b On the **Summary** tab, examine the value of the VM Storage Policy Compliance property under VM Storage Policies.

2 Examine the compliance status of the objects of the virtual machine.

   a In the vSphere Web Client, navigate to the Virtual SAN cluster.

   b On the **Monitor** tab, click **Virtual SAN** and select **Virtual Disks**.

   c Select a virtual machine object.

   d Examine the value of the Compliance Status property for the object. If the Compliance Status is different from Compliant, determine the cause for the noncompliance.

      - Examine the Operational State of the object to verify whether the object is healthy.
      - On the **Compliance Failure** tab, examine which requirements from the VM storage policy that the object cannot satisfy.
      - On the **Physical Disk Placement** tab, examine the state of the object components.

Accessibility of Virtual Machines Upon a Failure in Virtual SAN

If a virtual machine uses Virtual SAN storage, its storage accessibility might change according to the type of failure in the Virtual SAN cluster.

Changes in the accessibility occur when the cluster experiences more failures than the policy for a virtual machine object tolerates.

As a result from a failure in the Virtual SAN cluster, an virtual machine object might become inaccessible. An object is inaccessible if a full replica of the object is not available because the failure affects all replicas, or when less than 50 percent of the object's votes are available because the failure affects a replica and a witness.

According to the type of object that is inaccessible, virtual machines behave in the following ways:

Table 12-4. Inaccessibility of Virtual Machine Objects

Object Type	Virtual Machine State	Virtual Machine Symptoms
VM Home Namespace	■ Inaccessible ■ Orphaned if vCenter Server or the ESXi host cannot access the `.vmx` file of the virtual machine.	The virtual machine process might crash and the virtual machine might be powered off.
VMDK	Inaccessible	The virtual machine remains powered on but the I/O operations on the VMDK are not being executed. After a certain timeout passes, the guest operating system ends the operations.

Virtual machine inaccessibility is not a permanent state. After the underlying issue is resolved, and a full replica and more than 50 percent of the object's votes are restored, the virtual machine automatically becomes accessible again.

Capacity Device Not Accessible in a Virtual SAN Cluster

When a magnetic disk or flash capacity device fails, Virtual SAN evaluates the accessibility of the objects on the device and rebuilds them on another host if space is available and the number of failures to tolerate is defined equal to or greater than 1.

Component Failure State and Accessibility

The Virtual SAN components that reside on the magnetic disk or flash capacity device are marked as degraded.

Behavior of Virtual SAN

Virtual SAN responds to the capacity device failure in the following ways.

Parameter	Behavior
Number of failures to tolerate	If the number of failures to tolerate in the VM storage policy is equal to or greater than 1, the virtual machine objects are still accessible from another ESXi host in the cluster. If resources are available, Virtual SAN starts an automatic reprotection. If the number of failures to tolerate is equal to 0, a virtual machine object is inaccessible if one of the object's components is on the failed capacity device. Restore the virtual machine from a backup.
I/O operations on the capacity device	Virtual SAN stops all running I/O operations for 5-7 seconds until it re-evaluates whether an object is still available without the failed component. If Virtual SAN determines that the object is available, all running I/O operations are resumed.
Rebuilding data	Virtual SAN examines whether the hosts and the capacity devices can satisfy the requirements for space and placement rules for the objects on the failed device or disk group. If such a host with capacity is available, Virtual SAN starts the recovery process immediately because the components are marked as degraded. If resources are available, an automatic reprotect will occur.

A Flash Caching Device Is Not Accessible in a Virtual SAN Cluster

When a flash caching device fails, Virtual SAN evaluates the accessibility of the objects on the disk group that contains the cache device, and rebuilds them on another host if possible and the number of failures to tolerate is equal to or greater than 1.

Component Failure State and Accessibility

Both cache device and capacity devices that reside in the disk group, for example, magnetic disks, are marked as degraded. Virtual SAN interprets the failure of a single flash caching device as a failure of the entire disk group.

Behavior of Virtual SAN

Virtual SAN responds to the failure of a flash caching device in the following way:

Parameter	Behavior
Number of failures to tolerate	If the number of failures to tolerate in the VM storage policy is equal to or greater than 1, the virtual machine objects are still accessible from another ESXi host in the cluster. If resources are available, Virtual SAN starts an automatic reprotection. If the number of failures to tolerate is equal to 0, a virtual machine object is inaccessible if one of the object's components is on the failed disk group.
I/O operations on the disk group	Virtual SAN stops all running I/O operations for 5-7 seconds until it re-evaluates whether an object is still available without the failed component. If Virtual SAN determines that the object is available, all running I/O operations are resumed.
Rebuilding data	Virtual SAN examines whether the hosts and the capacity devices can satisfy the requirements for space and placement rules for the objects on the failed device or disk group. If such a host with capacity is available, Virtual SAN starts the recovery process immediately because the components are marked as degraded.

A Host Is Not Responding in a Virtual SAN Cluster

If a host stops responding because of a failure or reboot of the host, Virtual SAN waits for the host to recover before Virtual SAN rebuilds the components on the host elsewhere in the cluster.

Component Failure State and Accessibility

The Virtual SAN components that reside on the host are marked as absent.

Behavior of Virtual SAN

Virtual SAN responds to the host failure in the following way:

Parameter	Behavior
Number of failures to tolerate	If the number of failures to tolerate in the VM storage policy is equal to or greater than 1, the virtual machine objects are still accessible from another ESXi host in the cluster. If resources are available, Virtual SAN starts an automatic reprotection. If the number of failures to tolerate is equal to 0, a virtual machine object is inaccessible if the object's components is on the failed host.
I/O operations on the host	Virtual SAN stops all running I/O operations for 5-7 seconds until it re-evaluates whether an object is still available without the failed component. If Virtual SAN determines that the object is available, all running I/O operations are resumed.
Rebuilding data	If the host does not rejoin the cluster within 60 minutes, Virtual SAN examines whether some of the other hosts in the cluster can satisfy the requirements for cache, space and placement rules for the objects on the inaccessible host. If such a host is available, Virtual SAN starts the recovery process. If the host rejoins the cluster after 60 minutes and recovery has started, Virtual SAN evaluates whether to continue the recovery or stop it and resynchronize the original components.

Network Connectivity Is Lost in the Virtual SAN Cluster

When the connectivity between the hosts in the cluster is lost, Virtual SAN determines the active partition and rebuilds the components from the isolated partition on the active partition if the connectivity is not restored.

Component Failure State and Accessibility

Virtual SAN determines the partition where more than 50% of the votes of an object are available. The components on the isolated hosts are marked as absent.

Behavior of Virtual SAN

Virtual SAN responds to a network failure in the following way:

Parameter	Behavior
Number of failures to tolerate	If the number of failures to tolerate in the VM storage policy is equal to or greater than 1, the virtual machine objects are still accessible from another ESXi host in the cluster. If resources are available, Virtual SAN starts an automatic reprotection. If the number of failures to tolerate is equal to 0, a virtual machine object is inaccessible if the object's components are on the isolated hosts.
I/O operations on the isolated hosts	Virtual SAN stops all running I/O operations for 5-7 seconds until it re-evaluates whether an object is still available without the failed component. If Virtual SAN determines that the object is available, all running I/O operations are resumed.
Rebuilding data	If the host rejoins the cluster within 60 minutes, Virtual SAN synchronizes the components on the host. If the host does not rejoin the cluster within 60 minutes, Virtual SAN examines whether some of the other hosts in the cluster can satisfy the requirements for cache, space and placement rules for the objects on the inaccessible host. If such a host is available, Virtual SAN starts the recovery process. If the host rejoins the cluster after 60 minutes and recovery has started, Virtual SAN evaluates whether to continue the recovery or stop it and resynchronize the original components.

A Storage Controller Fails in a Virtual SAN Cluster

When a storage controller fails, Virtual SAN evaluates the accessibility of the objects on the disk groups that are attached to the controller and rebuilds them on another host.

Symptoms

If a host contains a single storage controller and multiple disk groups, and all devices in all disk groups are failed, then you might assume that a failure in the common storage controller is the root cause. Examine the VMkernel log messages to determine the nature of the fault.

Component Failure State and Accessibility

When a storage controller fails, the components on the flash caching devices and capacity devices in all disk groups that are connected to the controller are marked as degraded.

If a host contains multiple controllers, and only the devices that are attached to an individual controller are inaccessible, then you might assume that this controller has failed.

Behavior of Virtual SAN

Virtual SAN responds to a storage controller failure in the following way:

Parameter	Behavior
Number of failures to tolerate	If the number of failures to tolerate in the VM storage policy is equal to or greater than 1, the virtual machine objects are still accessible from another ESXi host in the cluster. If resources are available, Virtual SAN starts an automatic reprotection. If the number of failures to tolerate is equal to 0, a virtual machine object is inaccessible if the object's components is on the disk groups that are connected to the storage controller.
Rebuilding data	Virtual SAN examines whether the hosts and the capacity devices can satisfy the requirements for space and placement rules for the objects on the failed device or disk group. If such a host with capacity is available, Virtual SAN starts the recovery process immediately because the components are marked as degraded.

Stretched Cluster Site Fails or Loses Network Connection

A Virtual SAN stretched cluster manages failures that occur due to the loss of a network connection between sites or the temporary loss of one site.

Stretched Cluster Failure Handling

In most cases, the stretched cluster continues to operate during a failure and automatically recovers after the failure is resolved.

Table 12-5. How Stretched Cluster Handles Failures

Type of Failure	Behavior
Network Connection Lost Between Active Sites	If the network connection fails between the two active sites, the witness host and the preferred site continue to service storage operations, and keep data available. When the network connection returns, the two active sites are resynchronized.
Secondary Site Fails or Loses Network Connection	If the secondary site goes offline or becomes isolated from the preferred site and the witness host, the witness host and the preferred site continue to service storage operations, and keep data available. When the secondary site returns to the cluster, the two active sites are resynchronized.
Preferred Site Fails or Loses Network Connection	If the preferred site goes offline or becomes isolated from the secondary site and the witness host, the secondary site continues storage operations as long as it remains connected to the witness host. When the preferred site returns to the cluster, the two active sites are resynchronized.
Witness Host Fails or Loses Network Connection	If the witness host goes offline or becomes isolated from the preferred site or the secondary site, objects become noncompliant but data remains available. VMs that are currently running are not affected.

Troubleshooting Virtual SAN

Examine the performance and accessibility of virtual machines to diagnose problems in the Virtual SAN cluster.

Verify Drivers, Firmware, Storage I/O Controllers Against the *VMware Compatibility Guide*

Use the *VMware Compatibility Guide* to verify whether your hardware components, drivers, and firmware are compatible with Virtual SAN.

Using hardware components, drivers, and firmware that are not compatible with Virtual SAN might cause problems in the operation of the Virtual SAN cluster and the virtual machines running on it.

Procedure

1 Collect information about the drivers, firmware, and storage I/O controllers on the hosts in the cluster.

   a Browse to the host in the vSphere Web Client navigator.

   b On the **Summary** tab, examine the model of the physical machine.

   c On the **Monitor** tab, click **Hardware Status** and examine the storage device configuration on the host.

2 Verify the hardware against the *VMware Compatibility Guide*.

   a Open the *VMware Compatibility Guide* at http://www.vmware.com/resources/compatibility/search.php?deviceCategory=vsan.

   b In the **System / Servers** section, search the guide for the server model.

   c In the **Virtual SAN** section, search the guide for the storage devices and controllers.

   d Verify that the firmware and drivers on the hosts comply with the *VMware Compatibility Guide*.

Examining Performance in a Virtual SAN Cluster

Monitor the performance of virtual machines, hosts, of the Virtual SAN datastore to identify potential storage problems.

Monitor regularly the following performance indicators to identify faults in Virtual SAN storage, for example, by using the performance charts in the vSphere Web Client:

- Datastore. Rate of I/O operations on the aggregated datastore.
- Virtual Machine. I/O operations, memory and CPU usage, network throughput and bandwidth.

For detailed information about using performance data in a Virtual SAN cluster, see the *Virtual SAN 6.0 Troubleshooting Reference Manual*.

Network Misconfiguration Status in a Virtual SAN Cluster

After you enable Virtual SAN on a cluster, the datastore is not assembled correctly because of a detected network misconfiguration.

Problem

After you enable Virtual SAN on a cluster, on the **Summary** tab for the cluster the Network Status for Virtual SAN appears as `Misconfiguration detected`.

Cause

One or more members of the cluster cannot communicate because of either of the following reasons:

- A host in the cluster does not have a VMkernel adapter for Virtual SAN.
- The hosts cannot connect each other in the network.
- Multicast is not enabled on the physical switch.

Solution

Join the members of the cluster to the same network or enable multicast on the physical switch. See "Configuring Virtual SAN Network," on page 41.

Virtual Machine Appears as Noncompliant, Inaccessible or Orphaned in Virtual SAN

The state of a virtual machine that stores data on a Virtual SAN datastore appears as noncompliant, inaccessible or orphaned because of failures in the Virtual SAN cluster.

Problem

A virtual machine on a Virtual SAN datastore is in one of the following states that indicate a fault in the Virtual SAN cluster.

- The virtual machine is non-compliant and the compliance status of some of its object is noncompliant. See "Examine the Compliance of a Virtual Machine in Virtual SAN," on page 105.

- The virtual machine object is inaccessible or orphaned. See "Examine the Failure State of a Component," on page 103.

If an object replica is still available on another host, Virtual SAN forwards the I/O operations of the virtual machine to the replica.

Cause

If the object of the virtual machine can no longer satisfy the requirement of the assigned VM storage policy, Virtual SAN considers it noncompliant. For example, a host might temporarily lose connectivity. See "Object States That Indicate Problems in Virtual SAN," on page 104.

If Virtual SAN cannot locate a full replica or more than 50 percent of the votes for the object, the virtual machine becomes inaccessible. If a Virtual SAN detects that the `.vmx` file is not accessible because the VM Home Namespace is corrupted, the virtual machine becomes orphaned. See "Accessibility of Virtual Machines Upon a Failure in Virtual SAN," on page 105.

Solution

If the cluster contains enough resources, Virtual SAN automatically recovers the corrupted objects if the failure is permanent.

If the cluster does not have enough resources to rebuild the corrupted objects, extend the space in the cluster. See "Expanding Virtual SAN Cluster Capacity and Performance," on page 78 and "Add a Host to the Virtual SAN Cluster," on page 78.

Attempt to Create a Virtual Machine on Virtual SAN Fails

When you try to deploy a virtual machine in a Virtual SAN cluster, the operation fails with an error that the virtual machine files cannot be created.

Problem

The operation for creating a virtual machine fails with an error status: `Cannot complete file creation operation`.

Cause

The deployment of a virtual machine on Virtual SAN might fail for several reasons.

- Virtual SAN cannot allocate space for the virtual machine storage policies and virtual machine objects. Such a failure might occur if the datastore does not have enough usable capacity, for example, if a physical disk is temporarily disconnected from the host.
- The virtual machine has very large virtual disks and the hosts in the cluster cannot provide storage for them based on the placement rules in the VM storage policy

  For example, if the number of failures to tolerate in the VM storage policy is equal to 1, Virtual SAN must store two replicas of a virtual disk in the cluster, each replica on a different host. The datastore might have this space after aggregating the free space on all hosts in the cluster. However, no two hosts can be available in the cluster, each providing enough space to store a separate replica of the virtual disk.

  Virtual SAN does not move components between hosts or disks groups to free space for a new replica, even though the cluster might contain enough space for provisioning the new virtual machine.

Solution

- Verify the state of the capacity devices in the cluster.
  a In the vSphere Web Client, navigate to the Virtual SAN cluster.
  b On the **Monitor** tab, click **Virtual SAN** and select **Physical Disks**.
  c Examine the capacity and health status of the devices on the hosts in the cluster.

Stretched Cluster Configuration Error When Adding a Host

Before adding new hosts to a stretched cluster, all current hosts must be connected. If a current host is disconnected, the configuration of the new host is incomplete.

Problem

After you add a new host to a stretched cluster in which some hosts are disconnected, on the Summary tab for the cluster the Configuration Status for Virtual SAN appears as `Unicast agent unset on host`.

Cause

When a new host joins a stretched cluster, Virtual SAN must update the configuration on all hosts in the cluster. If one or more hosts are disconnected from the vCenter Server, the update fails. The new host successfully joins the cluster, but its configuration is incomplete.

Solution

Verify all hosts are connected to vCenter Server, and click on the link provided in the Configuration Status message to update the configuration of the new host.

If you cannot rejoin the disconnected host, remove the disconnected host from the cluster, and click on the link provided in the Configuration Status message to update the configuration of the new host.

Stretched Cluster Configuration Error When Using RVC to Add a Host

If you use the RVC tool to add a new host to a stretched cluster, the configuration of the new host is incomplete.

Problem

After you use the RVC tool to add a new host to a stretched cluster, on the Summary tab for the cluster the Configuration Status for Virtual SAN appears as `Unicast agent unset on host`.

Cause

When a new host joins a stretched cluster, Virtual SAN must update the configuration on all hosts in the cluster. If you use the RVC tool to add the host, the update does not occur. The new host successfully joins the cluster, but its configuration is incomplete.

Solution

Verify all hosts are connected to vCenter Server, and click on the link provided in the Configuration Status message to update the configuration of the new host.

Cannot Add or Remove the Witness Host in a Stretched Cluster

Before adding or removing the witness host in a stretched cluster, all current hosts must be connected. If a current host is disconnected, you cannot add or remove the witness host.

Problem

When you add or remove a witness host in a stretched cluster in which some hosts are disconnected, the operation fails with an error status: `The operation is not allowed in the current state. Not all hosts in the cluster are connected to Virtual Center`.

Cause

When the witness host joins or leaves a stretched cluster, Virtual SAN must update the configuration on all hosts in the cluster. If one or more hosts are disconnected from the vCenter Server, the witness host cannot be added or removed.

Solution

Verify all hosts are connected to vCenter Server, and retry the operation. If you cannot rejoin the disconnected host, remove the disconnected host from the cluster, and then you can add or remove the witness host.

Replacing Existing Hardware Components

Under certain conditions, you must replace hardware components, drivers, firmware, and storage I/O controllers in the Virtual SAN cluster.

In Virtual SAN, you should replace hardware devices when you encounter failures or if you must upgrade your cluster.

Replace a Flash Caching Device on a Host

You should replace a flash caching device if you detect a failure or when you must upgrade it. Before you physically unplug a flash device from the host, you must manually remove the device from Virtual SAN.

CAUTION If you decommission the flash caching device without removing it from Virtual SAN first, Virtual SAN uses smaller amount of cache than expected. As a result, the cluster performance becomes degraded.

When you replace a flash caching device, the virtual machines on the disk group become inaccessible and the components on the group are marked as degraded. See "A Flash Caching Device Is Not Accessible in a Virtual SAN Cluster," on page 106.

Prerequisites

- Verify that the storage controllers on the hosts are configured in passthrough mode and support the hot-plug feature.

  If the storage controllers are configured in RAID 0 mode, see the vendor documentation for information about adding and removing devices.

- If you upgrade the flash caching device, verify the following requirements:
  - If you upgrade the flash caching device, verify that the cluster contains enough space to migrate the data from the disk group that is associated with the flash device.
  - Place the host in maintenance mode. See "Place a Member of Virtual SAN Cluster in Maintenance Mode," on page 81.

Procedure

1 In the vSphere Web Client, navigate to the Virtual SAN cluster.

2 On the **Manage** tab, click **Settings** and select **Disk Management** under Virtual SAN.

3 Select the disk group that contains the device that you replace.

4 Select the flash caching device and click **Remove selected disk(s) from disk group**.

After the flash caching device is deleted from the Virtual SAN cluster, the cluster details reflect the current cluster capacity and configuration settings. Virtual SAN discards the disk group memberships, deletes partitions, and removes stale data from all devices.

What to do next

1 Add a new device to the host.

The host automatically detects the device.

2 If the host is unable to detect the device, perform a device rescan

Replace a Capacity Device

You should replace a flash capacity device or a magnetic disk if you detect a failure or when you upgrade it. Before you physically remove the device from the host, you must manually delete the device from Virtual SAN.

When you unplug a capacity device without removing it from the Virtual SAN cluster, the virtual machines on the disk group become inaccessible and the components on the group are marked as absent.

If the capacity device fails, the virtual machines become inaccessible and the components on the group are marked as degraded. See "Capacity Device Not Accessible in a Virtual SAN Cluster," on page 106.

Prerequisites

- Verify that the storage controllers on the hosts are configured in passthrough mode and support the hot-plug feature.

  If the storage controllers are configured in RAID 0 mode, see the vendor documentation for information about adding and removing devices.

- If you upgrade the capacity device, verify the following requirements:
  - Verify that the cluster contains enough space to migrate the data from the capacity device.
  - Place the host in maintenance mode. See "Place a Member of Virtual SAN Cluster in Maintenance Mode," on page 81.

Procedure

1 In the vSphere Web Client, navigate to the Virtual SAN cluster.

2 On the **Manage** tab, click **Settings** and select **Disk Management** under Virtual SAN.

3 Select the disk group that contains the device that you replace.

4 Select the flash capacity device or magnetic disk, and click **Remove selected disk(s) from disk group**.

What to do next

1 Add a new device to the host.

The host automatically detects the device.

2 If the host is unable to detect the device, perform a device rescan

Remove a Device from a Host by Using an ESXCLI Command

If you detect a failed storage device or if you upgrade a device, you can manually remove it from a host by using an ESXCLI command.

If you remove a flash caching device, Virtual SAN deletes the disk group that is associated with the flash device and all its member devices.

Prerequisites

Verify that the storage controllers on the hosts are configured in passthrough mode and support the hot-plug feature.

If the storage controllers are configured in RAID 0 mode, see the vendor documentation for information about adding and removing devices.

Procedure

1 Open an SSH connection to the ESXi host.

2 To identify the device ID of the failed device, run this command and learn the device ID from the output.

```
esxcli vsan storage list
```

3 To remove the device from Virtual SAN, run this command.

```
esxcli vsan storage remove –d device_id
```

What to do next

1 Add a new device to the host.

The host automatically detects the device.

2 If the host is unable to detect the device, perform a device rescan

Shutting Down the Virtual SAN Cluster

When necessary, you can shut down the entire Virtual SAN cluster.

If you plan to shut down the Virtual SAN cluster, you do not need to manually disable Virtual SAN on the cluster.

Procedure

1 Power off all virtual machines (VMs) running in the Virtual SAN cluster.

2 Place the ESXi hosts in maintenance mode.

a Right-click the host and select **Enter Maintenance Mode**.

b Select the **No data migration** evacuation mode and click **OK**.

3 In the Confirm Maintenance Mode wizard, deselect the **Move powered-off and suspended virtual machines to other hosts in the cluster** check box.

When you deselect this check box, Virtual SAN does not migrate the VMs to other hosts. If you plan to shut down the entire cluster and put all hosts in maintenance mode, you do not need to move or migrate the VM storage objects to other hosts or devices in the cluster.

4 Power off the hosts after they have successfully entered maintenance mode.

5 Power on the ESXi hosts.

   a On the physical box where ESXi is installed, press the power button until the power-on sequence begins.

      The ESXi host starts, locates its VMs, and functions normally.

   After you power on the hosts, the Virtual SAN is automatically recreated. ESXi

   If you browse to the ESXi host and click **Summary**, you might see that the Network Status of the cluster appears as `Misconfiguration detected`.

   You can ignore the status message if you did not make network configuration changes and the Virtual SAN cluster was working as expected before you shut down the cluster. The message disappears after at least three hosts join the cluster.

6 Take the hosts out of maintenance mode.

7 Restart the VMs.

Index

W