vSAN Frequently Asked Questions

Common questions and answers for vSAN

May 5, 2026

vmware by Broadcom

Table of Contents

General Information

What are the hardware requirements when running vSAN?

The hardware requirements will depend on the capacity and performance requirements for your environment. The newer, more powerful and efficient vSAN Express Storage Architecture (ESA) has different hardware requirements than the original storage architecture (OSA). For the latest minimum hardware requirements for vSAN ReadyNodes, see the "vSAN ESA ReadyNode Hardware Guidance" document. As of November 2025, vSAN ReadyNodes profiles certified for both vSAN HCI clusters and vSAN storage clusters have been changed, with much lower hardware minimums. See the post: "Driving Down Storage Costs with Lower Hardware Requirements for vSAN" for more details.

For a standard, single site topology, vSAN typically requires a minimum of three hosts, and can support as many as 64 hosts in a single cluster. Although, it does have a "2-Node" deployment option that is more appropriate for remote/edge environments. This will consist of two hosts storing the data resiliently, and a third host at a central location that helps determine availability of the data.

What are the typical hardware deployment options available?

vSAN runs on commodity servers running x86 processors. Hardware componentry that has been certified to work with vSAN will be available in the form of vSAN ReadyNodes. These are preconfigured hosts from your favorite server OEM vendor and purchased with a single SKU. They fall into three categories:

- **vSAN HCI ReadyNodes.** These are intended for use as an aggregated vSAN HCI cluster. Three profile sizes are designed to offer different performance capabilities.
- **vSAN Storage Cluster ReadyNodes.** These are intended for use as a vSAN storage cluster. Three profile sizes are designed to offer different performance capabilities
- **vSAN Cyber Recovery ReadyNodes.** These are intended for using a clean room cluster for cyber recovery. Three profile sizes are designed to offer different performance capabilities.

You can purchase the same hardware using discrete hardware components from these same OEM vendors. These are known as ReadyNode Emulated configurations.

Your existing servers running vSphere may be eligible for retrofitting that would turn it into a fully supported vSAN host. For more information, see the post: "The 2026 Structural Supply Crisis: Why VMware Cloud Foundation Is The Answer to the 2026 Hardware Crunch" and the "Repurposing ESX Servers for VMware vSAN" FAQs document. The post "Cost-Efficient VMware vSAN ReadyNodes Certified for Cyber Recovery Deployments" also details the very latest ReadyNodes certified for cyber recovery environments.

How much memory is required for a vSAN host?

The requirements will vary depending on workload demands, performance expectations of the workloads, and number of storage devices claimed by vSAN in a host. As of November 2025, vSAN ReadyNodes profiles certified for both vSAN HCI clusters and vSAN storage clusters have been changed, with much lower hardware minimums. See the post: "Driving Down Storage Costs with Lower Hardware Requirements for vSAN" for more details. For the latest minimum hardware requirements for vSAN ReadyNodes, see the "vSAN ESA ReadyNode Hardware Guidance" document. The amount of memory used by vSAN is dynamic. In production you may find that it is using substantially less than the amount required. When a vSAN cluster is running, one can view the amount of memory used by vSAN by highlighting the cluster, clicking **Monitor > vSAN > Support > Performance for Support > Memory > vSAN Memory**. Remember that hosts configured in a vSAN HCI cluster, as well as vSphere clusters mounting the datastore of a vSAN storage cluster can use VMware vSphere Advanced Memory Tiering. This will help you provide more memory resources to your guest workloads, and reduce your expenditures on costly RAM.

How many storage devices can I use in a vSAN host? How many should I use?

In general, vSAN ESA supports from 1 to 24 storage devices per host – with a minimum of 2 storage devices per host for storage clusters. Depending on the circumstances, the recommended number of storage devices used in a host will be higher. **Configuring hosts with 6 or more storage devices accomplishes several objectives**:

- **Ensures sufficient device throughput for the server.** The bus for each physical NVMe storage device has maximum throughput. While this does increase as new generations of PCIe devices are introduced (e.g. PCIe Gen 5), having an abnormally low number of storage devices may not allow vSAN ESA's storage stack inside the host to run at its optimal level of performance. This is also noted on the post: "Driving Down Storage Costs with Lower Hardware Requirements for vSAN."
- **Ensures a host can use the maximum vSAN component count.** vSAN ESA not only has a maximum component count per host (27,000), but per storage device (3,000 data components + 3,000 metadata components). In some circumstances, this per-device component limit may inhibit full storage utilization in configurations with fewer than 6 devices per host, especially if they are extremely high-density storage devices. For more information, see KB ID: 315533.
- **Provides enough storage devices for secondary level of resilience with 2-Node clusters.** These hosts must have at least three storage devices per host to provide a secondary level of resilience in a 2-Node environment.

The vSAN ESA ReadyNode Hardware Guidance document will show the "Maximum storage devices" for the respective ReadyNode profile. The lower performing, more value-focused profiles ("SM" and "MED") will show a number fewer than 24 devices. This is simply because the hosts have a lower amount of RAM. If you wish to populate a lower performance ReadyNode with more storage devices than the maximum listed for that ReadyNode profile, increase the RAM to the next ReadyNode accordingly.

What are the processor requirements for a vSAN host?

The requirements will vary depending on workload demands and performance expectations of those workloads. For the latest minimum hardware requirements for vSAN ReadyNodes, see the "vSAN ESA ReadyNode Hardware Guidance" document.

How are disks aggregated together by vSAN?

The vSAN Express Storage Architecture (ESA) uses a concept a "storage pool." This simply represents all devices in a host that have been claimed for the purpose of being used by the vSAN ESA. In vSAN ESA, all storage devices contribute to both capacity and performance. There is no dedicated caching or capacity tier, nor is there a funneling of I/O through a disk group. The design of vSAN ESA allows it to no longer need the concept of a disk group, and as the result offers far better performance, and improved durability of data. See the post: "The Impact of a Storage Device Failure in vSAN ESA versus OSA" for more information.

Are caching devices needed?

The newer vSAN Express Storage Architecture (ESA) does not use distinct tiers. All devices claimed by vSAN ESA contribute to performance and capacity.

Can I use existing storage arrays (block or file) in the same cluster as a vSAN cluster?

Yes, vSphere can access and use traditional VMFS and NFS datastores alongside vSAN, all in the same cluster. In most cases, vSphere Storage vMotion can be used to migrate VMs between these various datastore types. This feature makes it easy to migrate existing workloads when there is a need to perform maintenance or retirement of an older storage solution.

How can I size a vSAN cluster so that it meets my capacity and performance requirements?

Sizing of both performance and capacity was historically achieved through the vSAN ReadyNode Sizer. Paired with new simplified ReadyNodes with categories that only define performance, you can specify whatever capacity required for the

respective performance you desire. Tools such as Live Optics and RVTools can help you assess a current environment to begin the sizing exercise.

What can I change in a vSAN ReadyNode?

vSAN ReadyNodes are extremely flexible in their ability to be tailored to meet the requirements of your environment. For more information, see "What you can (and cannot) change in a vSAN ReadyNode." As of November 2025, vSAN ReadyNodes profiles certified for both vSAN HCI clusters and vSAN storage clusters have been changed, with much lower hardware minimums. See the post: "Driving Down Storage Costs with Lower Hardware Requirements for vSAN" for more details.

Can I add a host that does not have local storage devices to a vSAN cluster?

A host with no local storage can be added to a vSAN cluster but is not recommended. **Remember that vSAN is a cluster-based storage solution,** which defines a boundary of physical resources that can be used by a set of applications or data, so reasonable levels of resource symmetry of your compute, memory, and storage resources are ideal. Always strive to use uniformly configured hosts for vSAN deployments. While compute only hosts can technically exist in a vSAN cluster, VMware supports vSAN clusters with asymmetrical host configurations, VMware does not recommend having significantly unbalanced cluster configurations, and thus minimize significant levels of asymmetry. This will help prevent potential availability and capacity issues during failure conditions and minimize deviation in performance. See the post: "Asymmetrical vSAN Clusters - What is Allowed, and What is Smart" for more information.

Can a vCenter Server Appliance (VCSA) be installed on a single host on a new cluster?

Yes. VCSA deployment wizard includes the ability to claim disks and turn on vSAN on a single host. This enables administrators to deploy vCenter Server to a new environment where vSAN will be the only datastore. In vSAN for VCF 9.1, this becomes even easier because Auto-RAID will take care of the storage policy handling automatically during this initial bootstrap process. Previously, the "Force Provisioning" Rule needed to be temporarily used.

Does vSAN support VMware Advanced Memory Tiering?

Yes! vSAN hosts can be configured with NVMe devices dedicated to use with memory tiering. This can help reduce memory costs on your vSAN hosts by automatically offloading cold memory pages to a memory tier. It can be used in **vSAN HCI clusters, or vSAN compute clusters** (aka vSphere clusters) mounting a datastore from a vSAN storage cluster. Since memory tiering offloads active memory from VM instances, **using memory tiering within the hosts that make up a vSAN storage cluster will provide no meaningful benefit.** For more information on memory tiering, see the Advanced Memory Tiering Resource Hub.

Are there any vSphere features that are not supported with vSAN?

Nearly all vSphere features such as VMware vSphere vMotion, VMware vSphere Distributed Resource Scheduler (DRS), VMware vSphere High Availability (HA), VMware vSphere Network I/O Control (NIOC), and VMware vSphere Replication are compatible and supported with vSAN. VMware vSphere Fault Tolerance is supported for VMs running on vSAN except for stretched clusters.

The following vSphere features are not supported with vSAN:

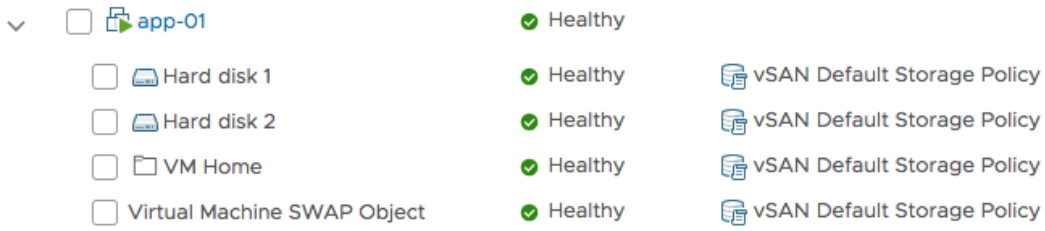
- VMware vSphere Distributed Power Management
- VMware vSphere Storage DRS VMware vSphere Storage I/O Control

Can I share a single vSAN datastore across multiple vSAN and vSphere clusters?

Yes. By default, a vSAN datastore is directly accessible only by the hosts and VMs in the vSAN cluster. When using vSAN's disaggregated offering, the datastore from one vSAN HCI cluster can be mounted to another vSAN HCI cluster. This is known as "vSAN HCI with Datastore Sharing" and typically used in an ad-hoc way to borrow storage from one cluster to another. It is not typically used as a part of a greenfield design. vSAN storage clusters are a vSAN deployment option that **provides a centralized shared storage solution** for your vSphere clusters and augment storage for your vSAN HCI clusters.

How does vSAN store objects such as VM configuration files and virtual disks?

vSAN stores data in a way that is very analogous to an object store. Items such as a VM's configuration (VMX) and virtual disks (VMDKs) are stored as objects. An object consists of one or more components. The size and number of components depend on several factors such as the size of the object and the storage policy used. The following figure shows common virtual machine objects.



Each object commonly consists of multiple components. Components are just an implementation detail, and not a manageable entity in vSAN. For more detail, see the blog post: vSAN Objects and Components Revisited.

When a VM is migrated to another host, are the VM's objects migrated with the VM?

No. The concept implied in the question is often referred to as "data locality". vSAN does not require host-based data locality to achieve excellent performance. vSAN does not tax the vSAN backend network by moving multiple gigabytes of data every time a VM is migrated to another host. There is no need to do this considering modern 10Gb and 25/100Gb networking paired with high performing storage devices.

Where can I find guidance on vSphere boot devices for hosts in a vSAN cluster?

The vSAN Design guide will provide more information on recommended boot devices. Using SD cards and USB devices as boot devices for the hypervisor is a trend that is falling out of favor, and will have limited support in future editions of vSphere. Using a device (such as an SSD, M.2 or BOSS card, etc.) that offers true persistent storage will allow crash dumps and logging to occur on the device, and is a much more useful and reliable configuration.

Should I create one really large vSAN cluster, or break that up into smaller vSAN clusters?

This topic is covered in the "vSAN Cluster Design – Large Clusters versus Small Clusters" document.

How is cost modeling different with vSAN versus traditional three-tier architectures?

One of the most compelling traits of vSAN is the incremental and linear scaling it possesses, and the smooth and predictable cost model. These two traits have historically been a challenge for data centers, because incremental growth can mean large capital expenditures when the environment reaches various resource thresholds. Traditional monolithic arrays, and the storage fabric that they run on, become much more costly as the demand of that shared storage unit increases. vSAN's architecture is different. A cluster serving your workloads can easily be scaled out one host at a time or even scaled up using more or higher density storage devices that can be purchased as a direct expense, as opposed to a large capital expenditure. All of this can be done easily in a non-disruptive manner that is prescriptive and predictable.

Recent TCO modeling has demonstrated that vSAN provides storage at a lower cost than other solutions. For more information, see the post: "Lower Storage TCO by 30% with VMware vSAN" and "Save Costs and Scale Efficiently with vSAN Deduplication in VMware Cloud Foundation 9.0." This is even better with vSAN in VCF 9.1 with global deduplication and enhanced compression.

Express Storage Architecture (ESA)

What is the vSAN Express Storage Architecture?

The vSAN Express Storage Architecture (ESA) is the newest architecture in vSAN that is designed to process and store data with all new levels of efficiency, scalability, and performance. This optional architecture is optimized to exploit the full

capabilities of the very latest in hardware. It was introduced in vSAN 8, and up to and including vSAN in VCF 9.1. It can be selected at the time of creating a cluster.

Does vSAN 9.1 include the Original Storage Architecture found in past editions of vSAN?

Yes! vSAN includes the new Express Storage Architecture (ESA) and the Original Storage Architecture (OSA). The OSA is the architecture that will be used for all in-place cluster upgrades, and new cluster installations using hardware that is not qualified for use with the ESA. The ESA can be used for new cluster installations using qualified hardware. The ESA has a much lower TCO, especially given recent announcements of a new entry-level ReadyNode classification, and the support of Read-Intensive storage devices.

Won't the use of NVMe-based storage devices make ESA more expensive?

No. The architecture and abilities of vSAN ESA will often make all NVMe-based configurations more affordable than clusters running the vSAN OSA using SAS devices. The drastically better TCO comes in multiple forms: 1.) Elimination of purchasing caching/buffering devices and storage controllers. 2.) Running space efficient erasure codes without any performance compromise and do so on clusters as small as three hosts. 3.) Better data compression, 4.) Much better deduplication, and 5.) Support for high-density QLC devices for some workloads. Items #1 & #2 are often enough to make running the ESA more cost efficient per Terabyte than running a similar cluster configured with SAS devices and running the vSAN OSA.

Does vSAN ESA use dedicated storage controllers?

No. **vSAN ESA only supports NVMe based flash devices directly connected to the PCIe slots in a server.** A dedicated storage controller like those found with SAS and SATA devices is not supported. NVMe devices have their own embedded controller, and for vSAN ESA, they should be directly attached to a PCIe slot in the server. (See KB 314305 for more information.) Some server vendors may inadvertently use a controller that supports multiple bus interfaces, known as a tri-mode controller. The use of tri-mode controllers (or any other external controller) when using NVMe storage devices is not supported in vSAN ESA, and will cause significant performance and availability issues as a result.

My vSAN cluster already runs all NVMe devices. Why should I consider running ESA?

Past editions of vSAN have supported the use of all NVMe-based storage devices. NVMe-based storage devices are certainly much higher performing than their SAS and SATA counterparts. While the original storage architecture (OSA) in vSAN could provide a fast storage platform, the ESA was built with these next generation devices in mind. When using vSAN ReadyNodes approved for the ESA, using ESA will be able to exploit the full potential of these storage devices, offering near device-level performance and consistency while improving operational simplicity, and driving down TCO.

There is a lot of talk about "efficiency" with ESA. What does this mean, and why is it so important?

vSAN ESA is designed for the capabilities of storage devices of today and tomorrow. All new technologies are emerging that may dramatically increase the capacity of a storage device. Even if a storage system has an extremely fast data path, a system must use a minimal amount of CPU resources per I/O, otherwise it could run out of CPU resources during high loads. Increased storage densities and new storage techniques can also place a burden in scaling metadata - the data about the data. The vSAN ESA was carefully designed to address these challenges, and as a result, "efficiency" may be one of the most compelling aspects of the vSAN ESA, and what it brings for our customers. An efficient system allows you to do more with what you already have.

Is the Original Storage Architecture (OSA) going away?

No, not in the foreseeable future. We recognize many of our customers have invested heavily in a wide variety of hardware, and customers can continue to use the vSAN OSA with confidence for those configurations. vSAN OSA will generally receive new management improvements that are applicable to both architectures. Continuing to upgrade your clusters to the latest version of vSAN using the OSA is a great way to use your **existing hardware** most effectively. Eventually, availability of hardware components certified for OSA may become more challenging due to industry trends and lack of certification from OEM vendors.

New clusters should always be designed for ESA, which will drive down TCO while benefiting from all the capabilities that vSAN ESA provides. vSAN **ESA can deliver new capabilities and levels of performance and efficiency that are simply not possible in with the OSA**.

Will vSAN ESA support the use of spinning media in a hybrid configuration?

No. vSAN ESA supports high-performing, NVMe-based flash storage devices in vSAN ReadyNodes approved for the ESA. Spinning media is simply not able to support the capabilities that make the ESA so special.

What do I need to run the vSAN Express Storage Architecture in my environment?

The Express Storage Architecture (ESA) can be used with vSAN ReadyNodes or emulated ReadyNodes approved for use with vSAN ESA. Servers must be vSAN ReadyNodes approved for use with the ESA. As shown on "vSAN ESA ReadyNode Hardware Guidance" there are approved ReadyNode when deployed as an aggregated vSAN HCI cluster, disaggregated vSAN storage cluster and Cyber Recovery clusters.

The ReadyNodes can also be "**emulated**" which means they are not purchased as an official ReadyNode using a single SKU, but are built using the same OEM server hardware and meets the minimum ReadyNode requirements. For more information, see the post "Support for ReadyNode Emulated Configurations in vSAN ESA." Some existing vSphere clusters may even be retrofitted to act as a fully supported vSAN host. See "Repurposing ESX Servers for VMware vSAN" for more information.

Can vSAN ESA support different storage device sizes in the same host?

Yes. While customers should strive for a relatively symmetrical cluster, we understand that market conditions over time can make purchasing storage devices of the same capacity challenging. The recommendations for cluster symmetry are very similar to using the original storage architecture. For more information, see the post: Asymmetrical vSAN Clusters - What is Allowed, and What is Smart. Initially, vSAN 8 can only be run on vSAN ReadyNodes approved for use with the ESA, which will naturally invite consistency and symmetry across a cluster.

While vSAN ESA supports both Read-Intensive (RI) and Mixed-Use (MU) devices, mixing devices with different endurance levels within a host, or across hosts within a cluster is not supported. For more information, see the post: "Expanded Hardware Compatibility for vSAN Express Storage Architecture." The additional support of QLC will announced with VCF 9.1 will only be for certain use cases, and cannot be mixed with other storage device types.

Is there a migration path to vSAN ESA?

For customers who have existing clusters running previous versions of vSAN, and wish to upgrade to vSAN 8, they can perform an in-place upgrade of the cluster as usual**. In this scenario, the cluster will then be upgraded to vSAN 8 and use the original storage architecture (OSA).** For customers with ReadyNodes approved for use with the ESA, a cluster can be created, and during the installation process, The "vSAN ESA" can be selected and it will proceed to install and configure the cluster as such. When the configuration is complete, customers can migrate the VMs using vMotion and Storage vMotion. vSAN in VCF 9.1 also introduces more flexibility in remote datastore sharing, which can also help with this transition.

For more information on transiting to the ESA, see the post: Migrating to the Express Storage Architecture in vSAN 8.

Does vSAN ESA look and operate in the same way as past vSAN versions?

Yes. Many of the aspects of the Express Storage Architecture are "under the hood" architectural changes. vSAN continues to operate in the same way as past editions and remains **the software you already know.** In many ways, operational aspects become much easier. For example, vSAN ESA does not use disk groups, so attending to a maintenance issue with a storage device becomes easier with a smaller area of impact. Storage policies are fundamentally easier with vSAN ESA in VCF 9.1, as the new "Auto-RAID" feature takes responsibility of ensuring an optimal level of resilience for all data on the datastore. This also translates to improved ease of use in other aspects, such as the new "Effective Capacity" view in vSAN for VCF 9.1

Note that the new "Auto-RAID" should be considered as the replacement to Auto-Policy Management capability found exclusively in vSAN ESA. More information on Auto-RAID can be found in these FAQs and other vSAN documentation.

I see the minimum hardware requirements, and it indicates that faster networking is required on the ESA for the majority of ReadyNode profiles. Does vSAN ESA use more network resources to process data?

No. While many ReadyNode profiles will define 25Gb or 100Gb as the minimum requirement, this requirement reflects the ESA's ability to deliver near device-level performance of the high-performing NVMe-based storage devices approved for use. Ensuring sufficient network resources allows for vSAN to exploit the full performance capabilities of the devices under maximum load. **If you are migrating production workloads from a vSAN OSA cluster to a vSAN ESA cluster, on average you will see fewer CPU and network resources used for those same workloads**. This is because the vSAN ESA uses fewer CPU cycles and fewer network resources to process and store I/O when compared to the vSAN OSA. While vSAN ESA is very efficient in transmitting data across the network, it is an extremely fast storage stack, and under heavy load may hit the physical network limits of your cluster. To understand the impact of this, see the post: "What to Look for in Network Switches for VMware vSAN." **Therefore, we recommend 25Gb or higher networking when using vSAN ESA.**

I'm looking for my favorite ReadyNode on the compatibility list for ESA, but do not see it. What am I doing wrong?

This may occur for a few reasons.

- **Completion of certification.** Server manufacturers are responsible for the certification of systems, and the certification status will be dependent on the desires of the server manufacturer.
- **Generation of ReadyNode.** Among other hardware requirements, the Express Storage Architecture requires ReadyNodes that use chipsets no older than Intel's Ice Lake family of processors. ReadyNodes using processors older than Intel Ice Lake (Debuted in April, 2021) use an older generation of PCIe, and also do not have sufficient number of PCIe lanes to support the quantity of NVMe storage devices supported by the ESA. These older servers will also be incapable of providing line-rate speeds of modern, 100Gb NICs.

For the latest list of ReadyNodes compatible with ESA, visit the Broadcom Compatibility Guide (BCG) for vSAN.

Does the Express Storage Architecture in vSAN use a caching device?

No. vSAN ESA does not use a dedicated caching device. The Express Storage Architecture is a flexible architecture using a single tier, so all performance and capacity needs are being performed by the same storage devices.

What is a storage pool in the vSAN ESA?

The vSAN ESA removes the concept of disk groups and discrete caching and capacity tiers, and replaces it with a "storage pool." This storage pool is comprised of all storage devices selected on the host to provide storage resources to vSAN. Since vSAN ESA does not have a dedicated caching tier to deliver performance, all selected storage devices in a storage pool on a host will contribute to capacity and performance. This improves the TCO as all devices contribute to capacity, and it also dramatically simplifies the provisioning process, and reduces the impact of a storage device failure. There can only be one storage pool per host. And these contribute to a single vSAN datastore per cluster.

Was vSAN rewritten when it introduced ESA?

No. Many aspects of vSAN Original Storage Architecture (OSA) are used throughout the Express Storage Architecture (ESA). vSAN has solved many of the great challenges of distributed storage systems, and we wanted to build off of these capabilities already built into vSAN. The ESA simply helps customers capitalize on the capabilities of this latest generation (and beyond) of high-performance hardware. Its architecture allows for the most efficient use of resources, so that vSAN is best positioned to power all application types.

How does vSAN ESA deliver the performance of RAID-1 mirroring while using RAID-5/6 erasure coding?

vSAN introduces a new, patented log-structured file system (LFS) and an optimized log-structured object manager to deliver significant efficiencies throughout the stack. The vSAN LFS allows us to ingest writes quickly, efficiently and in a durable manner, while preparing the data and metadata for an efficient, full stripe write. The new LFS in the ESA takes advantage of our approach to writing data resiliently by first quickly writing using a redundant mirror, and packaging it in a way that allows vSAN to write the data to a stripe with parity, all while maintaining the metadata in a very efficient and fast manner.

For more information, see the post: RAID-5/6 with the Performance of RAID-1 using the vSAN Express Storage Architecture. Stretched clusters will also see performance and efficiency improvements as well. For more information, see the post: Using the vSAN ESA in a Stretched Cluster Topology.

Does ESA in vSAN for VCF 9.1 and later offer better performance than the ESA found in vSAN 8?

Yes. Recent editions introduced several enhancements that drive better performance than the initial release of vSAN ESA.

Which storage policy data placement scheme (RAID-1, RAID-5, or RAID-6) should I use for VMs powered by a cluster running the ESA?

Since vSAN ESA eliminates the trade-off of performance versus deterministic space efficiency, the data placement scheme recommended is largely dependent on the size and type of vSAN cluster, and the level of resilience desired by the customer. **vSAN in VCF 9.1 changes the approach to data resilience using "Auto-RAID" and will determine this optimal resilience setting for you**, regardless of your cluster type or size.

For vSAN **ESA clusters prior to 9.1,** the guidance below will help you determine which data placement scheme makes the most sense for your organization.

- **For clusters with 7 or more hosts**. Select FTT=2 using RAID-6. This spreads the object data (and parity) across 6 hosts. It offers a very high level of resilience while being able to store the data in a space-efficient manner. With 7 hosts, one spare fault domain (host) will be available to regain the prescribed level of resilience in a host failure or maintenance condition.
- **For clusters with 6 hosts.** Select FTT=1 using RAID-5. This spreads the object data (and parity) across 5 hosts. It offers the resilience of data while providing supreme levels of space efficiency. With 6 hosts, one spare fault domain (host) will be available to regain the prescribed level of resilience in a host failure or maintenance condition.
- **For clusters with 3-5 hosts.** Select FTT=1 using RAID-5. This spreads the object data (and parity) across 3 hosts. It offers the resilience of data while providing great levels of space efficiency. With 4-5 hosts, one spare fault domain (host) will be available to regain the prescribed level of resilience in a host failure or maintenance condition.
- **For 2-Node clusters**. Select FTT=1 using RAID-1. This mirrors the data across two hosts and uses the virtual witness host appliance to help determine quorum.
- **Stretched clusters.** Select FTT=1 using RAID-1. This mirrors the data across two sites and uses the virtual witness host appliance to help determine quorum.

Beginning with vSAN 8 U2, the ESA offered the Auto-Policy Management feature that helps determine and configure a cluster-specific storage policy for a given vSAN cluster.

vSAN in VCF 9.1 introduces Auto-RAID. This is a new capability, and the successor to Auto-Policy Management that **automatically manages the optimal resilience settings for data in a cluster.** For all vSAN clusters upgraded to VCF 9.1, it will replace discrete specialized storage policies created manually or by Auto-Policy Management with a single policy. Each cluster will be able to dynamically adjust its ideal resilience settings as the cluster characteristics change.

I read that the vSAN ESA can support RAID-5 on 3 hosts? How does this work? I thought vSAN required 4 hosts at a minimum for RAID-5?

The older vSAN OSA uses a 3+1 stripe with parity data placement scheme, and as a result, requires 4 hosts minimum to run in a non-error state. The vSAN ESA uses two new RAID-5 erasure codes, that are automatically determined by vSAN based on the size of the cluster, and will adapt to the size of the cluster automatically.

- **6 or more hosts.** When RAID-5 is applied to objects in a cluster of this size, the vSAN ESA will use a 4+1 data placement scheme. This results in the capacity for data to be stored resiliently will only consume 1.25x the size of the original object. Note that this is no longer used in vSAN for VCF 9.1 when Auto-RAID is enabled. A 2+1 RAID-5 erasure code will always be used when it cannot meet the minimum requirements for a 4+2 RAID-6 erasure code.
- **3-5 hosts.** When RAID-5 is applied to objects in a cluster of this size, the vSAN ESA will use a 2+1 data placement scheme. This results in the capacity for data to be stored resiliently will only consume 1.5x the size of the original

object. This is much better space efficiency for storing data resiliently than the 2x required by a RAID-1 mirror, and you get to enjoy this guaranteed level of space savings with just a 3-host cluster.*

** While the RAID-5 erasure code in vSAN ESA allows it to be used on a cluster as small as 3 hosts (similar to the minimum for a RAID-1 mirror),* ***VMware still recommends sizing a cluster host count to one additional host beyond the requirements of the storage policies used.***

For more information on the new RAID-5 in ESA, see the post: Adaptive RAID-5 Erasure Coding with the Express Storage Architecture in vSAN 8.

I see vSAN objects have more components now. Should I be concerned with that?

vSAN ESA uses a modified data structure found in the original storage architecture of vSAN. As a result, objects in a vSAN ESA-powered cluster will typically have more components than the same objects in a vSAN OSA-powered cluster. To help accommodate for this, the vSAN ESA increases the host component count limit from 9,000 components to 27,000 components. There is however a per-device component count limit that encourages the use of 6 or more storage devices per host. In vSAN 8 u2, the limit of VMs per host has increased from 200 VMs to 500 VMs per host when using the ESA. The limit of 200 VMs per host in the OSA remains the same.

What is a "capacity leg" and "performance leg" in the vSAN ESA, and what do I need to know about them?

The ESA uses a modified object format to store data in an object in two legs: The performance leg, and the capacity leg. These are contained all within the same object, and interact with the vSAN log-structured filesystem (LFS). This uses the same underlying distributed object manager to store the payload of data and metadata but does so in a manner that delivers performance and capacity all from the same tier. **It is simply an implementation detail and does not impact the design or operations of an environment.** It is a great example of how VMware integrated a new, optional architecture within an existing code base.

I'm looking at a RAID-1 object in my vSAN ESA cluster, and I don't see any witness components. Where did they go?

When using the vSAN ESA, Storage policies that use RAID-1 mirroring will not have a dedicated witness component for an object, except for in 2-node and stretched cluster topologies, where a virtual witness host appliance is used. Quorum for these RAID-1 objects is determined by the voting of components that comprise both the capacity leg and the performance leg of an object, thus a witness component is no longer necessary. Note that as of vSAN in VCF 9.1, the use of RAID-1 mirroring only occurs in rare circumstances, thanks to the capabilities of "Auto-RAID."

I see compression capabilities in vSAN ESA in VCF 9.1 has moved back to a cluster-based service, and is always on by default. Why is this, and can I turn it off?

vSAN ESA in VCF 9.1 has introduced a new tuned compression technique that is much more space efficient while having minimal impact on CPU resources. As a result, data compression will always be on for all workloads in a cluster. It cannot be turned off in the UI.

I upgraded my cluster to vSAN for VCF 9.1, and am waiting for better compression rates, but only see limited improvement. Why is this?

Upgraded clusters will leave the data as-is with the older compression algorithm used in previous versions of vSAN. It will only use the new compression algorithm on existing data that is overwritten, and new data. Therefore, additional savings may take some time. It also assumes the data is compressible.

I'm using compression with deduplication in my vSAN ESA cluster. But the compression savings do not appear to be as much as I expected given the reported improvements in 9.1. Why is this?

When vSAN Global Deduplication is enabled on a vSAN cluster, vSAN will render the savings based on whatever space efficiency technology was able to save the most capacity. For example, some highly compressible data would show its initial savings in the UI under compression. But upon the next deduplication process, if vSAN determines it will be able to save

more space by deduplicating the data, it will do so. As a result, deduplication rates will go up, and compression rates may go down. The overall data reduction ratio improves due to it using a combination of both techniques. The only way to do a direct comparison of compression improvements is to ensure that deduplication is not enabled. This is described in more detail in the vSAN Space Efficiency Technologies document.

What is the theoretical maximum compression that vSAN ESA can deliver?

vSAN evaluates and compresses data per 4KB data block. The vSAN ESA can theoretically compress a data block as little as 8:1, or 8x, with more granularity (e.g. 7:1, 6:1, 5:1, etc.). This is on a per data block basis. Data compression is an opportunistic space efficiency feature, so the benefits that real workloads will see will depend on the type of data stored.

For more information, see the "vSAN Space Efficiency Technologies" document.

vSAN ESA supports data encryption. Does this mean that it supports data-at-rest or data-in-transit encryption, or both?

The vSAN ESA encrypts data high in the storage stack, prior to writing to other hosts. This means that when vSAN encryption is used with a cluster running the ESA, it will be encrypted in-flight, and at rest. We still provide the option to enable Data-in-Transit encryption for clusters using the ESA to ensure that all encrypted packets remain unique. The ESA also eliminates previous decrypt, and re-encrypt processes, which reduces overhead. In VCF 9.1, we introduced the capability of using vSAN Data-in-Transit encryption for clusters mounting remote vSAN datastore, such as a vSAN storage cluster. This ensures that data is encrypted from end to end.

For more information, see the "vSAN Encryption Services" document.

How much faster will vSAN ESA be than vSAN OSA?

This is a constantly evolving answer, but a good comparison of performance of the vSAN ESA and OSA in vSAN 8 U1 is provided in the blog post: vSAN 8 U1 Express Storage Architecture - Faster than Ever. Many of our customers typically see a storage solution that is 3x to 5x faster, with performance results surpassing traditional storage arrays. For more information, see the post: "vSAN ESA Beats Performance of Top Storage Array for Large Financial Firm."

The answer will depend on hardware configurations and workloads, and if synthetic tests are used, versus real-world workloads. For the latter, the measure of success will be monitoring the level and consistency of the latency, as provided by the vSAN performance service. Synthetic testing can be a useful exercise for stress tests but provide limited functional benefit for translating how real workloads will behave. An illustration of how synthetic generators compare to real workloads can be found in the blog post: Performance when using vSAN Encryption Services.

Achieving optimal performance will also be easier to do with vSAN ESA. For example, one will not need to adjust the stripe width storage policy rule to achieve better performance. For more information, see the post: Stripe Width Storage Policy Rule in the vSAN ESA. For general recommendations to optimize performance with vSAN ESA, see the post: Performance Recommendations for vSAN ESA.

Why does vSAN ESA include the ability to automatically manage vSAN-related network traffic?

With the tremendous efficiency of the vSAN ESA, storage traffic can be processed through the stack at nearly device-level rates. This higher rate of processing I/Os in a server means higher rates of I/O traversing the network, which can potentially lead to the network being the bottleneck when running highly demanding workloads. To help accommodate for this, the ESA in vSAN 8 includes an adaptive traffic shaping capability for vSAN I/O traversing a network. This helps ensure that when network contention occurs, vSAN will properly prioritize VM I/O over resynchronization activity. This can help deliver more consistent performance for these demanding workloads that may otherwise be saturating a network link.

For more information, see the post: Adaptive Network Traffic Shaping in the vSAN Express Storage Architecture.

How do I see if my new cluster running the vSAN ESA is performing better than my other OSA based vSAN clusters?

Performance can be checked by using the metrics provided by the vSAN performance service. **Monitoring latency as seen by the guest VM is a good way to determine if the storage can meet the demands of the workloads.** VMs with higher latencies on other clusters will have the potential to perform better when running on a cluster powered by the vSAN ESA. A VM demonstrating low, consistent latency is the desired outcome for all workloads. Sometimes if you have applications performing batch processes, the time taken to complete the batch processing can be compared from one cluster to another. A shorter time to complete a batch process is a good indicator of improved performance.

How do I see if my new cluster running the vSAN ESA is more efficient than my other OSA based vSAN clusters?

The "efficiency" of a cluster can be measured two different ways. 1.) The amount of computational resources (CPU) across the cluster for a given set of workloads. 2.) The amount of storage capacity across the cluster to protect the data resiliently. Prior to migrating workloads running on an existing vSAN cluster running the OSA, look for the average CPU consumption across the cluster, then compare it to the CPU consumption on the new cluster. The same can be performed for capacity. Although, note that performing a capacity comparison will only be accurate if you maintain the same FTT level for all VMs.

Why did VMware introduce a new snapshotting capability within the vSAN ESA?

The introduction of vSAN ESA means that vSAN can now manage the data in new and interesting ways. Building a new, native snapshotting capability was a great way to exploit the capabilities of the ESA, and help support the needs and use-cases of our customers. Customers will be able to create point-in-time states of data with minimal degradation in the performance of a VM, no matter how many snapshots are taken. The new native snapshot capability is integrated directly in vSphere and fully supports our broad backup partner community with the continued support of VADP backup integration. For more information, see the post: Scalable, High-Performance Native Snapshots in the vSAN Express Storage Architecture.

vSAN 8 U3 introduces all-new capabilities with ESA snapshots through its vSAN Protection and Recovery capability (previously known as "vSAN Data Protection" in versions prior to 9.1. More information on vSAN Protection and Recovery can be found in this FAQ, as well as the post: "vSAN Data Protection in VMware Cloud Foundation – The Solution You Already Own." New enhancements in 9.1 can be found at: "VMware vSAN Protection and Recovery Enhancements for VCF 9.1."

I've always heard vSAN described as analogous to an object store, but I hear vSAN ESA uses a new file system. I'm confused. What is used in vSAN ESA?

vSAN (OSA and ESA) uses a data structure that most analogous to an object store, which is an ideal approach for a distributed storage solution like vSAN. The ESA in vSAN 8 introduces a new "log-structured file system" known as the "vSAN LFS." This does not refer in any way to a traditional file system such as NTFS, ext4, or a cluster file system like VMFS. It is a common industry reference to a method of how data and metadata are written and appended to a circular log buffer and persisted to the storage subsystems. For the vSAN ESA, its LFS helps vSAN ingest data quickly and efficiently and allows data and metadata to be prepared and stored with high levels of efficiency, scalability, and performance.

How much capacity overhead is consumed when using the vSAN ESA?

The amount of overhead used is about the same as the vSAN OSA. But may be effectively less as data can be stored in more efficient ways. For a more detailed answer, see the post: Capacity Overheads for the ESA in vSAN 8 as well as Improved Capacity Reporting in VMware Cloud Foundation 5.1 and vSAN 8 U2.

vSAN in VCF 9.1 introduces new technologies that eliminate the need to factor in overheads. **The result is that when using vSAN in VCF 9.1, you will see the actual "effective capacity" that can be used in a vSAN cluster**, without the complexity of understanding raw capacity and calculating overheads.

Are there any features or capabilities in vSAN that are not available when using vSAN ESA?

As of vSAN in VCF 9.1, all features' gaps between the two architectures have been eliminated, with most new capabilities only being available in vSAN ESA.

I want to make sure my new ESA cluster is running as fast as possible. What steps should I take to ensure that I'm getting optimal performance from my ESA cluster?

We have provided a collection of Performance Recommendations for vSAN ESA to help users ensure that they are getting the most out of their environment. Be sure to refer to this resource often, as it will continue to be updated based on the capabilities of the ESA.

Note that with vSAN, **performance of a VM is derived from the host hardware and the network** used to interconnect the hosts in the vSAN cluster, **not the cluster host count**. While increasing the host count of a cluster will increase the aggregate IOPS and bandwidth achieved by the cluster, in most cases it will not improve the discrete performance capabilities observed by the VM. VM performance will be a function of the host hardware and network connectivity. See the post: "What to Look for in Network Switches for VMware vSAN" for more information.

Availability

What happens if a host fails in a vSAN cluster?

vSAN will wait for 60 minutes by default and then rebuild the affected data on other hosts in the cluster. The 60-minute timer is in place to avoid unnecessary movement of large amounts of data due to a temporary issue. As an example, a reboot takes the host offline for approximately 10 minutes. It would be inefficient and resource-intensive to begin rebuilding several gigabytes or terabytes of data when the host is offline briefly. vSAN will also write all subsequent updated blocks of data to an additional host, in addition to the writes that are being committed to the other replica object. These are known as "durability components" and helps improve durability of data in both planned, and unplanned events. See the vSAN Availability Technologies document for a better understanding on how vSAN handles various conditions in a cluster.

How does vSAN handle a dividing or isolation of parts of a network, known as a network partition?

vSAN uses a quorum voting algorithm to help protect against "split-brain" scenarios and ensure data integrity. An object is available for reads and writes as long as greater than 50% of its components are accessible.

As an example, a VM has a virtual disk with a data component on Host1, a second mirrored data component on Host2, and a witness component on Host 3. Host1 is isolated from Host2 and Host3. Host2 and Host3 are still connected over the network. Since Host2 and Host3 have greater than 50% of the components (a data component and a witness component), the VM's virtual disk is accessible.

However, if all three hosts in our example above are isolated from each other, none of the hosts have access to greater than 50% of the components. vSAN makes the object inaccessible until the hosts can communicate over the network. This is a protection mechanism that helps ensure data integrity. See the vSAN Availability Technologies document for a better understanding on how vSAN handles various conditions in a cluster.

What happens if a storage device fails in a vSAN host?

vSAN can not only handle host failures with ease, but also storage device failures. When a device is degraded and error codes are sensed by vSAN, all the vSAN components on the affected drive are marked degraded and the rebuilding process starts immediately to restore redundancy. If the device fails without warning (no error codes received from the device), vSAN will wait for 60 minutes by default and then rebuild the affected data on other disks in the cluster. With vSAN ESA, the boundary of failure of a single storage device failure is limited to just that storage device. For more information, see the post: "The Impact of a Storage Device Failure in vSAN ESA versus OSA."

What happens if there is not enough free capacity to perform all the component rebuilds after one or more host failures?

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

When additional resources come back online, vSAN will continue the repair process to comply with storage policies. We recommend maintaining enough reserved capacity for rebuild operations and other activities such as policy changes, etc.

What happens if there are multiple failures (loss of hosts, etc.) that exceed the configured threshold of failures?

Some vSAN objects will become **inaccessible** if the number of failures in a cluster exceeds the failures to tolerate (FTT) setting in the storage policy assigned to the given object. If a VM object is using a storage policy that uses FTT=2, and three of the hosts that contain this object fail simultaneously, the object will be offline (but not lost) to preserve the integrity of the data. **The FTT level does not specify the total failures that the cluster can tolerate, but rather, specifies the failure that the assigned VM object can tolerate**. For example, a large cluster can have multiple failures, but if only one of the hosts for a given object is impacted, the data remains available. See the post: "Erasure Codes in VMware vSAN versus Storage Arrays" for more information.

When using the vSAN Express Storage Architecture (ESA), use the more resilient FTT=2 using RAID-6 and enjoy the benefits of space efficient storage and high levels of resilience without any performance cost.

What does vSAN do to protect against discrete device failure?

vSAN ESA implements several techniques to assist with the potential failure of storage devices.

- **Utilization of S.M.A.R.T. Data.** Unlike SAS and SATA interfaces, S.M.A.R.T data for NVMe provides very good telemetry data about the storage devices themselves.
- **Storage Device endurance tracking via vSAN Health.** Device endurance is now tracked in ESA, and health checks will be triggered in the event that the devices approach the number of drive writes per day warrantied by the manufacturer.
- **vSAN Proactive Hardware Management (PHM) framework.** This framework allows OEMs to use vLCM to surface relevant device-level telemetry data to vSphere, providing better visibility. **OEMs will then guarantee an RMA for devices that are predicted to fail.**
- **Low-level metadata resilience to protect against UREs. vSAN** ESA has unique levels of resilience to address "Unrecoverable Read Errors" or UREs. This helps correct data retrieval issues with minimal effort.
- **DDH to identify suspicious device patterns.** DDH has been adjusted to better account for behaviors of flash devices when running vSAN ESA. It can better predict failure, and reduces false positives.

How do I protect VMs residing on vSAN?

Many third-party data protection products use VMware vSphere Storage APIs for Data Protection (VADP) to provide efficient, reliable backup and recovery for virtualized environments. These APIs are compatible with vSAN just the same as other datastore types such as VMFS and NFS. Nearly all these solutions should work with vSAN.

vSAN Protection and Recovery in VCF 9.1 (previously named "vSAN Data Protection" in VCF 9.0) uses ESA's snapshotting mechanism to protect workloads locally, and remotely VMware Site Recovery or VMware Cyber Recovery (previously known as VMware Live Recovery, or VLR). See the "Protection and Recovery" section in this FAQs for more information.

Protection mechanisms can often be layered on top of each other to address specific use cases. For example. One could have traditional VADP-based backups to provide backups that follow the typical 3-2-1 rule of protecting data (three copies of data across two types of media, and one copy offsite) paired with vSAN Protection and Recovery that would augment this effort by providing an easier and faster way to recover workloads. Or if vSAN Protection and Recovery can meet your protection requirements on its own, your process and toolset for your environment becomes easier.

Does vSAN work with both products within the VMware Advanced Cyber Compliance (ACC) license – Site Recovery and Cyber Recovery?

Yes. In VCF 9.1, workloads can be protected to a remote location using VMware Site Recovery Manager (SRM) or VMware Cyber Recovery. This gives the ability to remotely protect data to a customer owned infrastructure, for the purposes of disaster recovery, ransomware recovery, or both.

Is there a way to stop vSAN data resynchronizations?

vSAN provides ways to gracefully stop resynchronizations, which means that in some cases, existing resynchronizations will finish to completion while others are halted prior to starting. vSAN also has protections in place that will pause resync operations if disk space usage meets or exceeds critical thresholds. Subsequently, the operations are resumed when sufficient capacity is made available.

How is vSAN impacted if vCenter Server is offline?

vCenter operates as the primary interface to manage and monitor vSAN. However, vCenter does not affect the data plane i.e., the VM I/O path. When vCenter Server is offline, vSAN continues to function normally. VMs continue to run, and application availability is not impacted. Management features such as changing a storage policy, monitoring performance, and adding a disk group are not available.

vSAN has a highly available control plane for health checks using the VMware Host Client—even if vCenter Server is offline. Hosts in a vSAN cluster cooperate in a distributed fashion to check the health of the entire cluster. Any host in the cluster can be used to view vSAN Health. This provides redundancy for the vSAN Health data to help ensure administrators always have this information available. One may wish to deploy a management cluster for additional durability of their data center. For more information, see the post: "Using vSAN as a Management Cluster."

Does vSAN support application-level clustering?

Yes. Infrastructure-level availability of applications and their data provided by vSAN can be combined with application-level clustering capabilities. Although this may add additional complexity if it is not serving a very specific need. For more information, see the post: "Application Versus Infrastructure-Level High Availability with vSAN in VMware Cloud Foundation."

Many of these clustered applications use a shared quorum disk. vSAN supports multiple approaches to this, through the use of a multiwriter flag for Oracle RAC, and through SCSI3-PR shared disks for Windows Server Failover Clusters (WSFC). vSAN in VCF 9.1 supports the ability to expand, or "hot-extend" these quorum disks. For more information, see the post: "Expand Shared VMDKs with Clustered Applications in VMware vSAN for VCF 9.1."

What happens if a vSAN cluster loses power?

The architecture of vSAN ensures that writes are always written to qualified persistent media in a redundant way - ensuring the availability and integrity of data. While vSAN strives for durability under the harshest of environments, a data center design that uses redundant power or standby temporary power is always encouraged.

Once power is restored to the entire cluster, the hosts will begin to power on. The initial booting of ESXi may take a little longer than usual to improve service states and data. This time difference will be relatively insignificant if the cluster is using vSAN ESA. When the vSAN hosts complete the boot process, HA will initiate the power-on of VMs previously turned on, and vSAN will begin any resynchronizations to ensure full data resilience and honor the storage policies prescribed to the data.

For environments that only have temporary power during times of sustained power loss, vSAN has shutdown workflows in the UI and programmatically to help provide for a speedy shutdown. See the post: "Automation Improvements using PowerCLI 13.1 with vSAN 8 U1" for more information.

Does vSAN store data in a crash consistent manner?

Yes. vSAN ensures that data has persisted to disk before write acknowledgements are ever sent back to the guest VM. It is also written in a resilient way defined by the resilience setting in the prescribe storage policy. vSAN uses special protocols to achieve consistency of data written to multiple locations.

Upon an unexpected failure, such as a power loss of a host where some of the VM data resides, the VM will continue to write data without issue, but in a less resilient way until the data is automatically reconstructed elsewhere. If a power loss occurred on a host where the VM instance was running, then vSphere HA will restart that VM elsewhere, and work with the data persisted to disk.

When a VM is protected (backed up) using our APIs for data protection (3rd party solutions using VADP, or vSAN Protection and Recovery built into vSAN), the data and its data structure is persisted to disk in a state that is preserved at the time of the crash. This may or may not include preserving the memory state of the VM. This is what is referred to as "crash consistency." When referring to three-tier architectures using traditional storage arrays, the additional step of a "stun" may need to occur on all VMs residing in a LUN. This would pause the I/O so that the storage array could create a LUN-based snapshot in a consistent manner. For crash-consistent backups using vSAN Protection and Recovery, **a VM stun is unnecessary for vSAN** as we are fully aware of the I/Os and if they are persisted to disk. This makes snapshots in vSAN Protection and Recovery fast, efficient, and scalable. For more information, see the post: "Superior Snapshots using VMware vSAN Data Protection." Note that a stun may occur if change block tracking (CBT) is enabled.

Crash consistency is what most backup vendors provide, because it is purely working with the data that has been fully committed to disk. It is sufficient in most cases. "Application consistency" is a more sophisticated step that involves quiescing the data and application state performs a coordinated flush of items in memory to disk to ensure that changes aren't currently being held in memory. Some Operating Systems have mechanisms in place that allow for application or file-based consistency (such as Microsoft's Volume Shadow Copy Service, or VSS), while other operating systems must use pre and post backup scripts to quiesce an application. Backup applications often need to coordinate with these mechanisms to provide application consistent backups.

The ability for vSAN to provide crash consistency applies to both the VM object data, as well as any associated snapshots using VADP, or vSAN Protection and Recovery as a part of VCF 9.1.

I understand that vSAN provides infrastructure-level availability, but how is this different than application-level availability, and which one should be used?

See the post: "Application Versus Infrastructure-Level High Availability with vSAN in VMware Cloud Foundation" for more details.

My storage array vendor states their array can tolerate more failures than vSAN. Is this true?

Not in the way that you would envision. Storage arrays have proven themselves to be a good option for enterprise storage but do suffer from some technical challenges due to their design. For example, some can tolerate three storage device failures in a storage array, but because there are so many storage devices, the probability of failure is worse than other solutions protecting against two storage devices failures. vSAN's distributed storage architecture and use of erasure coding offers tremendous levels of availability. For more information, see the post: "Erasure Codes in VMware vSAN versus Storage Arrays."

What is "Auto-RAID" found in the latest edition of vSAN for VCF?

Auto-RAID is a new capability introduced in vSAN for VCF 9.1. It automatically manages optimal resilience settings for data stored in a vSAN cluster. This optimal resilience setting is applied and enforced across the cluster and will dynamically adjust as the characteristics of the cluster change. It is enabled by default on all new vSAN clusters and can be enabled on clusters after they are upgraded to 9.1. Auto-RAID should be considered as the successor to "Auto-Policy Management." The latter served as a helpful step in simplifying storage policy management but was limited in its potential. Only Auto-RAID should be used in environments running VCF 9.1 and newer. For more information, see the post: "Auto-RAID in VMware vSAN for VCF 9.1 - Comprehensive System-Managed Data Resilience."

Does Auto-RAID mean that granular settings with storage policies are going away?

No. Storage policy settings such as IOPS limits and object space reservations (OSR) still exist, and can be applied on an as-needed basis. Auto-RAID simply takes control of the resilience settings so that you can be certain that data has the highest potential level of resilience based on the characteristics of the cluster. The behavior of Auto-RAID also helps deliver

consistent overheads for data stored on clusters, which offers customers the ability to view capacity usage by actual effective capacity, instead of raw capacity. This is achieved through the "effective capacity" view in 9.1 for clusters with Auto-RAID enabled.

For Auto-RAID enabled clusters, **other storage policy rules that are no longer relevant to vSAN ESA will not show up within the policy**, as they are no longer applicable. This includes:

- Force provisioning (now automatically handled)
- Number of disk stripes per object (irrelevant)
- Flash read cache reservation (irrelevant)
- Disable checksum (irrelevant)
- Compression (now an always-on cluster service in VCF 9.1)

Cloud-Native Storage

What is Cloud-Native Storage, or CNS?

Cloud-Native Storage (CNS) is a term used to describe the storage that can be provisioned to Cloud-Native Applications (CNAs). These CNAs are typically containerized, deployed and managed by a Container Orchestrator like Kubernetes, Mesos, Docker Swarm, etc. The storage consumed by such apps could be ephemeral or persistent, but in most cases, it is required to be persistent. CNS is supported when using the vSAN OSA, and when using the ESA in vSAN 8 U1 (block based RWO volumes only) and both block and file when using the ESA in vSAN 8 U2.

What is a Container Storage Interface, or CSI?

Container Storage Interface (CSI) is a standardized API developed for container orchestration platforms to interface with storage plugins. This API framework enables vSAN to be able to provision persistent volumes to Kubernetes based containers running on vSphere.

Can a vSAN datastore be used to provision persistent storage for a Kubernetes cluster?

Yes, vSAN supports provisioning persistent volumes to Kubernetes based workloads. A brief walkthrough is available here - Cloud-Native Storage on vSAN.

For vSAN in VCF 9.1, much of the provisioning of persistent volumes is achieved through the various services (VKS service, VM service, etc.) offered by VCF Automation.

Does vSAN provide any type of multitenancy?

Yes. When using VCF 9.1, one can provide tenant services such as VM services or VMware Kubernetes Services (VKS) to multiple tenants. This can be achieved in a VCF instance where the Supervisor has been enabled. The supervisor will be able to provide volume services using VCF Automation. It will allow for the creation of regions, organizations, as well as projects and namespaces to provide all the logical boundaries necessary for multitenancy. This same approach applies to the new native S3 compatible object storage using vSAN available as a technology preview in 9.1.

S3 Compatible Object Storage

Does vSAN offer S3 compatible object storage capabilities?

Yes. vSAN in VCF 9.1 includes an all new S3 compatible object storage capability built directly into vSAN. This will give VCF administrators the capability to deploy highly scalable and resilient S3 storage all courtesy of vSAN. It is initially available as a Technology Preview and will be generally available in a later version. It will initially be ideal for containerized workloads, but use cases will certainly expand. For more information, see the post: "Native S3 Compatible Object Storage in VMware vSAN for VCF 9.1."

Does vSAN native S3 compatible object storage require any additional licensing?

Licensing information will be provided when it becomes generally available in a later edition of vSAN for VCF.

Can a vSAN cluster storing VM data also be used to provide S3 compatible object storage?

Yes. A customer can provide block, file, and S3 compatible object storage all on the same cluster.

Does Native S3 compatible object storage support all S3 APIs?

No. The initial Technical Preview aimed to support a large number of S3 API calls primarily focused on Dev Ops functionality (CRUD, ACL, etc.), and will work to expand the support of S3 APIs incrementally in future releases.

How can native S3 compatible object storage be configured and provisioned?

Configuration and provisioning of object storage primarily occur through VCF Automation. A Provider Admin will install the object store services, and assign it to the desired tenant organizations, while the Tenant Admin will provision S3 bucks for consumption by a Tenant User. S3 compatible object storage can also be consumed in environments not using VCF Automation. The management and consumption would occur through the vSphere Supervisor. For more information, see the post: "Native S3 Compatible Object Storage in VMware vSAN for VCF 9.1."

vSAN File Services

How is vSAN File Services integrated into vSAN?

vSAN File Services is powered and managed by the vSphere platform that deploys a set of containers on each of the hosts. These containers act as the primary delivery vehicle to provision file services and are tightly integrated with the hypervisor. **In vSAN for VCF 9.0, file services in vSAN ESA support up to 500 shares per cluster.** File Services in VCF 9.1 introduces several enhancements to improve the effective performance of file services.

Can I run VMs on top of a file share provided by vSAN File Services?

No, it is not supported to mount NFS to ESXi for the purpose of running virtual machines. The NFS shares may be used to mount NFS directly to virtual machines running on the vSAN cluster, but may not be used to store VMDKs for running virtual machines.

vSAN has evolved to help with storing vSphere content libraries and ISOs. See page 60 of the vSAN Operations Guide for more information.

What is the minimum number of hosts required in a cluster to deploy vSAN File Services?

For a standard vSAN cluster, a minimum of 3 hosts is required to configure vSAN File Services. It will run with as many as 2 remaining hosts. vSAN File Services will auto-scale 1 container per host up to 64 per cluster. vSAN File Services will work on a 2-Node cluster however, which only has two data nodes.

Is vSAN File Services supported on a stretched cluster and 2-Node cluster?

Yes, both topologies are supported.

What is the estimated resource overhead of each host when running vSAN File Services?

One protocol services VM instance runs on each host in the cluster. Each VM instance is configured with 4GB of RAM and 4vCPU. By default, there are no reservations applied to the resource pool associated with the entities required for vSAN File Service.

How is vSAN File Services monitored?

vSAN File Service can be monitored with vSAN Skyline Health Services. A health check called "File Service - Infrastructure health" monitors several parameters and includes an automated remediation option.

What protocols and authentication methods are supported?

NFSv3, NFSv4.1, SMB v2.1 and SMBv3 are supported. Both NFS and SMB file shares are now able to use Kerberos based authentication when using Microsoft Active Directory.

Can a single share provide access using NFS and SMB at the same time?

Simultaneous access of a single share using both NFS and SMB is not supported at this time.

How can snapshot functionality be used in vSAN File Services?

The snapshot mechanism introduced in vSAN 7 U2 allows for our backup partners and Independent Software Vendors (ISVs) to provide point-in-time backup and recovery capabilities of files in vSAN file service shares into their backup solutions. The full capabilities of the snapshot mechanism are accessible via API. Using file share snapshots as a backup source requires backup products supporting the functionality. Backup vendors are currently working to provide support for vSAN file shares. Until then, organizations can use PowerCLI and backup vendor PowerShell modules to add newly created snapshots as a backup source.

Note that this snapshot mechanism is completely independent and different from the snapshotting mechanism within vSAN ESA.

Do I need to migrate or manage the file services VMs?

No. vSAN automatically manages the File Server VM. The containers are automatically shut down and removed. It will be recreated once an available host is no longer in maintenance mode.

Do I need to create or add vmdks or objects to expand storage to vSAN File Services?

vSAN File service uses elastic scalability and will create additional components as needed without any manual intervention.

How can I limit the consumption of shares provided by vSAN File Services?

Soft and hard share quotas can help manage capacity consumption.

- **Hard quotas** prevent users from writing data to disk. Hard quotas automatically limit the user's disk space, and no users are granted exceptions. Once users are about to reach their quota, they must request help.
- **Soft quotas** send alerts when users are about to exceed disk space. Unlike hard quotas, there is no physical restriction to prevent users from saving their data. However, you do get alerts and can create a corporate policy to help manage data.

Can I provision file shares to Cloud-Native workloads?

Yes, vSAN File Services can be used to provision file shares to container workloads as well as traditional workloads.

How do NFS shares recover from host failure or migrate during upgrades?

vMotion and vSphere HA are not used as part of migration, or failure recovery. Services within vSphere monitor for failure or maintenance activities and drive the relocation of services. The containers powering vSAN file services, automatically restart on other hosts, independent of vSphere HA.

While by default you will have one container per host, additional containers will run on a host in a case where a host (or hosts) have failed. When a host enters maintenance mode the container powering a given share or group of shares is recovered on a different host.

How is vSAN File Services updated?

An updated OVF can be automatically downloaded or manually updated to the vCenter managing the cluster. A non-disruptive rolling upgrade will proceed across the cluster replacing the old containers with the new version.

Disaggregated Storage using vSAN Storage Clusters

What is disaggregated storage in vSAN?

Disaggregated storage in vSAN uses a unique software-based approach to disaggregate, or decouple compute and storage resources. **vSAN HCI aggregates storage and compute resources into a single cluster, while disaggregation decouples storage and compute resources.** First introduced in vSAN 7 U1, it has since been enhanced up to and including vSAN in

VCF 9.1. In the latest version of vSAN, datastore sharing using traditional vSAN HCI clusters allows the datastore of a vSAN HCI cluster (a "server cluster") to be mounted by other vSAN HCI clusters (considered a "client cluster") for the purposes of consuming storage resources. Disaggregation with vSAN storage clusters (previously known as vSAN Max) is a deployment option where a dedicated cluster provides storage resources to one or more vSphere clusters. This provides full independence of scaling between your compute resources and your storage resource, but still using all of the technology of vSAN ESA with consistent unified management across all configuration types.

vSAN storage clusters can be used as principal storage. It supports all ReadyNodes certified for vSAN storage clusters, as well as all cluster sizes recommended. At this time, it can be used for workload domains. The use of vSAN storage clusters as principal storage in a management domain is not supported.

With vSAN in VCF 9.1, vSAN and vSphere clusters can mount remote vSAN datastores regardless of their underlying architecture (e.g. ESA or OSA). This means that OSA clusters can mount both ESA and OSA clusters concurrently. ESA clusters can also mount both ESA and OSA clusters concurrently. And finally, compute clusters (vSphere clusters) can mount both ESA and OSA clusters concurrently.

What capabilities does disaggregation in vSAN add to data center environments?

Full disaggregation of compute and storage resources using vSAN storage clusters **provide dedicated, centralized storage** for vSphere clusters, or even vSAN HCI clusters. A vSphere cluster participating in this topology would be referred to as a "compute cluster" and would mount a remote vSAN datastore to consume the storage resources of that remote vSAN datastore. It uses native vSAN protocols for supreme levels of efficiency and functionality.

Another deployment option is "vSAN HCI with datastore sharing." It allows a vSAN HCI cluster to consume resources from other vSAN HCI clusters. This helps use stranded capacity between clusters and use unique data services or hardware capabilities provided by a given cluster (such as Data-at-Rest Encryption). It also allows administrators and architects the ability to scale compute and storage independently, easing design and operational complexities. For more information on which deployment option may be best for your environment, see: "vSAN HCI or vSAN Max - Which Deployment Option is Right for You?"

What are vSAN storage clusters? (previously named “vSAN Max”)

vSAN storage clusters are a deployment option in vSAN that provides highly flexible disaggregated storage for vSphere clusters. It is powered by the vSAN Express Storage Architecture, or ESA. It provides our customers the ability to deploy a highly scalable storage cluster to be **used as primary storage for vSphere clusters**, or augment storage for traditional vSAN HCI clusters. This makes vSAN the premier storage platform for powering VMware Cloud Foundation. See the post "Introducing vSAN Max" for more information. vSAN in VCF 9.1 introduces more flexibility for storage clusters. For more information, see the post: “Greater Flexibility and Security with VMware vSAN Storage Clusters in VCF 9.1.”

What would be some common use cases for vSAN storage clusters?

Some common uses cases would include, but are certainly not limited to the following:

- **Cost optimization for infrastructures and applications.** It is not unusual for customers to want to configure clusters to help minimize application costs. vSAN storage clusters will allow customers to right-size their compute resources to minimize these licensing costs. One can tailor the vSAN storage cluster to whatever the business needs in terms of performance, capacity, and data services provided.
- **Operational simplicity with unified storage.** vSAN storage clusters can be used to extend the life of blade servers and other older hardware that was not ideal for HCI. If you've wanted to take advantage of vSAN but wanted to retain compute and storage resources independent from each other, vSAN storage clusters is for you. **It can easily serve as your primary, centralized shared storage platform for your data center**, and be scaled easily and incrementally as your needs grow. It is a true, distributed scale-out storage solution that can be used as centralized shared storage by any of your vSphere clusters. assets.

- **Maintain alignment with your organization.** Many organizations wish to maintain a dedicated storage team who is responsible for the provisioning and maintenance of storage. vSAN storage clusters allow customers to maintain this aspect of their organizational model just as if they were using a traditional storage array.

How are vSAN storage clusters licensed?

Licensing for storage in VMware Cloud Foundation is based on two elements, 1.) The number of CPU cores licensed, and 2.) the amount of additional capacity (per TiB) purchased. **For every core licensed in VMware Cloud Foundation, a complimentary 1 TiB of vSAN capacity is provided**. vSAN can be deployed as an aggregated "vSAN HCI" or a disaggregated "vSAN storage cluster" using these vSAN entitlements that are associated with your licensed VCF cores Any additional capacity beyond the sum capacity provided by the core count would be through a capacity add-on license.

Which deployment option you choose is entirely up to you. While one may be preferable over the other based on your environment, the new licensing model for VMware Cloud Foundation largely removes licensing as a factor for considering one deployment option for another. When a vSAN HCI cluster is provisioned, all cores and capacity licensing would reside on that cluster. But if a disaggregated approach was chosen, some licensing cores will reside on the vSphere clusters, while other licensing cores and capacity reside on the vSAN storage cluster. When using a disaggregated deployment option, **fewer cores are needed for each respective vSphere host, since vSphere hosts are only responsible for running VM instances, and vSAN storage cluster hosts are only responsible for processing storage.**

Do hosts in a vSAN storage cluster need to be licensed with VCF? If so, why is this?

Yes, all hosts that comprise a vSAN storage cluster will need to be licensed with VCF. Think of VCF as an operating system (OS) that needs to run wherever some aspect of the OS is needed - in this case, storage. When considering the typical hardware costs of storage-dense servers, the proportion of cost of the VCF licenses for these hosts in a storage cluster are quite small - often single digit percentages.

The benefit of the VCF licensing model is that it **allows you to apply the vSAN storage entitlement of 1 TiB per VCF core for any of your vSphere clusters running VCF to the storage cluster**. This is illustrated in a very simple example on Figure 2 in the post: "Starting Small with vSAN Storage Clusters."

Isn't a vSAN storage cluster just a vSAN HCI cluster without any running VM instances?

Even though vSAN storage clusters share many characteristics with a vSAN HCI cluster, they are not the same. When a vSAN storage cluster deployment option is chosen, special tunings are automatically made that change the storage stack to accommodate for the fact that the vSAN storage cluster is going to be predominantly processing I/O and storing data, not running a substantial amount of VM instances. The tuning helps maintain more metadata in memory, which helps both read and write operations. ReadyNodes certified for vSAN storage clusters are also adjusted with lower requirements to reflect that they are not running VM instances on the cluster.

vSAN storage clusters in VCF 9.0 and later also provide a new "network traffic separation" capability. This allows one VMkernel port to be tagged for backend vSAN traffic while another VMkernel port is tagged for frontend traffic going to and from the vSphere hosts mounting the datastore. This offers substantial performance improvements, as well as more efficient network isolation.

Are ReadyNodes certified for vSAN storage clusters the same as ReadyNodes certified for vSAN HCI?

No. vSAN storage clusters will use vSAN ReadyNode profiles that share similarities to ReadyNodes for vSAN HCI clusters, but with much higher capacities. For the latest comparison, see: "vSAN ESA ReadyNode Hardware Guidance." For more information, see the post: "ReadyNode Profiles Certified for vSAN Max." Additional flexibility in supported server configurations and host counts in vSAN storage clusters were also introduced recently. See the post: "Driving Down Storage Costs with Lower Hardware Requirements for vSAN."

Are there general design, sizing, and other recommendations for vSAN storage clusters?

Yes, the "vSAN Storage Clusters Design and Operational Guidance" document covers planning and sizing requirements and considerations, Day-0 deployment and configuration guidance, and Day-2 Operational guidance to help guide you through the process of providing a high performing, robust storage environment using vSAN storage clusters.

What versions of vSphere can be used to connect to a vSAN storage cluster datastore?

The hosts in a vSphere cluster attempting to mount a vSAN storage cluster datastore must be running **vSphere 8 or later**. vSAN clusters connecting to a vSAN storage cluster datastore must be running vSAN 8 or later, and MUST be using vSAN ESA. With new capabilities such as support for Data-in-Transit encryption between compute clusters and the storage cluster, these servers must be running VCF 9.1 or newer.

vSAN storage clusters are built using vSAN ESA. When the server cluster (either a vSAN storage cluster deployment option, or a traditional vSAN HCI deployment option) is powered by vSAN ESA, initially only client clusters that are running vSAN ESA can mount the datastore of a server cluster. In VCF 9.1, support for connecting OSA and ESA clusters to a vSAN storage cluster now makes connectivity much easier. For more information, see the post: "Greater Flexibility and Security with VMware vSAN Storage Clusters in VCF 9.1."

How much additional CPU and memory is required for hosts in a vSphere cluster to communicate with a vSAN storage cluster?

The vSAN client service activated on a vSphere cluster is purpose-built for the role of communicating with a server cluster. The resources required on the hosts in the client cluster are relatively small, but will vary based on a number of factors, including the number of server clusters mounted, features used and workload activity. One may see up to about 17GB of memory used per host in the client cluster connecting to a vSAN storage cluster. Typically, the CPU demands on the hosts in the client cluster will be less than one core.

What cluster types and connectivity are supported with vSAN storage clusters?

vSAN storage clusters can be configured either as a single site cluster, or stretched across two sites to provide site-level resilience. Client cluster types can include vSphere clusters (also known as vSAN compute clusters), vSAN HCI clusters, and vSAN HCI clusters that are stretched or in 2-node topologies.

Client Cluster Type	Server Cluster Type	Supported	Notes
vSAN HCI clusters (ESA) in a stretched cluster configuration.	vSAN storage cluster or vSAN HCI cluster (ESA) in a stretched cluster configuration	Yes	Provides resilience of data and high availability of running VM instances.
vSAN HCI clusters (ESA) when it resides in one of the data sites where the vSAN storage cluster resides.	vSAN storage cluster or vSAN HCI cluster (ESA) in a stretched cluster configuration	Yes	Provides resilience of data but no high availability of running VM instances.
vSphere clusters stretched across two sites using asymmetrical* network connectivity.	vSAN storage cluster or vSAN HCI cluster (ESA) in a stretched cluster configuration	Yes (in VCF 9.0)	Not supported in versions prior to vSAN 9.0.
vSphere clusters stretched across two sites using symmetrical* network connectivity.	vSAN storage cluster or vSAN HCI cluster (ESA) in a stretched cluster configuration	Yes	Supported, but less common, as it would require the same network capabilities (bandwidth and

			latency) between fault domains defining each site.
vSphere clusters when it resides in one of the data sites that the vSAN storage cluster resides.	vSAN storage cluster or vSAN HCI cluster (ESA) in a stretched cluster configuration	Yes	Provides resilience of data but no high availability of running VM instances.
Any type of client cluster running **vSAN OSA**	vSAN storage cluster or vSAN HCI cluster (ESA) in a single site or stretched cluster configuration	Yes (in VCF 9.1)	**Not supported in versions prior to vSAN 9.1**

* Asymmetrical network connectivity would describe a topology where the network capabilities (latency & bandwidth) connecting the two sites (fault domains) would be less than the network capabilities between the client cluster and the server cluster within each site. This is most common with stretched cluster configurations using and inter-site link (ISL) between sites. Symmetrical network connectivity would describe a topology where the network capabilities connecting the two sites would be the same as the network capabilities between the client cluster and server cluster within each site. This configuration is less common, but might be found in environments where the two fault domains defining the sites in the stretched topology are simply server racks or rooms sitting adjacent to each other using the same network spine.

Aren't traditional storage arrays "disaggregated?" And if so, how is this any different?

Yes, a traditional three-tier architecture using a storage array provides storage resources that are "disaggregated" from compute resources. But that approach has inherent limitations. The design of a monolithic storage array connected to a dedicated storage fabric using a pair of redundant controllers must funnel all I/O through those controllers to a clustered filesystem like VMFS, then rely on locking mechanisms to prevent simultaneous access. vSAN storage clusters are quite different. It is a fully distributed storage system that, because of its design, offers incremental scalability of both capacity and performance, a common management plane, all running on commodity servers. This helps overcome most of the challenges with a traditional three-tier architecture using a storage array.

See the post: "vSAN Max and the Advantage of Scalability" for more information.

Can disaggregation be used to maintain cluster homogeneity of server vendors?

Absolutely! Disaggregation will allow you to keep servers from the same manufacturer in the same cluster, while consuming that storage from other vSphere or vSAN clusters. This can be ideal for organizations using multiple server vendors.

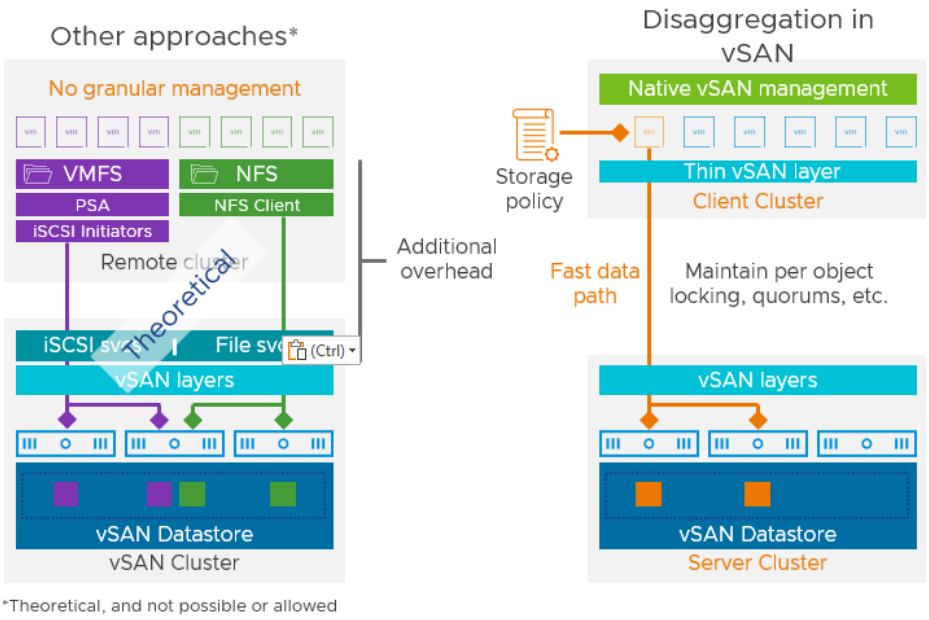
How is disaggregation with vSAN different than composable/modular infrastructures?

vSAN's disaggregation capabilities use a software-based approach to disaggregation that can be implemented on any certified hardware. Composable infrastructure is based off a hardware-centric approach that requires specialized hardware in unique or proprietary form factors. Unlike other solutions, our approach treats the disaggregation at the cluster level. This helps avoid the challenges associated with stop-gap approaches such as storage-only nodes found with other solutions. This is why **most customers find using rack mounted 2U servers the most flexible and affordable for any type of vSAN cluster**.

Recently, VMware partnered with Samsung to develop a proof of concept using vSAN Max that connected to dedicated NVMe-based JBOF enclosures to extend the concept of disaggregation. This is still a proof of concept at this time, and not commercially available.

Which protocol and data path does disaggregation in vSAN use?

It uses vSAN's native protocol and data path for cross-cluster connections, which preserves the vSAN management experience and provides the most efficient and optimized network transport.



Using an approach as illustrated on the left side of the image above would have a complex data path that would inhibit performance, efficiency, scalability, and compatibility objectives. vSAN's stack was developed for the specific needs of a distributed system and has been adapted to a disaggregated architecture.

Can a vSAN storage cluster be mounted to vSphere clusters?

Yes, this is the primary use case, where vSAN storage clusters provide centralized shared storage resources for vSphere clusters. When configuring a vSphere cluster for connection to a vSAN storage cluster, a thin layer of vSAN is activated on the hosts participating in the client cluster and is used to communicate natively with the server cluster that is providing the storage capacity and services. For communication, disaggregation in vSAN uses native vSAN protocols for supreme levels of efficiency and functionality.

Do hosts in client clusters and the vSAN storage cluster need to be using the same CPU manufacturer?

No. The CPU manufacturer (e.g. Intel, AMD) used in hosts in a client cluster can be different than the CPU manufacturer of the hosts that comprise a vSAN storage cluster. For example, a vSphere client cluster using AMD CPUs can mount the datastore of a vSAN storage cluster comprised of hosts using Intel CPUs. The vSAN storage cluster must use hardware certified for use with vSAN storage clusters, while a client vSphere cluster only needs to adhere to the basic hardware compatibility list for vSphere.

Do hosts in client clusters (vSphere clusters) need to be certified vSAN ReadyNodes?

No. Hosts in client clusters that are mounting an external vSAN **datastore only need to comply with the Broadcom Compatibility Guide (BCG) for vSphere, not the BCG for vSAN.**

We do recommend that network connectivity between a client cluster and server cluster be as fast as possible, since this represents the fabric that storage traffic is transmitted. See the document "vSAN Storage Cluster Design and Operational Guidance" for further information on connectivity recommendations from client clusters to server clusters powered by vSAN storage clusters.

What are some of the scaling capabilities with vSAN storage clusters?

Just like vSAN HCI clusters, vSAN storage clusters can be scaled incrementally to add more capacity and performance. Scale up by adding more or higher density storage devices in the hosts or scale out by adding more storage nodes. Cluster capacities are based on what is available in ReadyNodes certified for vSAN storage clusters. A client cluster can mount up to a maximum of 5 remote vSAN datastores, and a server cluster can export its datastore up to a maximum of 10 client

clusters. Up to 128 hosts can connect to a remote vSAN datastore - when counting the hosts from the client cluster(s) and the server cluster.

VMware has guidance on various design maximums. See the document "vSAN Storage Cluster Design and Operational Guidance" for more information.

Can VMs be provisioned to span across multiple remote datastores?

All objects related to VM (VMDKs, VM Home, etc.) **must reside in a single datastore** - which can be either local or remote. For example, a VM cannot have one of its objects residing in a local datastore, while another one of its objects resides on a remote datastore.

Does disaggregation integrate with other vSAN features?

Yes. As of vSAN in VCF 9.1, most features found in vSAN HCI deployments are available in vSAN storage clusters. This includes but is not limited to data services such as vSAN File Services, data compression, deduplication, data-at-rest encryption, data-in-transit encryption, etc.

What are the network recommendations to implement disaggregation with vSAN?

Recommendations follow into two categories: 1.) intra-cluster "east/west" traffic and 2.) client cluster "north/south" traffic. For the former, low-latency and high bandwidth network topologies are recommended for optimal performance, which means 25Gb or 100Gb network with no network oversubscription. For the latter, client clusters mounting the datastore generally require less network bandwidth, and can get away with using 10GbE connectivity. 25GbE or faster is preferred. See the document "vSAN Storage Cluster Design and Operational Guidance" for more information specific to connectivity to a vSAN storage cluster.

More generalized guidance for vSAN HCI to vSAN HCI cluster disaggregation includes the following:

- Design for redundancy everywhere for highest availability (multiple links, NICs, TOR/spine switches, etc.)
- For all but the very smallest of vSAN storage cluster ReadyNode profiles, use NICs and storage-class switches capable of at least 25Gbps for connectivity between the hosts that make up the vSAN storage cluster. The **vSphere clusters do not need to meet this minimum**, but faster networking will help ensure that it is not the performance bottleneck. For more information, see the post: "What to Look for in Network Switches for VMware vSAN."
- Keep the storage cluster in a single rack, if you are not using the vSAN Fault domains feature. This helps optimize the network traffic for your environment. For more information, see the post: "vSAN Networking – Optimal Placement of Hosts in Racks."
- vSphere clusters ("compute clusters") acting as the client cluster can use 10GbE connectivity to the server cluster. This is to accommodate older server types such as blades where customers want to extend the life of the servers in use. However, we still highly recommend 25GbE or 100GbE for client cluster to server cluster communications.
- Use NIOC with vSphere Distributed Switches
- **Use only the networking teaming policy of "Active/Standby" with "Route based on originating virtual port ID."** Other teaming options such as active/active using LBT are not appropriate for vSAN clusters. For more information, see the posts: "vSAN Networking – Teaming for Redundancy" and "vSAN Networking – Teaming for Performance."
- The mounting of a vSAN storage cluster or vSAN HCI datastore will run a precheck and look for network connectivity between the client clusters and the server cluster of 5ms or less. Since this connection represents storage traffic, it is best to strive for 1ms or less for network connectivity between client clusters and a server cluster.

Are there any availability considerations with disaggregation in vSAN?

vSAN's disaggregation offering uses existing vSphere HA and vSAN availability concepts to provide both compute and storage high availability. vSphere HA will provide compute availability on the client cluster, and vSAN storage policies with FTT=N configured will provide storage availability on the server cluster. If the network connection between a client and server cluster is severed, the remote vSAN datastore on a client host will enter APD 60 seconds after the host becomes

isolated from the server cluster. After that, vSphere will follow the current HA mechanisms for APD events and attempt to restart VMs after a 180 second delay.

Are storage policies integrated with disaggregation?

Yes, all VMs consuming storage on vSAN on a vSAN HCI cluster, or a vSAN storage cluster, are controlled by storage policies. With vSAN in VCF 9.1, ensuring optimal levels of resilience on the VMs stored on the remote datastore becomes quite easy thanks to Auto-RAID, as it will automatically set and manage the optimal resilience for all of the data stored on the datastore.

Is cross-cluster vMotion (without storage vMotion) supported with disaggregation in vSAN?

Yes. VMs can be migrated between two clusters sharing a single datastore and are fully supported with disaggregated topologies.

Does the configuration of a vSAN storage cluster require the use of DRS and HA? And what about Virtual Distributed Switches?

Even though a vSAN storage cluster does not host any user-created guest VMs, some vSphere configuration settings are necessary for proper functionality.

- Virtual Distributed Switches (vDS) should be used to simplify management.
- DRS and HA should be enabled.
- vMotion interface should be configured to ensure mobility of management VMs.

For more information on configuration of a cluster, see the "Preparing the vSAN storage cluster for its initial configuration" section in the vSAN Storage Cluster Design and Operational Guide.

What happened to HCI Mesh?

HCI Mesh was the name for the early disaggregation capabilities of vSAN. It represented connectivity from one vSAN HCI cluster to another vSAN HCI cluster for the purpose of **cross-cluster capacity sharing**. As noted above**, this type of configuration still exists**, and is now referred to as "**vSAN HCI with Datastore Sharing**." vSAN storage clusters (previously known as “vSAN Max”) is a new first-class citizen in vCenter Server, giving the unique ability to provide a "storage-only" cluster to serve resources to vSphere clusters, and even vSAN HCI clusters. These two offerings were not well represented by the name of HCI Mesh. The term is no longer used in recent releases of vSAN.

What is the difference between vSAN storage clusters and the old HCI Mesh?

HCI Mesh represented connectivity from one vSAN HCI cluster to another vSAN HCI cluster for the purpose of cross-cluster datastore sharing. This type of capability still exists between vSAN HCI clusters but is not referred to in the UI as HCI Mesh anymore. It is now referred to as "vSAN HCI with Datastore Sharing." This option will also allow compute clusters to mount the datastore of a vSAN cluster, but for a centralized shared storage solution, vSAN storage clusters are built using vSAN ESA, and will be the best way to achieve supreme levels of scalability and flexibility.

vSAN storage clusters are our fully disaggregated deployment option where a vSAN storage cluster provides the storage services only to vSphere clusters, and even vSAN HCI clusters. vSAN storage clusters are integrated in the UI in a way that gives customers the ability to easily configure and manage the vSAN storage cluster and provide easy connectivity and management to vSphere clusters. vSAN storage clusters are also tuned specifically for serving storage I/O. It can even take advantage of additional network ports using “network traffic separation.”

Can vSAN HCI with Datastore sharing be used with 2-Node clusters?

Yes. As noted above, a 2-Node cluster can be used as a client cluster that mounts the datastore of another vSAN HCI cluster, or a vSAN storage cluster. A 2-Node, vSAN HCI cluster may also be used as a server cluster. While it is possible to use this type of cluster to serve resources to other clusters, this type of cluster will have a reduced ability to tolerate multiple host failures when compared to a vSAN HCI cluster with a larger host count.

With the introduction of vSAN storage clusters, is an aggregated vSAN HCI approach no longer preferable?

The introduction of vSAN storage clusters is about choice and flexibility to tailor systems to the specific needs of your organization. Since vSAN HCI clusters treat storage as an exclusive resource of the cluster, they can be ideal for specialized needs (R&D, Finance, etc.) or even multi-tenant environments where physical separation of resources is required. We believe **aggregated vSAN HCI clusters and disaggregated vSAN storage clusters can provide a powerful combination** for your enterprise needs.

I want to make vSAN storage clusters as fast as possible. How do I do this?

Use higher performance hardware to achieve faster storage performance. With vSAN, performance of a VM is derived from the host hardware and the network used to interconnect the hosts in the vSAN cluster, not the cluster host count. While increasing the host count of a cluster will increase the aggregate IOPS and bandwidth achieved by the cluster, in most cases it will not improve the discrete performance capabilities observed by the VM. VM performance will be a function of the host hardware and network connectivity. Paying attention to your network topology will be critical in delivering the highest, most consistent performance possible. **Ensure that your spine-leaf topology is not oversubscribed.** For more information, see the post: "vSAN Networking – Network Oversubscription." Using network traffic separation with vSAN storage clusters will also improve performance.

I've heard conflicting information on features like vSAN encryption services and the new global deduplication capabilities, and if they are supported in vSAN storage clusters. What is the stance of availability and support of these data services?

As of vSAN ESA in VCF 9.1

- vSAN Data-at-Rest Encryption: Supported in vSAN HCI clusters and vSAN storage clusters
- vSAN Data-in-Transit Encryption: Supported in vSAN HCI clusters and vSAN storage clusters. Also supports the use of DiT encryption from the client clusters to the vSAN storage cluster.
- vSAN Global Deduplication: Supported in vSAN HCI clusters and vSAN storage clusters
- vSAN Data-at-Rest Encryption with Global Dedup: Supported in vSAN HCI clusters and vSAN storage clusters

As of vSAN ESA in **VCF 9.0 P01**

- vSAN Data-at-Rest Encryption: Supported in vSAN HCI clusters and vSAN storage clusters
- vSAN Data-in-Transit Encryption: Supported in vSAN HCI clusters. Not supported in vSAN storage clusters
- vSAN Global Deduplication: Supported in vSAN HCI clusters and vSAN storage clusters (upon TQR approval)
- vSAN Data-at Rest Encryption with Global Dedup: Not supported in vSAN HCI or vSAN storage clusters.

As of vSAN ESA in **VCF 9.0**

- vSAN Data-at-Rest Encryption: Supported in vSAN HCI clusters and vSAN storage clusters
- vSAN Data-in-Transit Encryption: Supported in vSAN HCI clusters. Not supported in vSAN storage clusters
- vSAN Global Deduplication: Not available in vSAN 9.0
- vSAN Data-at-Rest Encryption with Global Dedup: Not available in vSAN 9.0

Stretched Clusters and 2-Node Clusters

What is a 2-Node or 2-Host vSAN cluster and how does it work?

A vSAN 2-node cluster is a special type of vSAN cluster that consists of 2 hosts storing the data, and one host in the form of a virtual witness host appliance that provides quorum voting capabilities to determine availability and prevent split-brain scenarios.

Why do vSAN stretched clusters and 2-Node clusters need a third location for a witness?

In a stretched cluster topology, the third location that holds a witness host appliance will determine quorum for the data in the stretched cluster. This prevents failure scenarios that could result in a split-brain configuration: The same VM running and updating data in two locations simultaneously. For 2-node clusters, the same principles apply, where the third entity is simply a witness host appliance at a remote location (typically the primary datacenter) and determines quorum for the two nodes in the cluster. For more information, see the post: "Application Versus Infrastructure-Level High Availability with vSAN in VMware Cloud Foundation."

What are the hardware requirements for running a vSAN stretched cluster or a vSAN 2-Node cluster?

vSAN 2-node configurations have two physical hosts, with a third witness host in a third location. A stretched cluster can have up to 40 physical hosts (20 at each site) with one witness host in a third location.

Can a witness host be shared across multiple deployments?

2-Node clusters can share a witness host, supporting up to 64 2-Node clusters per witness host. See the post "New Design and Operation Considerations for vSAN 2-Node Topologies" for a better understanding on the considerations of sharing a witness host across multiple 2-Node clusters.

vSAN stretched clusters cannot share a witness host appliance. This would create a failure domain that would be unnecessarily large only for the purpose of saving a minimal amount of CPU and memory.

Can the witness host appliance be deployed in the Cloud?

The **witness host appliance is packaged as an OVA**. This virtual witness host appliance must run on a physical, licensed ESXi host. More than one witness host appliance can run on a physical ESXi host. Any cloud provider issuing a licensed ESXi host for use can be used for hosting the virtual witness host appliance.

Does the ESXi host version powering the virtual witness host appliance need to be the same version as the appliance?

The vSAN Witness Host is contributing to the vSAN cluster, therefore it is recommended to be the same build as the vSAN Data Nodes. It is generally required to be the same release as vSAN.

A vSAN Witness Appliance is provided with each release of vSAN, but not always updated with every patch of ESXi, so the exact build numbers may not match. Upon initial deployment of the vSAN Witness Appliance, it is required to be the same as the version of vSAN.

What are my options for redundancy in a stretched cluster configuration?

In a stretched cluster configuration, data can be mirrored across sites for redundancy. Additionally, a secondary level of resilience can be assigned to data that resides within each site. Assuming enough hosts on each site to do so (as few as 3 hosts on each data site), the secondary level of resilience will provide additional resilience in the event of host outages in the data sites. Assigning site-level protection and a secondary level of protection is all achieved in a single storage policy.

Using vSAN ESA in a stretched cluster environment has some compelling advantages. For more information, see the post: Using the vSAN ESA in a Stretched Cluster Topology.

Can stretched clusters maintain data availability if there is a failure of one site and the witness host appliance?

If they are simultaneous failures, no, as this would not meet the levels sufficient for quorum. In vSAN 7 U3 and later, data in one site can fail or be taken down for maintenance, followed by a subsequent outage of the witness host appliance, and the data will remain available. This enhancement also applies to 2-node clusters as well. vSAN in VCF 9.1 does introduce a "manual site takeover" capability that improves the ability to recover data from a data site if one data site is in maintenance mode, and the witness site fails. This allows you to regain availability of data and VMs for site in maintenance mode following permanent failure of the other data site.

Do secondary levels of resilience in a stretched cluster help maintain availability if there is an outage of a data site and a witness site?

No. Secondary levels of resilience are intended to improve availability of data within each site if one of the data sites is unavailable. It does not help maintain availability in the event of a double failure.

Can 2-Node clusters provide a secondary level of resilience?

Yes, secondary levels of resilience exist for 2-Node clusters. This can help maintain availability of data in the event of a host failure, and discrete storage device failures on the remaining host. Due to the hardware requirements, it is much more efficient and economical to use the secondary level of resilience on 2-Node clusters running ESA.

Can I use "vCenter HA" with vSAN stretched clusters?

While vCenter HA (VCHA) can be used with vSAN Stretched Cluster using DRS affinity/anti-affinity to pin VMs to sites a cost/benefit analysis must be done in order to understand if it will work in your environment and operational model.

VCHA requires a maximum of 5ms latency between all three VMs, the primary, secondary, and witness - as such, placing the VCHA witness on the same site as the vSAN stretched cluster witness host would mean that site can be a maximum of 5ms away, not 200ms as is supported by a vSAN stretched cluster. If your vSAN witness sites are more than 5ms away an option is to co-locate the VCHA witness with either the primary or the secondary VCHA VMs, however, this also means that if the co-located site fails, the vCenter Server will be offline.

A **vSAN stretched cluster is natively integrated with vSphere HA**, offering automated failover and startup for all VMs on it - including vCenter. The benefit of using VCHA over a single vCenter VM with vSphere HA to automate the restart is in the order of a minute or two in startup time - as such, it should be considered that if the extra startup time is acceptable, use vSphere HA and a single vCenter VM for operational simplicity.

If not, use VCHA with your chosen topology: either the VCHA witness on the same site as the vSAN Witness provided it is less than 5ms away or co-located with the primary or secondary VCHA VMs taking into consideration the failure scenarios that go with such a co-located topology.

Is vSAN File Services supported on a stretched cluster and 2-Node clusters?

Yes. File services are supported in the vSAN ESA as of vSAN 8 U2.

Does disaggregation work with vSAN stretched cluster and 2-Node clusters?

Yes, disaggregation is compatible with vSAN stretched clusters and 2-node clusters when using vSAN 8 U1 (OSA), and vSAN 8 U2 in the ESA in a vSAN HCI with Datastore sharing deployment. As of VCF 9.0, vSAN storage clusters support a stretched cluster topology where the vSphere clusters are also stretched across the same two sites that the vSAN storage cluster is stretched. See the post: "Flexible Topologies with vSAN Max." for more information.

Are there recommendations for vSAN stretched cluster network connectivity?

The vSAN Stretched Cluster guide and the vSAN Stretched Cluster Bandwidth Sizing guide contains more information and recommendations specific to stretched clusters networking. While the Inter-site link (ISL) can be a significant contributor to the effective performance of a vSAN stretched cluster, it isn't the only factor. See the post "Performance with vSAN Stretched Clusters" and "Using the vSAN ESA in a Stretched Cluster Topology" for more information.

Can a standard single site vSAN cluster be converted to a vSAN stretched cluster?

Yes, it is easy to convert a standard (non-stretched) vSAN cluster to a stretched cluster. This is performed in the "Fault Domains & Stretched Cluster" section of the vSAN UI. More details can be found in the vSAN Stretched Cluster Guide.

Miscellaneous

Where can I find technical blog posts related to vSAN?

As of January 2025, all vSAN-related technical blog posts can be found at: https://blogs.vmware.com/cloud-foundation/technical/storage/

Where can I find technical white papers and design guides related to vSAN?

As of January 2025, all vSAN-related technical white papers, design guides, operational guides, etc. can be found at: https://www.vmware.com/resources/resource-center

Filter by "VMware vSAN" under the "Product" section.

Where can I find step-by-step "how to" product documentation on vSAN? (equivalent to the old docs.vmware.com platform)

As of January 2025, all product documentation can be found at: https://techdocs.broadcom.com/us/en/vmware-cis/vsan/vsan/8-0.html

Where can I find product documentation on VMware Cloud Foundation?

As of January 2025, all product documentation for VCF can be found at: https://techdocs.broadcom.com/us/en/vmware-cis/vcf.html

I see some requirements or statements in support and compatibility that are different in the VCF documentation than they are in the vSAN documentation. Why is this, and which one is correct?

VCF provides a more prescriptive deployment than a stand-alone instance of one of its sub-components like vSAN. As a result, VCF may have additional restrictions. If an environment using the official VCF product stack through the SDDC Manager, then the documentation for VCF should be followed. For other environments where perhaps vSAN is running independently, then the vSAN-specific product documentation applies.

Some new features are noted as having "Limited Availability." What does this mean, and how do I qualify for a feature with this status?

Features are occasionally introduced with limited availability, which can occur for various reasons. In the past, VMware required approval for use of the feature through a "Request for Product Qualification" (RPQ). Under Broadcom, a similar process is available and called a "Technical Qualification Request" (TQR). For any feature that is labeled as "Limited Availability" a TQR must be submitted, and your request must be approved to use the given feature.

Where can I find the latest release notes for vSAN?

The release notes for vSAN in VCF 9.0.2.0 can be found on our TechDocs site. Release notes for other versions can be found on that same link.

Help me! I'm confused with some of the name changes for vSAN related functionality. Can you help?

Yes, we understand your confusion as we change toward a more functional naming style. Here are a few quick references to remember

- **vSAN Protection and Recovery.** This name replaces "vSAN Data Protection". This is the capability built into your VCF license that provides the ability for local protection and recovery of your VMs.
- **Site Recovery (or Site Recovery Manager.** This name replaces "VMware Live Site Recovery (VLSR) which was a part of the "VMware Live Recovery" (VLR) add-on license. SRM can acquired on its own as an add-on license to VCF, or come as a part of the more comprehensive "Advanced Cyber Recovery" (ACC) add-on license to VCF, which provides both DR and Cyber Recovery functionality.
- **Cyber Recovery.** This replaces "VMware Live Cyber Recovery (VLCR) which was a part of the "VMware Live Recovery" (VLR) add-on license to VCF. This is part of the VMware ACC add-on license to VCF, which provides both DR and Cyber Recovery functionality.
- **vSAN storage clusters.** This name replaces "vSAN Max." vSAN storage clusters are simply a deployment option in vSAN, and a part of your VCF licensing.

Networking

What are the networking requirements for running vSAN?

The networking requirements for vSAN will depend on the vSAN ReadyNode profile class used, as stated on the vSAN ESA ReadyNode Hardware Guidance document. Hosts participating in a vSAN cluster must be connected to the network using at least one network interface card (NIC). Multiple NICs with one or more ports are recommended for redundancy. For more information for teaming for redundancy, see the post: "vSAN Networking – Teaming for Redundancy."

The technical minimum bandwidth needed is 10Gbps. Given the current performance capabilities of vSAN ESA, 10Gb is the accepted minimum for **only the entry-level ReadyNode**. A realistic minimum is 25Gbps. For more information, see "vSAN ESA ReadyNode Hardware Guidance."

Are there recommendations for vSAN network connectivity?

The vSAN Network Design Guide contains more information and recommendations for network connectivity in a vSAN cluster. The "Design and Operational Guidance for vSAN Storage Clusters" document will also have network guidance that relates to both vSAN HCI clusters, as well as vSAN storage clusters.

A recent series of blog posts that summarizes considerations and recommendations for vSAN networking can be found below:

- vSAN Networking - Network Topologies
- vSAN Networking - Network Oversubscription
- vSAN Networking - Optimal Placement of Hosts in Racks
- vSAN Networking – Teaming for Performance
- vSAN Networking – Teaming for Redundancy
- vSAN Networking – Is RDMA Right for You?
- What to Look for in Network Switches for VMware vSAN

Does vSAN support RDMA?

vSAN supports the RoCE v2 implementation of RDMA, using qualified hardware and configurations. Other implementations of RDMA such as iWarp and InfiniBand are not supported. For more information on whether vSAN over RDMA is right for you, see the post: "vSAN Networking – Is RDMA Right for You?"

To ensure a consistent, predictable, and supported experience, strict adherence to approved RDMA adapter cards on the HCL is required. vSAN clusters using RDMA may be subject to additional limitations of supported features or functionality, including, but not limited to:

- vSAN cluster sizes are limited to 32 hosts
- vSAN cluster must not be running the vSAN iSCSI services
- vSAN cluster must not be running in a stretched cluster configuration
- vSAN cluster must be using RDMA over Layer 2. RDMA over Layer 3 is not supported.
- vSAN cluster running vSAN over RDMA is **not supported** with disaggregated vSAN topologies, such as **vSAN storage clusters**.
- vSAN cluster **must not** be using a teaming policy based on IP Hash or any active/active connection where sessions are balanced across two or more uplinks.

Can multiple VMkernel ports tagged for vSAN be used?

The use of multiple VMkernel ports tagged for vSAN is not an officially supported configuration at this time. In vSAN 9.0 and later, vSAN storage can use a "network traffic separation." Capability. This allows one VMkernel port to be tagged for backend vSAN traffic while another VMkernel port is tagged for frontend traffic going to and from the vSphere hosts mounting the datastore. This offers substantial performance improvements, as well as more efficient network isolation.

Does NIC teaming improve performance in vSAN?

NIC teaming aims to achieve two potential benefits. Improved failover/resilience and improved performance. NIC teaming using LACP can offer marginal performance improvements, but is complex to deploy, maintain and troubleshoot. It is also not supported at this time in VCF. NIC teaming using **active/active with Load Based Teaming (LBT) which is currently a default setting for VCF, was not intended for the deterministic requirements of storage traffic** and is **not recommended**. The most reliable and robust configuration at this time is using an active/standby arrangement using "route based on originating virtual port ID." This will provide a more deterministic path for storage I/O. **It is recommended that all vSAN deployments (whether they are stand-alone vSAN clusters, or as a part of VCF) use active/standby with "route based on originating virtual port ID" as method of connectivity.**

For more information, see the post: "vSAN Networking – Teaming for Performance."

Do faster network switches and interface cards improve vSAN performance?

Yes. In general, a faster networking fabric will improve performance to a vSAN cluster when the existing network fabric is a significant contributor to contention. This is especially true for clusters running the vSAN Express Storage Architecture (ESA) introduced in vSAN 8 and enhanced in vSAN 8 U1 and U2. As storage devices become faster, this can shift the primary point of contention to the network. 25/100Gb networking is quickly becoming commonplace as a result.

Does vSAN require storage fabric host bus adapters (HBAs)?

No, vSAN uses standard network interface cards (NICs) found in nearly every x86 server platform. There is no need to provision and implement specialized storage networking hardware to use vSAN.

Can I run vSAN traffic through a network overlay, firewall, IDS, or NSX?

While front-end VM traffic can run through network overlays, **we highly recommend that all VMkernel traffic (including vSAN traffic) has as simple of a path as possible.** Firewalls and IDS/IPS systems can inadvertently block this mission critical storage I/O in a manner that could cause substantial impacts on the performance or availability of data.

Can vSAN support direct (switchless) connection of hosts with clusters greater than two hosts?

No. This type of switchless connection is only supported in a 2-node topology. Attempting to do this with more than two hosts would create networking loops and violate general Layer 1/2 networking principles that apply to all environments, not just vSAN. Without a switch, there would be no spanning tree protocol (STP) to prevent network loops, and a simple node failure could easily break quorum.

What should I look for in a switch to be used with vSAN?

vSAN is a distributed storage system, so the network, and the attributes of the network switches become very important. Look for port speeds of 25Gbps or 100Gbps, but also be sure to look for the advertised "switch capacity." This represents the total amount of data a switch can handle at any given time, and is usually expressed in Gbps or Tbps. The higher this value is, the better as it represents more processing throughput capabilities of the switch. Next is to look at the "port buffers." This is the amount of memory on the switch that is used for incoming packets. It is usually expressed in Gbps. Typically, you will want to look for the "ultradeep" classification of buffers, which is about 4GB or higher. Larger port buffers do not inherently make the processing of packets faster, but they can reduce the amount of packets discarded under times of heavy contention. For more information, see the post: "What to Look for in Network Switches for VMware vSAN."

Capacity

How much capacity will I need in my vSAN cluster?

The amount of capacity you need depends on the amount of capacity you plan to use for the workloads that will reside on the vSAN datastore. See the "*How Much Effective Capacity to Expect from a vSAN Cluster*" section of the "vSAN Space Efficiency Technologies" document for more information. Your vSAN ReadyNodes can then be tailored to meet your capacity requirements. See the "vSAN ESA ReadyNode Hardware Guidance" for more information.

New to vSAN in VCF 9.1 is the "effective capacity" view. This will now render vSAN capacity in terms of the cluster's actual usable capacity. It eliminates the complexities of estimating overheads and raw capacity calculations, making vSAN storage management behave very similar to a traditional storage array. **This new view requires the use of vSAN Auto-RAID on a cluster** to ensure the highest level of accuracy. For more information, see the post: "Simplifying Storage with the New Effective Capacity View in VMware vSAN for VCF 9.1."

For all versions prior to VCF 9.1, vSAN presents storage capacity statistics for the cluster in raw form. As noted in the post: "Demystifying Capacity Reporting in vSAN" the cluster capacity advertised as available does not reflect the capacity available for data residing in a resilient manner. The amount of additional capacity needed depends on the type of storage policy applied, and whether you are using vSAN ESA or OSA.

How much free capacity should I maintain in a vSAN cluster?

The new "effective capacity" view paired with Auto-RAID in vSAN for VCF 9.1 eliminates the considerations around maintaining free capacity, as this overhead is now built in. What is displayed as usable capacity in the capacity view is indeed usable capacity.

Prior to vSAN in VCF 9.1, the amount of free raw capacity required varied based on a myriad of conditions but generally ranged about 20-25% (up to 30% for much older versions).

vSAN requires additional space for operations such as host maintenance mode data evacuation, component rebuilds, rebalancing operations, and VM snapshots. Activities such as rebuilds and rebalancing can temporarily consume additional raw capacity. Host maintenance mode temporarily reduces the total amount of raw capacity a cluster has. This is because the local drives on a host that is in maintenance mode do not contribute to vSAN datastore capacity until the host exits maintenance mode.

What is the new "Effective Capacity" view in vSAN for VCF 9.1, and how is it different than past capacity views?

For all clusters with Auto-RAID enabled, vSAN in VCF 9.1 provides a new "Effective Capacity" view that shows a cluster's actual, usable capacity. This means that vSAN moves away from showing capacity in raw form and expecting you to factor in all of the capacity overheads, and moves toward a usable capacity view that is analogous to traditional storage arrays. This new elegant view is comprised of effective capacity used and remaining, as well as a space-efficiency breakdown, so you can see how much capacity you are reclaiming by deduplication, compression, and thin provisioning. As a result, you will also see that VM capacity statistics for VMs stored on vSAN now show VMDK consumption statistics, not raw capacity statistics. For more information, see the post: "Simplifying Storage with the New Effective Capacity View in VMware vSAN for VCF 9.1."

I see references for Auto-Policy Management in VCF 9.1. Should I use Auto-RAID or Auto-Policy Management?

Always use "Auto-RAID" for clusters running VCF 9.1 or later. There is no need to enable the "Auto-Policy Management" feature for new clusters in 9.1. Auto-RAID is the much more capable successor to "Auto-Policy Management" and allows for simple, managed resilience across all your vSAN clusters, and is used by default on all new vSAN clusters in VCF 9.1.

References to "Auto-Policy Management" remain in 9.1 to help with compatibility. Clusters upgraded to 9.1 will continue to use their previous self-managed policies, which may include the custom policy created by Auto-Policy Management. This can be addressed in one of two ways:

- **Option 1: Change all objects to the "vSAN ESA Auto-RAID Policy."** This is the preferred method, as it is cleaner, and old storage policies no longer used may be removed from vCenter Server. The default datastore policy will need to be changed to "vSAN ESA Auto-RAID Policy" as well.
- **Option 2: Enable "Apply Auto-RAID to all objects."** This is a catch-all mechanism to ensure that all objects are using Auto-RAID regardless of previous, self-managed storage policies. While this method is quick and easy, old storage policy names may continue to be tied to some of your clusters. It also means that toggling off the setting

may reconfigure objects back to their old storage policy setting. The approach described in "Option 1" makes the toggle meaningless.

For more details, see the post: "Auto-RAID for vSAN in VCF 9.1 – Comprehensive System-Managed Data Resilience."

Should I enable "Host Rebuild Reserve" and "Operation Reserve" toggles in all of my vSAN clusters?

No. The Host Rebuild Reserve (HRR) and Operations Reserve (OR) toggles are optional capacity management toggles as a part of the "Reserved Capacity" capability applying versions of vSAN prior to 9.1. While it can simplify capacity management, in some cases it may not either be supported, or advised. The "Reserved Capacity" capability currently does not support clusters in stretched cluster configurations, 2-Node deployments, or clusters using vSAN's "Fault Domains" feature. For more information, see the post: "Understanding Reserved Capacity Concepts in vSAN."

Note that the "Host Rebuild Reserve" and "Operations Reserve" are constructs of capacity management prior to vSAN in VCF 9.1. With Auto-RAID and the "Effective Capacity" view in vSAN for VCF 9.1, those terms and toggles are no longer relevant.

How can I add storage capacity to a vSAN cluster?

Storage capacity can be increased in a few ways. Hosts containing local storage devices can be added to a vSAN cluster (scale out), or additional storage devices can be added to existing hosts (scale up).

For vSAN clusters running the Express Storage Architecture (ESA), storage devices can be easily claimed and added to the storage pool that resides on each host (one and only one storage pool on each host). Storage devices added to the storage pool will all contribute to capacity resources of the single vSAN datastore presented by the cluster.

This scale-out approach of adding more hosts is most common and adds compute capacity to the cluster. More storage devices can be added to existing hosts assuming there is room in the server's chassis to add these devices. Unlike traditional storage solutions, vSAN enables a "just-in-time" provisioning model. Storage and compute capacity can be quickly provisioned as needed.

vSAN supports the TRIM/UNMAP space reclamation options. How can this be monitored?

vCenter UI can be used to monitor IOPS and throughput generated through TRIM/UNMAP commands. This is available under **Monitor > vSAN > Performance > Backend**. Alternatively, this can also be monitored through vsantop. Looking at UNMAP latency on its own can be misleading, as it may show a relatively high latency with no UNMAP operations.

Space Efficiency

Is deduplication available in vSAN ESA?

Yes, as of vSAN in VCF 9.1, global deduplication is generally available to all customers. For more information, see the post: "More Capacity with VMware vSAN Compression and Global Deduplication in VCF 9.1."

How is deduplication in ESA different than OSA, and why is it so much more effective?

Several characteristics of vSAN ESA's deduplication make it different, and much more effective than what was found in vSAN OSA. It allows you to save costs and scale efficiently in ways not previously possible. Global deduplication in vSAN ESA differs from deduplication in OSA in the following ways:

- **Larger deduplication domain.** vSAN ESA's global deduplication will deduplicate data if finds across the entire cluster. vSAN OSA had a deduplication domain that was limited to a single disk group within a host. And unlike traditional storage arrays, the deduplication domain scales as the cluster host count grows.
- **Post processed.** Unlike vSAN OSA, deduplication in ESA is performed asynchronously. vSAN will perform these tasks when the cluster workloads are less busy, and will dynamically assign resources based on what is available. There will be minimal performance impact to the workloads.
- **Scalable.** As the cluster grows in size, so does vSAN's ability to increase its effectiveness.

- **No change in failure domain.** vSAN ESA will treat a discrete storage device failure in the same way as a cluster not using ESA's global deduplication. A single device failure will only impact the data on that device, which offers much better data durability than vSAN OSA.

For more information, see the post: "Global Deduplication in vSAN ESA for VMware Cloud Foundation 9.0."

How much capacity savings should I expect to see with deduplication in vSAN ESA?

Deduplication is an opportunistic space efficiency feature, so its effectiveness will depend highly on the type of data being stored, the likeness of the data, and the amount of data being stored. Internal testing has demonstrated data reduction ratios that meet or exceed popular storage arrays on the market. For more information, see the post: "Save Costs and Scale Efficiently with vSAN Deduplication in VMware Cloud Foundation 9.0."

What are some of the limitations with global deduplication in vSAN ESA?

As of vSAN in VCF 9.1, global deduplication is not supported in stretched clusters, and 2-Node clusters.

The capacity utilization page of my vSAN cluster gives a deduplication and/or compression savings that has changed. How should I interpret the savings?

Notation of savings or space efficiency capabilities (deduplication, compression, or both) compressibility can be easily conflated with the different methods of stating its absolute efficiency, or its efficiency relative to the previous implementation of compression. For example:

- **Expressed as a ratio.** [starting unit size]:[ending unit size]. A ratio of 10:1 would indicate that data that would otherwise consume 100TB would only consume 10TB after the space efficiency techniques are applied. **This is the way it will be rendered in vSAN for VCF 9.1.**
- **Expressed as a multiplier of reduction.** A multiplier of 10x would indicate that data that would otherwise consume 100TB would only consume 10TB after the space efficiency techniques are applied. This is the way it was rendered in vSAN for VCF 9.0 and earlier.
- **Expressed as a percentage of original size:** A percentage of 10% would indicate that data that would otherwise consume 100TB would only consume 10TB after the space efficiency techniques are applied.
- **Expressed as a percentage of savings:** A percentage of 90% would indicate that data that would otherwise consume 100TB would only consume 10TB after the space efficiency techniques are applied.

I enabled global deduplication on my vSAN ESA cluster, and my advertised compression savings ratio now shows as lower, even though deduplication has reclaimed a lot of capacity. Why does this occur?

This is simply due to how the savings from each space efficiency technique is calculated, and is expected. Calculating discrete ratios of deduplication and compression can be complex due to the wide variety of conditions for each data block. What matters most is the overall "Data reduction ratio" shown in the new capacity UI in vSAN for VCF 9.1. The overall data reduction is the most meaningful metric to help determine how much storage capacity is saved. This is addressed in more detail in the "vSAN Space Efficiency Technologies" document.

How is vSAN ESA more space efficient than vSAN OSA?

vSAN ESA achieves better space efficiency in several ways:

- ESA can use space efficient RAID-5/6 erasure coding without any compromise in performance.
- ESA uses a much more space-efficient method of data compression, enhanced in VCF 9.1.
- More effective space reclamation using TRIM/UNMAP
- ESA uses an all-new cluster-wide deduplication technique (introduced in vSAN 9.0) that is dramatically better than vSAN OSA.

Does vSAN support TRIM/UNMAP space reclamation techniques?

Guest Operating systems use commands known as TRIM/UNMAP for the respective ATA and SCSI protocols, to reclaim space that is no longer in use. This helps the guest operating systems be more efficient with storage space usage. Space reclamation will not only free up capacity, but it will help reconcile capacity reporting differences. See the post: "The Importance of Space Reclamation for Data Usage Reporting in vSAN." TRIM/UNMAP can be enabled from the CLI, or, beginning in vSAN 8, enabled in vSphere Client. For more information, see the "TRIM/UNMAP Space Reclamation" section in the vSAN Space Efficiency Technologies document.

Can space efficiency services such as deduplication and/or compression be enabled on an existing vSAN cluster?

In vSAN for VCF 9.1, compression is a cluster-based feature that is enabled by default and always on. Global deduplication, once enabled, can only be paused.

Will deduplication impact storage performance?

With vSAN ESA, **there is no material impact on performance when using compression, and vSAN's global deduplication is a post-processing activity and is not expected to produce any meaningful impact on guest VM performance.**

My storage array vendor states they use a more space efficient erasure code than vSAN. Is this true?

It may be, but that does not tell the complete story. Storage arrays use erasure codes of all types to deliver their specified amount of data resilience. But typically, the nature of their architecture drives their need to use erasure codes that trades data availability for space efficiency. vSAN's approach provides the best of both worlds. For more information, see the post: "Erasure Codes in VMware vSAN versus Storage Arrays."

Operations

What is the primary user interface (UI) used to configure and monitor vSAN?

The vSphere Client is used to perform nearly all configuration and monitoring tasks. For automation and repeatable tasks at scale, PowerCLI is a good option. Additional insight can be achieved using VCF Operations, which provides enhanced analytics from the data it collects from vCenter Server.

How do I monitor the health of a vSAN cluster?

vSAN features a comprehensive health service appropriately called Skyline Health for vSAN that actively tests and monitors many items such as hardware compatibility, verification of storage device controllers, controller queue depth, and environmental checks for all-flash and hybrid vSAN configurations. In vSAN 8 U1, we introduced a new Skyline Health Scoring, Diagnostics and Remediation dashboard that helps solve the challenge of prioritizing identified issues in a vSAN cluster. See the post "Skyline Health Scoring, Diagnostics and Remediation in vSAN 8 U1" for more information.

"Skyline" is a legacy term. In vSAN for VCF 9.1, you may notice that it has been renamed to simply "vSAN Health."

What is the health cluster scoring dashboard, and how does it work?

Introduced in vSAN 8 U1 (OSA & ESA), the new cluster health status and troubleshooting dashboard helps answer the basic questions of, "*Is my cluster and the workloads it serves in a health state?" and if not, how sever is the condition? …and should the issue be resolved?* It provides information that traditional health alerts alone cannot. For each vSAN cluster, the Skyline Health for vSAN will provide a quick at-a-glance score of the condition of a cluster so an administrator can easily determine if all is good, and if not, how impactful any identified issues are. The mechanism uses a sophisticated method of weighing triggered health checks, and aligns them with common pillars of responsibility, such as data availability, performance, capacity utilization, efficiency and compliance. It will then provide the most important, impactful triggered health findings in an order of priority so that an administrator can resolve issues quickly and easily. See the post "Skyline Health Scoring, Diagnostics and Remediation in vSAN 8 U1" for more information.

What vSphere maintenance mode should I use in vSAN?

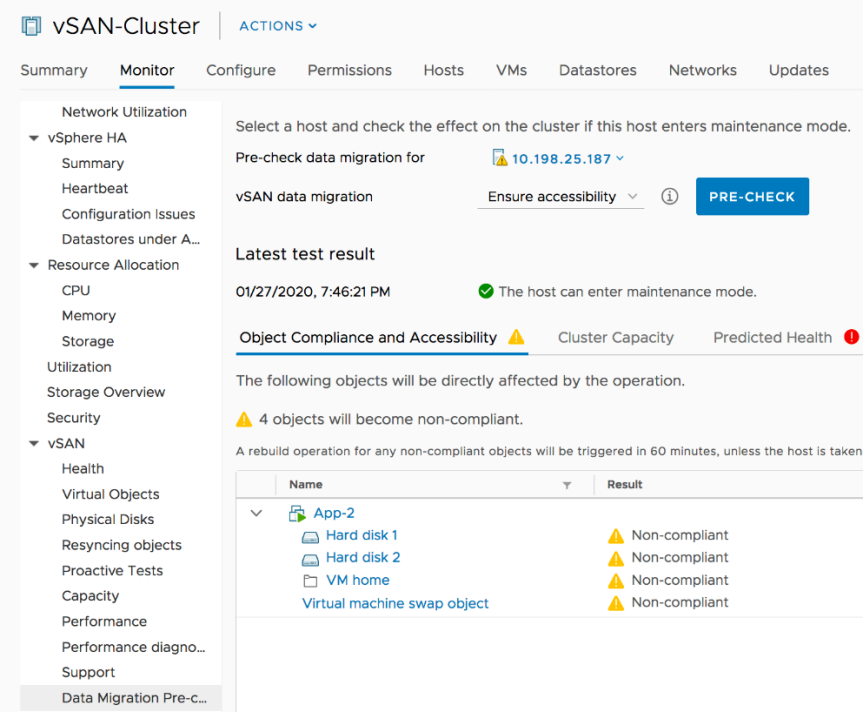
When a host that is part of a vSAN cluster is put into maintenance mode, the administrator is given three options concerning the data (vSAN components) on the local storage devices of that host. The option selected has a bearing on a couple of factors: The level of availability maintained for the objects with components on the host and the amount of time it will take for the host to enter maintenance mode. The options are:

- Ensure accessibility (default)
- Full data migration
- No data migration

In most cases, simply use the default "Ensure Accessibility." This is the most flexible option, and minimize the movement of data. Details on how data is handled are provided in the vSAN documentation. In summary, the default option, "Ensure accessibility," is used when the host will be offline for a shorter period of time. For example, during maintenance such as a firmware upgrade or adding memory to a host. "Full data migration" is typically appropriate for longer periods (hours or days) of planned downtime or the host is being permanently removed from the cluster. "No data migration" commonly allows the host to enter maintenance mode in the shortest amount of time. However, any objects with an FTT=0 with components on the host going into maintenance mode are inaccessible until the host is back online.

How would I know what VMs, and objects would be impacted when a host enters maintenance mode?

Before moving a host into maintenance mode, an administrator can use the Data Migration Pre-Check feature to assess the impact of the maintenance mode option. This will help you determine the impact on Object Compliance and Accessibility, Cluster capacity, and Predicted Health.



Can vSAN upload information about my environment to help improve a support case opened?

Yes, vSAN allows customers to upload anonymous information about their environments to VMware, which provides several benefits including:

Time spent on gathering data is reduced when a support request (SR) is opened with VMware Global Services (GS - previously known as "GSS"). A GS Technical Support Engineer (TSE) can utilize vSAN Support Insight <link> to view current and historic data about a customer's environment and start troubleshooting efforts sooner, which leads to faster resolution

times. vSAN online health checks identify issues specific to a customer's environment and suggest resolutions. These online health checks can also make recommendation changes that adhere to VMware "best practices."

VMware receives anonymous data from a large number of environments that can be utilized to identify trends, potential software bugs, and better understand how products are used. Bug fixes can potentially be developed faster, and improvements are implemented to provide a better overall customer experience.

In vSAN 8 and newer, one can benefit from some of the offerings through enrolling in the CEIP without ticking the checkbox. vSAN 8 introduces "Proactive Insights" that will provide improved health awareness for environments with a vCenter Server with connectivity to the internet, but not enrolled in CEIP.

Can isolated environments use the built-in health features for vSAN found in vCenter Server?

Yes. Earlier versions of vSAN introduced the Health Diagnostics tool allows administrators of isolated environments to manage an environment in a way that is like a fully cloud connected environment. The tool can gather the latest signature libraries at a time and frequency that is best for the customer and run periodically in an environment to detect the latest alerts and updates provided by VMware. It will also allow VMware Global Support Services to resolve issues in isolated environments more effectively, as it can assist with the data gathering process.

Does vSAN work with VMware vSphere Lifecycle Manager (vLCM)?

Yes, vSAN is fully integrated with vLCM.

In stretched cluster and 2-Node environments, should I back up a vSAN virtual witness host appliance?

No. Backups, restores, clones, and snapshots of a Witness Host are not supported. Use the 'Change witness host' function in the vSphere Client to deploy a new Witness Host when there is an issue with the existing Witness Host.

How can I gracefully power down a vSAN cluster?

vSAN 7 U3 introduced an automated workflow that will guide the user through the process of gracefully shutting down a cluster. It includes several prechecks and other guidance to ensure that the shutdown and power-up process is simple and predictable. vSAN 8 enhanced the workflow to ensure there are no circular dependencies, and in vSAN 8 U1, and PowerCLI 13.1, new cmdlets are available to perform this task via PowerCLI. See the post "Automation improvements using PowerCLI with vSAN 8 U1" for more information.

Does upgrading from vSAN 8 U3 to vSAN 9.0 introduce an on-disk format (ODF) change?

Yes. vSAN clusters running vSAN 8 U3 use an ODF format version of 20, and SAN in VCF 9.0 will use an ODF of 21. vSAN in VCF 9.1 will use an ODF format of 22. This ODF upgrade is a simple metadata update and will not invoke any rolling reformat or data movement. For more information, see the post: "Upgrading On-Disk and Object Formats in vSAN."

Performance

What is the "Number of Disk Stripes per Object" rule in a vSAN storage policy?

For the vSAN ESA, this storage policy rule can be ignored. For more information, see the post: Stripe Width Storage Policy Rule in the vSAN ESA. For vSAN in VCF 9.1, it is not even available for Auto-RAID enabled clusters.

What is the recommended way to test vSAN performance?

VMware provides a tool called HCIBench. It is essentially an automation wrapper around popular and proven synthetic test utilities. With HCIBench, you can either invoke Vdbench or Flexible I/O tester (FIO) to automate performance assessment in an HCI cluster.

HCIBench simplifies and accelerates proof-of-concept (POC) performance testing in a consistent and controlled manner. The tool fully automates the process of deploying test VMs, coordinating workload runs, aggregating test results, and collecting data for troubleshooting purposes. The output from HCIBench can be analyzed by the Performance Diagnostics feature. See this VMware Knowledge Base article for more information: vSAN Performance Diagnostics (2148770)

HCIBench can be used to evaluate the performance of vSAN and other HCI storage solutions in a vSphere environment.

Use HCIBench to run performance tests across the cluster instead of running a workload from a single VM. HCIBench can be configured to deploy and distribute multiple VMs across the hosts in an HCI cluster to provide more realistic and accurate test results. It is also recommended that HCI bench be run on a cluster prior to introducing it into production, then save the results for reference, if needed. This can help identify if there are any issues prior to introducing the cluster into production and can serve as a nice reference point if there are any issues later, and one wants to compare current test results to test results after initial deployment.

How does vSAN minimize the impact of data resync operations when a device or host fails?

There are several mechanisms in place to dynamically balance and prioritize virtual machine and resync I/O. It is important to maintain adequate performance while providing resources for resync operations to restore resilience.

When there is contention for I/O in the hardware storage stack, vSAN guarantees approximately 20% of the bandwidth to resync operations while virtual machines utilize the remaining 80%. If there is no contention for bandwidth, resync operations can consume more bandwidth to reduce resync times. Virtual machines can use 100% of the bandwidth when there are no resync operations occurring. For more information, see "Adaptive Resync in vSAN."

When using the Express Storage Architecture (ESA), VMware introduces adaptive network traffic shaping for resynchronizations. This is due to the tremendous efficiency of the vSAN ESA, and its ability to push higher rates of data through the network stack. See the post: "Adaptive Network Traffic Shaping with the vSAN Express Storage Architecture" for more information.

vSAN ESA also minimizes the impact of a discrete storage device failure, thus greatly reducing the amount of data need to be resynchronized. For more information, see the post: "The Impact of a Storage Device failure in vSAN ESA versus OSA."

Does vSAN require manual intervention to balance data across the cluster?

No. Users can configure proactive rebalancing as an automated action at the cluster level to let the cluster balance the data out as it sees fit. When toggled on, this will replace the previous behavior of proactive rebalancing, which included tripping an alert in the health service, followed by the need for manual intervention by the administrator to perform the rebalance.

What is the best way to troubleshoot performance issues in vSAN?

The "Troubleshooting vSAN Performance" guide will provide a framework for proper isolation of performance issues for more accurate diagnostics of the performance issue. vSAN contains several built-in tools to help gather data in the troubleshooting performance effort, including the vSAN performance metrics, I/O Insight, vsantop, and in the VM I/O Trip Analyzer.

How do I get more detailed performance metrics for vSAN?

vSAN 8 U1 introduces the ability to view time-based performance metrics that use a **30 second sampling interval**. This is a 10x improvement over the 5-minute sampling interval used in previous versions. This will produce performance graphs that are more representative of system behavior and will help to identify when performance issues occur. This is available in both the OSA and ESA of vSAN 8 U1. See the post: "High Resolution Performance Monitor in vSAN 8 U1" for more information.

Security

Is encryption supported with vSAN?

Yes, vSAN supports Data-At-Rest Encryption and Data-in-Transit Encryption as cluster-based toggles, and use an AES 256 cipher. For vSAN Data-at-Rest encryption, data is encrypted as it is persisted to disk. For deduplication activity, this data is decrypted using special threads low in the stack for optimal performance. vSAN Data-In-Transit encryption will encrypt the data as it is transmitted between hosts in a cluster, and/or between a vSphere cluster mounting a vSAN datastore.

Does vSAN encryption require special hardware?

No, vSAN Encryption does not require any specialized hardware such as Self-encrypting drives (SEDs). Some drives on the vSAN Compatibility Guide may have SED capabilities, but the use of those SED capabilities is not supported.

Should vSAN encryption be enabled when first creating a cluster, or after workloads have been migrated?

If you are interested in using vSAN encryption, it is most efficient to enable this during the initial configuration of a vSAN cluster, prior to any VM migration.

What are the prerequisites to enable vSAN Data-at-Rest Encryption?

vSAN data-at-rest and data-in-transit encryption requires a key provider is required to enable and use vSAN encryption. Nearly all KMIP-compliant KMS vendors are compatible, with specific testing completed for vendors. vSAN can also use the vSphere Native Key Provider, which is the common key provider for vSAN Data-at-Rest Encryption for customers with or without external KMS solutions. This gives basic key provider functionality built directly into vSphere, and generally provides a more flexible and robust method of key distribution for vSAN clusters. The Native Key Provider does not provide KMIP services for things external to vSphere. See the vSAN Encryption Services document for more information.

For vSAN, turning on encryption is a simple matter of clicking a checkbox. Encryption can be enabled when vSAN is enabled or after, with or without virtual machines (VMs) residing on the datastore.

Always use Trusted Platform Modules (TPMs) to cryptographically store issued keys on host, for better resilience if a key provider is offline.

How does vSAN Encryption differ from vSphere VM Encryption?

vSAN encryption operates at the storage level and encrypts the vSAN datastore. VM Encryption operates on a per-VM basis and performs encryption on in-flight I/O i.e., encrypts IO as it is generated from the VM.

VMs encrypted with vSphere VM encryption can be deployed to a vSAN datastore just like other datastore types such as VMFS and NFS. However, vSAN space efficiency features such as deduplication and compression will provide little to no benefit with these encrypted VMs. **vSAN encryption services are highly preferred over vSphere encryption capabilities when VMs are powered by vSAN.**

Can I used vSAN Encryption and vSphere VM Encryption at the same time?

It is possible, but is not recommended. You will lose all space efficiency capabilities by enabling vSphere Encryption on top of vSAN Encryption Services, and introduce unnecessary complexity. Stick with vSAN Data-at-Rest Encryption and vSAN Data-in-Transit Encryption for your encryption needs. See the vSAN Encryption Services document for more information.

How is performance impacted when using vSAN encryption services?

A detailed explanation on this matter can be found on the blog post: "Performance when using vSAN Encryption Services." (this post predates the introduction of vSAN ESA). The level of impact will depend also if a cluster is running the Original Storage Architecture (OSA) in vSAN, or the Express Storage Architecture (ESA). vSAN **ESA is inherently more efficient at resource utilization for these types of data services, and will have a substantially less impact on resources when compared to vSAN OSA.** We've found that enabling Data-at-Rest Encryption on vSAN ESA results in little to no impact on the performance of your workloads, with minimal increases in CPU resources.

Does encryption in the vSAN ESA perform better than in the OSA?

Yes, much better. Due to the architectural design of the ESA, it uses fewer CPU cycles and I/Os to encrypt a given amount of data. This translates to improved performance, and reduced utilization.

Does enabling encryption consume any additional capacity overhead?

No. There will not be any additional capacity overhead used when vSAN encryption is enabled on a cluster.

Does vSAN encrypt object data with different keys?

In more recent versions of vSAN ESA, yes. While vSAN ESA uses a cluster-wide Disk Encryption Key (DEK), the object data has object-specific keys that are used. **This ensures that object data in VM uses different keys than any other VM.** Note however that a rekey operation will perform a rekey across all data on the cluster, and does not give a choice for rekeying discrete objects.

Should I deploy an external Key Management Service (KMS) server on the vSAN datastore that will use the same KMS for key management?

Deploying a third-party external KMS solution on a vSAN datastore encrypted by vSAN Data-at-Rest encryption that uses that KMS is discouraged. When a vSAN host with encryption enabled is restarted, it requests a new Host Key and Key Encryption Key (KEK) from the KMS. If the KMS is not online to provide these keys, the host will not be able to read the encrypted data. This creates a circular dependency resulting in no access to encrypted data. For clusters using an external KMS, keys can be persisted to the hosts with additional configuration settings, but may compromise some of your security objectives. Regardless, we do recommend Trusted Platform Modules (TPM) are installed on each host in your environment. This ensures that if keys are stored locally, that they are cryptographically secured on the TPM of each host. See the vSAN Encryption Services document for more information.

Should I deploy the vSphere Native Key Provider (NKP) on the vSAN datastore that will use the same NKP for key management?

Running the vCenter server that provides the NKP services on a vSAN datastore that will use the same NKP for key management is indeed supported. The NKP avoids some of the technical challenges and circular dependencies associated with attempting the same configuration using an external KMS. We do recommend Trusted Platform Modules (TPM) are installed on each host in your environment. This ensures that if keys are stored locally, that they are cryptographically secured on the TPM of each host. See the vSAN Encryption Services document for more information.

What is vSAN Data-in-Transit encryption?

Data-in-transit encryption is a cluster-wide feature that transmits vSAN storage traffic in-flight between vSAN hosts and/or the hosts mounting a vSAN datastore. As of vSAN in VCF 9.1, Data-in-Transit Encryption is supported in aggregated vSAN HCI deployments and in vSAN storage clusters.

Data-in-transit securely encrypts vSAN data traffic that traverses across hosts using 140-3 Cryptographic modules.

Does vSAN Data-in-Transit encryption require the use of a key provider such as an external KMS or NKP?

No, Data-in-transit encryption does not require a key provider such as an external KMS or NKP. With DiT encryption, the vSAN hosts are responsible for creating the symmetric keys. This is a process that happens transparently, without any administrative effort.

What happens when a vCenter server managing a vSAN datastore with encryption enabled is offline?

There is no impact on the virtual machines running on the vSAN datastore with encryption enabled. After vSAN Encryption is configured, vSAN hosts communicate directly with the Key Management Server (KMS) cluster. If the original vCenter Server cannot be recovered, a new vCenter Server should be deployed as soon as possible.

For a more robust environment during failure conditions of a key provider, we recommend the use of Trusted Platform Modules (TPM) on each server. This will allow for the keys distributed to the hosts to be securely stored on persistent media (the TPM) and ensure keys can be used under these types of failure conditions.

What is the impact to the VMs running on a vSAN datastore with encryption enabled if the KMS is offline?

Key Encryption Key(KEK) is cached on the ESXi hosts' memory on booting. Hence there is no impact to the virtual machines till the hosts remain powered-on. If the hosts are restarted, the encrypted disk groups are unmounted and cannot be mounted until the KMS is restored. For a more robust environment during failure conditions of a KMS server or the vSphere Native Key Provider, we recommend the use of Trusted Platform Modules (TPM) on each server. This will allow for the keys distributed to the hosts to be securely stored on persistent media (the TPM) and ensure keys can be used under these types of failure conditions.

Do items such as backup and recovery work with vSAN encryption services?

Yes, as vSAN encryption was designed to maintain compatibility with other vSAN and vSphere features, as well as, 3rd-party products including data protection solutions. Data is encrypted or decrypted just above the physical storage device layer. APIs such as vSphere Storage APIs for Data Protection (VADP) and vSphere APIs for IO Filtering (VAIO) that are used for data protection and other solutions are located higher in the storage stack. **Data at this layer is not yet encrypted**. Therefore, compatibility with these solutions is maintained when vSAN encryption is enabled, and VMs can be protected, migrated, or replicated without issue.

If I use the datastore browser in the vSphere Client to download a VMDK from a datastore browser using vSAN Data-at-Rest Encryption, will the downloaded file be encrypted or unencrypted?

It will be downloaded unencrypted. Data-at-Rest encryption is intended to cryptographically secure data as it resides on the datastore, in an effort to prevent data leakage from disk or server theft. Any form of access using formal APIs (including downloading from the datastore, replication, etc), will fetch the data and save it to your target location in an unencrypted format. To reduce the threat of unauthorized downloading of files, used Role Based Access Control (RBAC) to secure the management plane.

Is two-factor authentication supported in vSAN?

2-factor authentication methods, such as RSA SecurID and Common Access Card (CAC), are supported with vSAN, vSphere, and vCenter Server.

Is vSAN part of a DISA STIG?

Yes, VMware vSAN is part of the VMware vSphere STIG Framework. The DISA STIG defines secure installation requirements for deploying vSAN on DoD networks. VMware worked closely with DISA to include vSAN in the existing vSphere STIG.

Has vSAN achieved FIPS certification?

In 2017, the VMware VMkernel Cryptographic Module achieved FIPS 140-2 validation under the National Institute of Standards and Technology (NIST) Cryptographic Module Validation Program (CMVP). In 2024, the validation for FIPS 140-3 was achieved.

vSAN consumes the VMware VMkernel Cryptographic Module when providing data-at-rest encryption thanks to the tight integration between vSAN and the ESXi kernel. When vSAN Encryption is enabled, the vSAN datastore is encrypted with FIPS Approved AES-256 utilizing the validated VMware VMkernel Cryptographic Module. This delivers FIPS compliance without the need for costly self-encrypting drives (SEDs).

How can storage devices used in a vSAN cluster be safely decommissioned, removing any residual data?

vSAN has ways to securely wipe storage flash devices decommissioned from a vSAN cluster. This is done through a set of PowerCLI commands (or API), providing an efficient and secure methodology to erase data in accordance with NIST standards. See the post: "vSAN – A Secure Fortress for your Data" for more information.

Does vSAN support the use of the vSphere Native Key Provider (NKP)?

Yes. vSAN supports the use of the vSphere NKP for key management. The vSphere Native Key Provider is a simple way for vSphere to create and manage encryption keys without using a traditional Key Management Server (KMS). It can be ideal for customers who have simple security requirements, have not already enabled encryption for their vSphere clusters, and want to do so in a secure and supported manner.

Can I still use my existing KMS for key management of a vSAN environment?

Yes. Full featured KMS solutions offer features and capabilities that are beyond the intention of the vSphere Native Key Provider. Depending on the environmental and customer security requirements, an external KMS may be the only way to achieve the specific customer security requirements. Customers with existing external KMS solutions often find they prefer the durability and simplicity of the NKP, and use the NKP as the key provider for vSAN Data-at-Rest Encryption.

Does vSAN support TLS?

Yes. See the link: Enable or Disable TLS Versions on ESXi Hosts for more information on how to use a specific version supported.

Should I use the vSphere NKP instead of a full-featured KMS solution?

It depends on the customer requirements, but in many cases, the specific customer requirements related to key management extend beyond the basic capabilities of the vSphere Native Key Provider (NKP). The NKP is ideal for customers who have simple security requirements and need basic key management for vSphere and/or vSAN only. The NKP is not a full featured KMS and does not support many of the capabilities found in a KMS.

How much bandwidth does a KMS introduce into an environment?

The key exchange process is relatively lightweight, introducing approximately 100KB/s between the KMS, the hosts and the managing vCenter Server. For reliable key distribution, persistent availability of keys is much more important than the bandwidth required. To ensure keys are readily available, VMware recommends the use of Trusted Platform Modules (TPM) installed on each host in a vSAN cluster. This will cryptographically store issued keys on the host should there be any connectivity issues with the KMS.

vSAN Protection and Recovery

What is vSAN Protection and Recovery?

vSAN Protection and Recovery (previously known as "vSAN Data Protection") is a capability that allows customers to easily protect and recover virtual machines. It uses the highly efficient, high performance snapshotting capabilities of the Express Storage Architecture (ESA) paired with a new intuitive User Interface to make local protection and restoration of VM's easy. It was released in vSAN 8 U3 and provided local protection of VMs on a vSAN datastore. It was enhanced in VCF 9.0, offering remote protection of VMs to a remote vSAN datastore, when paired with the use of the appropriate addon license. For more information, see the post: "vSAN Data Protection in VMware Cloud Foundation – The Solution You Already Own." Additional enhancements were introduced in vSAN for VCF 9.1.

What is new with vSAN Protection and Recovery for VCF 9.1?

In vSAN for VCF 9.1, vSAN Protection and Recovery adds new logical snapshot retention mechanisms to make snapshot retention easier and more scalable. Often known as "grandfather, father, son" (GFS), this new tiered snapshot retention model allows you to specify retention by hours, days, weeks, and months. vSAN Protection and Recovery also now supports vSphere tags to associate VMs with protection groups. For environments that are replicating data to a remote target, vSAN Protection and Recovery now supports the use of manual seeding for the initial replica, using a portable storage device. For more information, see the post: "VMware vSAN Protection and Recovery Enhancements for VCF 9.1."

My storage array can replicate data. Why is vSAN Protection and Recovery better?

The following are noticeable advantages that vSAN Protection and Recovery has over array-based solutions.

- **Granularity.** When replicating data using vSAN Protection and Recovery, you achieve fully granular per-VM protection and recovery.
- **Asymmetry.** The source and the target clusters can be dissimilar in size and hardware configuration.
- **Autonomy.** The replicas created by vSAN Protection and Recovery are fully autonomous copies at each site, meaning there is no dependency on one site versus the other – each of which can have up to 200 snapshots per VM per site.
- **Low RPO.** The architecture also allows for an extremely efficient 1-minute recovery point objective (RPO).

For practical example of how some of these capabilities can be easily used, see the post: "vSAN Data Protection in VMware Cloud Foundation – The Solution You Already Own."

What is required to use vSAN Protection and Recovery?

vSAN Protection and Recovery is a part of a VCF licensing entitlement and does not require any extra licensing. Some aspects of extending vSAN Protection and Recovery's abilities (such as vSAN-to-vSAN replication) may require add on licensing through Site Recovery Manager, or VMware Advanced Cyber Compliance (ACC).

Can I replicate VMs to a remote vSAN cluster while using a single vCenter Server?

No. The source and the target locations must use their own vCenter Server instance.

Is vSAN Protection and Recovery simply using vSphere Replication to protect VMs remotely?

No. While vSAN Protection and Recovery pairs a light-weight delta (LWD) mechanism of vSphere Replication with ESA snapshots, it offers abilities that are not possible with vSphere Replication alone. For example, vSphere Replication was limited to a snapshot limit of 24, while vSAN Protection and Recovery supports 200 snapshots per VM at each site. vSphere Replication was limited to protecting data on a per-VM basis, where vSAN Protection and Recovery uses the concept of protection groups. vSphere Replication didn't have any recovery API, whereas vSAN Protection and Recovery has a full protection API. And finally, vSphere Replication was not integrated into the vCenter Server Ui, where vSAN Protection and Recovery is fully integrated.

What would some typical examples of how vSAN Protection and Recovery could be used?

vSAN Protection and Recovery can be an ideal solution to augment your existing data protection strategies. The most common use cases for local protection might include:

- **Reverting existing VMs.** This helps address situations such as accidental VM misconfigurations, unsuccessful VM OS upgrades, or suspected malicious activity.
- **Restore existing VMs no longer registered in vCenter Server.** VMs that are accidentally deleted or moved can be easily restored to the existing cluster.
- **Clone VMs.** Cloning VMs can be an ideal way to support basic test and development workflows for all different types of teams, including Development and IT staff.

In vSAN for VCF 9.0, remote vSAN to vSAN replication (through the SRM or ACC add on license) opened new use cases, including:

- **Disaster recovery.** The use of VLR, and specifically VMware Live Site Recovery (VLSR) allow for data to be replicated to a remote site, and failover and failback be orchestrated through VLSR
- **Isolated recovery.** With the ability to specify local snapshots, or remote replicas with remote snapshots, this remote replication could provide an isolated environment for testing and recovery scenarios.
- **Long term archiving.** Remote replication will allow data to be stored with longer retention periods than may be possible for local protection.

vSAN in VCF 9.1 adds the ability to replicate from multiple source types (VMFS, NFS, vSAN) to a single vSAN target. Not only can this expand recovery use cases, but also serve as a way to migrate data from legacy storage when storage vMotion is not possible. For more information, see the post: "VMware vSAN Protection and Recovery Enhancements for VCF 9.1."

Can vSAN Protection and Recovery protect VMs from a remote location?

vSAN Protection and Recovery in vSAN 8 U3 (VCF 5.2) was limited to local protection only. The ability to protect VMs remotely is a part of VCF 9.0 and later when using the addon license. vSAN Protection and Recovery in VCF 9.1 introduces new features, such as multi-source replication. This gives the ability for source clusters using different storage types (e.g. VMFS, NFS, vSAN) the ability to replicate data to a target vSAN cluster.

Does vSAN Protection and Recovery protect against ESXi hosts compromises?

No. vSAN Protection and Recovery aims to provide an additional level of protection and flexibility to your VM data protection strategy. Concerns that regard host security should be addressed with the vSphere Security Hardening Guide.

What is a protection group?

vSAN Protection and Recovery uses the concept of Protection Groups, which allows administrators to achieve two objectives.

1. Group multiple VMs for easy and repeatable snapshot creation and management
2. Define and execute a policy of outcomes, such as the frequency of protection and retention schedules.

Protection groups also define if you wish to have local snapshots, local and remote snapshots, or just remote snapshots. This offers maximum levels of flexibility.

Can VMs participate in more than one protection group?

Yes. A VM can be a part of up to three protection groups

Can protection groups consist of multiple schedules?

Yes. A protection group can have as many as 10 schedules.

How can VMs be associated with a protection group?

One or more VMs can be individually added to a protection group. Or they may be added by dynamic naming assignment, where "*" and "?" characters can be used to include all VMs that match a naming pattern. In VCF 9.1, vSphere tags can be used to associate VMs with a given protection group.

Are the snapshots of VMs in a protection group taken at precisely the same time?

No. At this time, while a protection group will strive to take snapshots of the VM's that are a member of that protection group in accordance with the defined schedule, it will be a best effort operation, meaning that the snapshots may not be taken at precisely the same time.

What is snapshot immutability, and why does it exist?

In the context of vSAN Protection and Recovery, snapshot immutability refers to the inability of the snapshot to be modified or deleted. Snapshots can be made immutable by an optional toggle within the settings of a protection group. They exist to preserve the integrity of the snapshot against malicious activities and can serve as a basic way of recovering VMs during a ransomware attack.

Why not make all protection groups immutable?

While the immutability setting is extremely helpful for some use cases, it is not ideal for all situations. For example, when a protection group has the immutability setting enabled, one cannot edit or delete the protection group, change the VM membership, or edit/delete the snapshot.

How many snapshots can be created for a VM?

vSAN Protection and Recovery supports up to 200 snapshots per VM when snapshots are created using its UI and APIs. Note that if one uses the traditional UI or APIs, this will still be limited to 32 snapshots per VM. When replicating the VMs from one site to another, a VM can have up to 200 snapshots per VM at each site.

What happens when VMs reach the 200-snapshot limit?

The oldest snapshots will automatically expire when the maximum number of snapshots for a VM are hit. vSAN Protection and Recovery in VCF 9.1 introduces a new approach to snapshot retention that will make long term retention more logical and achievable.

How does the system protect against capacity management issues when allowing for so many snapshots?

vSAN Protection and Recovery will automatically pause snapshots if 70% of the cluster capacity is reached.

Once a VM is cloned from an existing snapshot, can it be protected using vSAN Protection and Recovery?

No. When new VMs are created from a snapshot, these are a linked clone, not a fully autonomous, independent clone. Therefore, snapshots from the linked clone cannot be taken. While the linked clones created from the cloning process are not eligible for protection using vSAN Protection and Recovery, these VMs can be protected using backup applications that use VADP, as manual snapshots of these linked clones can still be performed in the UI, or via API, such as VADP.

Can a VMDK from a VM that was cloned from a snapshot in vSAN Protection and Recovery be detached and attached to another VM?

No. VMs that are cloned from a snapshot via vSAN Protection and Recovery can have their virtual hardware adjusted, including the detaching of a VMDK. These detached VMDKs cannot be attached or mounted to other VMs.

Are there any disadvantages to having a system perform a lot of snapshots?

The snapshotting engine in vSAN ESA is extremely efficient and generally has very little impact on the performance of a VM. However, there can be cases where the frequency and number of snapshots may have an impact on the cluster. Snapshot frequency may dramatically **increase the rate of capacity usage if your data change rate is high.** As the capacity utilization of a cluster increases, it may need to perform more garbage collection processes, which may **temporarily impact performance** especially as available capacity becomes scarce.

It is recommended to choose modest levels of snapshot frequencies and retention rates so that change rate behavior can be better understood in your specific environment.

I see vSAN Protection and Recovery uses a virtual appliance. Won't this be a single point of failure for snapshots?

No. The virtual appliance, which runs PhotonOS powering a few containers, is used to help orchestrate snapshot activities and interact with vCenter Server. Metadata about the snapshot created by the snapshot service will be persisted to disk. This allows for the retention of information about the snapshots even if the snapshot appliance is unavailable, or corrupted.

In VCF 9.0 and later, a **single virtual appliance** is responsible for vSAN Protection and Recovery, replication, and DR orchestration courtesy of SRM or VMware Advanced Cyber Recovery.

What happens if the virtual appliance is accidentally deleted? Do I lose all my snapshots?

All information about the snapshots (metadata) is stored on disk, independent from the virtual appliance. A new appliance can be recreated. With the assistance of our Global Support team, the new appliance can be configured to understand the existing state of all existing snapshots and other metadata.

Why is this called "vSAN Protection and Recovery" if I've always been told that snapshots are not backups?

vSAN Protection and Recovery helps you protect and recover data to a previous state using data stored locally and can augment your existing backup and recovery strategy by offering all new levels of convenience and flexibility.

The confusion in the question stems from terms such as "backup" and "protection" that have been overly generalized by the industry and can have multiple meanings depending on the context in which they are used. When paired with extremely simplified statements such as "*Snapshots are not backups*" this leads to more misunderstanding, as the statement lacks the level of detail necessary to be accurate in all cases.

The most accepted data protection strategy generally involves a "3-2-1" rule. This refers to the notion of having three copies of data, with two backups using different media types or targets, and one copy living independently, outside the domain of failure. Organizations generally achieve this through VADP-based backup applications triggering hypervisor snapshots so that it can capture the state of a VM at a specific point in time and copy it to a target location that lives on different media and/or different locations, which achieves the objectives of a 3-2-1 rule.

vSAN Protection and Recovery uses the same snapshotting engine in ESA that is used by VADP-based backup vendors. vSAN Protection and Recovery can help augment protection and recovery strategies by creating, storing, and managing locally housed snapshots, as well as remote replicas and snapshots. It can provide convenient and flexible data recovery scenarios such as reverting a VM to a previous state, restoring a VM that has been deleted from inventory, or cloning a VM for other operational workflows.

Are vSAN snapshots crash consistent?

Yes, since vSAN Protection and Recovery is integrated in the I/O path of vSAN ESA, the snapshot data is automatically committed in a crash consistent manner, **without the need to stun the VM**. For more details, see the FAQs in the "Availability" section of the vSAN FAQs.

Can vSAN snapshots taken with vSAN Protection and Recovery create an application-consistent snapshot?

No. At this time, vSAN Protection and Recovery can only create crash consistent snapshots. The option of creating an application-consistent snapshot is only available performed within the vSphere Client UI by highlighting the VM, or using a VADP-based 3rd party backup. Application consistent snapshots are achieved when “Quiesce guest file system” is selected, **and** the supporting VM operating system has the mechanisms in place to achieve the quiescing (e.g. “Microsoft Volume Shadow Copy Service” or VSS).

Can Site Recovery Manager be used with vSAN Protection and Recovery?

Yes. VMware Site Recovery Manager uses many elements of vSAN Protection and Recovery, but adds the ability for full site failover and failback through recovery plans, etc. VMware Cyber Recovery also uses vSAN Protection and Recovery, but adds the ability to establish an on-premises, customer owned clean room for cyber security events.

I’m confused about what capabilities come with my VCF licensing. Can you provide a quick reference?

The following is a brief overview of licensing as it relates to protection capabilities. See the appropriate licensing guides for more details.

- VCF licenses. Includes local protection through vSAN Protection and Recovery (previously known as “vSAN Data Protection.”)
- VCF + SRM license. Includes local protection, remote replication, and DR orchestration abilities.
- VCF + ACC. Includes local protection, DR orchestration that comes with SRM, and full cyber recovery capabilities.