IMPLEMENTING VMWARE VSAN ON HPE PROLIANT SERVERS WITH VMWARE VSPHERE

Best practices and general build information for deploying and configuring VMware vSAN 6.7 on the HPE ProLiant platform

CONTENTS

INTRODUCTION

VMware vSAN™ provides a software-defined storage solution that supports the storage needs of virtual machines in a vSphere cluster and provides support for key vSphere capabilities such as VMware vSphere® vMotion®, VMware vSphere® High Availability (vSphere HA), and VMware vSphere® Distributed Resource Scheduler™ (DRS). vSAN aggregates the local server storage of multiple HPE ProLiant servers within a vSphere cluster into a shared storage pool that is available to all hosts within that vSphere cluster. Hard disk drives (HDDs) and flash-based storage, such as solid-state drives (SSDs), PCIe accelerators, and NVMe storage from servers in a vSAN-enabled vSphere cluster contribute to the performance and capacity of the vSAN solution.

NOTE
For the remainder of this paper, flash-based storage will be generalized and referred to as SSDs.

This paper focuses on the requirements for a successful vSAN implementation on HPE ProLiant servers with vSphere 6.7. For earlier versions of vSAN, please refer to this White paper. When planning a vSAN implementation, consider the configuration and support restrictions discussed in this paper. Additionally, this paper assumes that you are familiar with vSphere and HPE ProLiant servers.

VMWARE VSAN OVERVIEW

VMware vSAN, as its name implies, provides an alternative to traditional shared storage solutions, such as iSCSI, Fibre Channel, and Fibre Channel over Ethernet (FCoE), by utilizing the local storage available to each HPE ProLiant server in the vSphere cluster. vSAN enables you to seamlessly extend virtualization to storage, creating a hyper-converged solution that simply works with your existing tools, skillsets, software solutions and hardware platforms. Operations become easier with fewer tasks and intelligent automation that can be managed through one tool and a unified team, allowing you to respond to business demands faster and more intelligently. vSAN also helps enterprises prepare for tomorrow's IT demands—whether expanding to the public cloud or rapidly deploying the latest flash and server technologies.

Data Access
When configuring vSAN, vSphere administrators define disk groups in the VMware vSphere® Web Client. Up to five disk groups can be configured on an VMware ESXi™ host in a vSAN cluster with each disk group being comprised of one caching tier SSD and up to seven-capacity tier HDD/SSD. Each disk group is presented as a vSAN shared datastore, which are utilized to store virtual machine objects such as virtual disks, configuration files, and snapshots.

The data on a vSAN datastore is accessed in one of two methods. If the requested data is local to the host, then the I/O is serviced directly from the local disks. However, if the data is remote, then the data is read from or written to the remote host over the Ethernet network.

In a hybrid configuration, the SSDs in the caching tier form a caching layer with 70 percent dedicated for read cache and 30 percent for write buffer. In an all-flash configuration, the caching tier only serves as a write buffer while the capacity tier is used for reading cache and storing data. The read cache keeps a list of commonly accessed disk blocks in order to reduce I/O read latency. If the requested block is not in cache, it is retrieved directly from the capacity layer. Important to note is that a host can access the read cache of any of the hosts in the same cluster. vSAN maintains a directory of cached blocks that can be referenced to find the requested block and then retrieved over the vSAN network.

Write operations go into the write buffer first and later de-staged to the disks in the capacity layer. When a write is initiated, it is sent to the local cache on the host and additionally to the write cache on a remote host. In the event of a failure, this ensures a copy of the data is available. Once those writes land on both hosts, the write operation is acknowledged back to the application. The data then is de-staged to the capacity layer at regular intervals.

High Availability
vSAN high availability is addressed by replicating the data to one or more additional hosts within the vSAN-enabled cluster. In addition to acting as read/write cache, the SSDs in the caching layer also support the vSAN replication activities by acting as the initial write copy target for the replicated data.

vSAN provides the ability for an administrator to define the number of complete mirror copies of data, referred to as replicas, within the vSAN cluster. The administrator defines the desired number of replicas on a per-VM basis, which provides the flexibility to adjust the desired level of protection based on the criticality of the VM data.

Scale-out Storage

As new HPE ProLiant servers are added into a vSAN cluster, the potential capacity of existing shared datastores is increased. Up to 32 nodes can be added into the cluster in vSAN 5.5, which matches the number of hosts supported in a vSphere cluster. vSAN 6.x supports up to 64 nodes.

In addition to extending existing shared datastores, scale-out is possible by creating new datastores through the addition of new hosts into the cluster, by adding additional disks into existing members of the cluster, or by performing a combination of those two options. Depending on the vSphere administrator preference, the configuration and addition of the new storage can be performed automatically by VMware vCenter Server® or manually by the administrator.

Limitations

As discussed previously, vSAN supports shared datastores and key vSphere capabilities such as VMware vSphere vMotion, vSphere High Availability (vSphere HA), and vSphere Distributed Resource Scheduler. However, there are some limitations that should be considered when determining the appropriate storage solution for your environment. The following list describes key limitations:

- vSAN does not support Distributed Power Management (DPM).
- Based on the cluster size, disk group, and disk count restrictions, there is a physical limit to the capacity support of a vSAN configuration.
- vSAN and VMware NSX® can co-exist on the same vSphere infrastructure but NSX does not support vSAN network traffic.
- vSAN requires a minimum of three vSphere hosts in each cluster.
- vSAN storage is only available to the hosts within the vSAN cluster.
- Up to 100 VMs per host in vSAN 5.5. Up to 200 VMs per host in vSAN 6.x (Refer VMware's vSphere® Configuration Maximums page).
- Blade systems with direct attached storage systems are supported in vSAN 6.x. This support does not exist in vSAN 5.5.
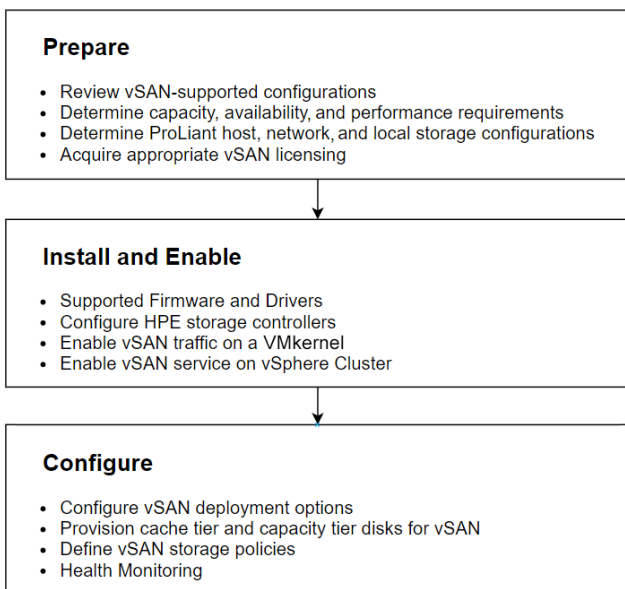
VSAN ENABLEMENT PROCESS

Prepare
- Review vSAN-supported configurations
- Determine capacity, availability, and performance requirements
- Determine ProLiant host, network, and local storage configurations
- Acquire appropriate vSAN licensing

Install and Enable
- Supported Firmware and Drivers
- Configure HPE storage controllers
- Enable vSAN traffic on a VMkernel
- Enable vSAN service on vSphere Cluster

Configure
- Configure vSAN deployment options
- Provision cache tier and capacity tier disks for vSAN
- Define vSAN storage policies
- Health Monitoring

FIGURE 1. vSAN enablement process

Prepare

Determining the hardware components and physical configuration for a vSAN implementation is not a trivial task. A large number of factors must be taken into account. This section provides an overview of the planning steps required to implement a supported vSAN configuration, highlights unique factors to consider for HPE solutions and provides references to additional resources. The following steps define a rough outline and recommended steps to implement a successful vSAN configuration. It will become clear as the administrator performs the steps that the factors and requirements discussed in each step have multiple interdependencies that influence each other. In most cases, it will be an iterative process to finalize the configuration.

1. Review possible options for building a supported configuration: To implement a fully supported vSAN configuration, a vSAN Ready Node configuration can be implemented or a custom configuration (Build Your Own based on Certified Components) can be defined if the following requirements are met:

    – HPE ProLiant server is listed on the VMware® Compatibility Guide for Systems/Servers with support for the targeted vSphere release.

    – HPE ProLiant server supports the targeted storage controller and drives. This can be determined by working with an HPE sales representative.

    – Storage controllers, HDDs, and SSDs are listed on the VMware Compatibility Guide for vSAN. The VMware Compatibility Guide (VCG) has specific hardware listed along with their supported firmware and drivers. VMware is emphasizing the importance of following the VCG to avoid issues stemming from unsupported configurations.

    vSAN-certified solutions

    HPE has certified several vSAN Ready Node configurations. HPE Ready Nodes are certified combinations of HPE ProLiant servers, RAID/Storage controllers, NICs, and drives specific for vSAN usage. These configurations are categorized into Ready Node Profiles based on VMware's vSAN Hardware Quick Reference Guide. This guide provides a reference for selecting the right Ready Node server for your environment.

    To build your own server configuration, HPE server components have been certified and listed on the VMware Compatibility Guide. See the "Build Your Own based on Certified Components" section of the VCG.

    Refer to the VMware Compatibility Guide for the latest list of vSAN certified offerings.

2. Size and define the characteristics of the vSphere cluster: Initially, the traditional planning exercises required to define the compute, memory, and networking requirements of a vSphere cluster should be completed. At the end of this step, a rough number of ESXi hosts, their CPU and memory configuration, and the physical and virtual networking should be defined. In addition, the virtual machines/application storage capacity and storage performance requirements along with their I/O profile should be well understood. The configuration defined in this step should be viewed as a starting point, which will likely be altered when the vSAN requirements and needs are considered. See VMware's TCO and Sizing Calculator vsantco.vmware.com/vSAN/si/siev.

    – A minimum of three hosts are required for vSAN, although in a ROBO configuration you can have two hosts with an external witness. This minimum configuration only allows for a single host failure. HPE recommends having four or more hosts for additional availability options (failures, maintenance, etc.).

    – The minimum memory requirement for a host is 6 GB.

    – A dedicated 1 Gb or 10 Gb network for vSAN traffic is required with a recommendation for 10 Gb networks in most environments. For all flash configurations, the 10 Gb network is required due to the increase in network traffic.

    – Load balancing is not supported for vSAN network traffic. However, it is recommended to include more than one adapter and to use NIC teaming to provide redundancy. This means configuring physical NICs in an active/standby configuration.

    – vSAN requires multicast traffic be enabled on the Layer 2 network between ESXi hosts in a cluster. Layer 3 multicast is not required.

    – USB and SD devices are supported as ESXi boot devices in vSAN.

3. Size and define the characteristics of the vSAN configuration: Capacity, performance, and availability requirements for storage, VMs, and networking are the primary factors to consider in this step. The VMware white paper, "VMware vSAN Design and Sizing Guide" provides a detailed analysis of each of the factors and provides formulas to assist in the planning process.
    See storagehub.vmware.com/t/vmware-r-vsan-tm-design-and-sizing-guide-2/.
    Information to take into account during the planning exercises in this step:

    – If ESXi will be installed on a local disk, that disk cannot be utilized for a vSAN disk group. Any RAID level supported by the storage controller can be used on the install disk(s).

    – It is not required for all ESXi hosts in the cluster to contribute storage to the vSAN cluster; however, a minimum of three hosts must contribute storage.

    – Achieving optimal performance depends on the I/O load. HPE recommends ensuring that the storage controllers have a queue depth of at least 256.

    – vSAN supports storage controllers in two modes: either pass-through mode or RAID 0 mode. Pass-through mode is also commonly referred to as HBA mode.

    – vSAN supports multiple HPE controllers on a single host. This allows for disk configurations beyond the limit of a single controller.

    – HPE SAS expanders are supported with vSAN with maximum of 16 drives.

    – HPE recommends having the cache size of the vSAN datastore be at least 10 percent of the anticipated consumed capacity. This ratio should be calculated before the number of failures to tolerate setting is considered. This applies to both hybrid and all-flash configurations.

    – Up to five disk groups can be configured on an ESXi host in a vSAN cluster with each disk group being comprised of one caching tier SSD and up to seven-capacity tier HDD/SSD.

    – With the limitation of one caching tier SSD per disk group, HPE recommends having multiple disk groups to improve performance and to protect against failure.

    – SSDs are categorized by their endurance and performance classes. HPE offers a variety of vSAN certified SSDs (SAS, SATA, NVMe, and PCIe). See VMware's VCG here.

    At the end of this step, the questions below should be answered, and the initial design adjusted accordingly:

    – How many shared datastores are required and what are the minimum raw capacity requirements for each of those?

    – What is the desired caching-to-capacity disk ratio?

    – How many VMs can be supported by the cluster and per host?

    – What is the expected host CPU and memory overhead for vSAN processing?

    – What are the networking requirements for vSAN?

    – What is the minimum number of ESXi hosts required?

4. Define the hardware configuration: In this step, the ESXi host compute, memory, and storage requirements should be mapped to hardware. As mentioned in step 1, the VMware Compatibility Guide should be referenced to determine the supported hardware with vSAN. Keep in mind that depending on the HPE ProLiant server and drive tray configuration, the number of physical disks that can be supported by each system will vary. See VMware's vSAN Hardware Quick Reference Guide for more information. HPE supports vSAN configurations with both HPE Smart Array controllers and Smart Host Bus Adapters (HBAs). In either case, when a PCIe standup controller is utilized, it should be placed in a PCIe Gen3 slot for optimal performance.

5. Procure appropriate vSAN licensing: vSAN licenses have a per CPU capacity. When a license is applied to a cluster, the total number of CPUs in the participating hosts count towards that license capacity. See VMware's vSAN Licensing Guide for additional information.

Install and enable

Firmware and drivers: To ensure you are running a supported vSAN configuration, update the drivers and firmware on the storage controllers, SSDs and/or HDDs to match the supported versions listed on the VMware Compatibility Guide for vSAN.

See HPE's white paper on Deploying and updating VMware vSphere on HPE ProLiant Servers.

Storage controllers: Before enabling vSAN, the storage controllers must be configured in either pass-through or RAID 0 mode. VMware recommends running in pass-through mode when possible so that the hypervisor is given direct access to the physical disks. vSAN can then setup a distributed configuration across all hosts in a cluster. In RAID 0 mode, single disk RAID 0 sets will need to be configured for each disk using HPE Smart Storage Administrator. In this case, vSAN will not manage hot-plug operations (instead relying on the storage controller's software to manage that), and local SSDs will need to be tagged as SSD before they can be claimed in vSAN. HPE SSA can be used to configure the RAID 0 sets. HPE Smart Storage Administrator CLI is also supported and can be used:

Create a single disk RAID 0 set:
```
# esxcli ssacli cmd -q "ctrl slot=<slotnumber> create type=ld drives=<driveaddress> raid=0"
```

For more information about SSACLI, see the HPE VMware Utilities User Guide.

Additionally, in RAID 0 mode, it is recommended to configure the storage controller cache for 100 percent read cache.
```
# esxcli ssacli cmd -q "ctrl slot=<slotnumber> modify cacheratio=100/0"
```

This effectively disables write caching to minimize conflicts with vSAN's caching layer. In RAID 0 mode, there are two supported vSAN Smart Array configurations that can be configured using SSACLI commands:

1. Lower q-depth writes—generally Smart Array default (disables drive write cache):
   ```
   # esxcli ssacli cmd -q "ctrl slot=<slotnumber> modify dwc=disable"
   ```

2. High q-depth writes (enables drive write cache):
   ```
   # esxcli ssacli cmd -q "ctrl slot=<slotnumber> modify dwc=enable"
   ```

Networking: The vSAN service must be enabled on a VMkernel port for each host. HPE recommends dedicating a VMkernel to vSAN traffic and separating it from the management traffic. This happens automatically if the vSAN toggle is checked when creating a new cluster.

If vMotion, Storage vMotion, or HA is desired on the cluster, an additional VMkernel port is needed. vSAN traffic requires a 1 Gb or 10 Gb network, and standard switches are supported but vSphere Distributed Switches are recommended.

Enable vSAN on a cluster: When creating a new cluster in vSphere, check the vSAN radio button.
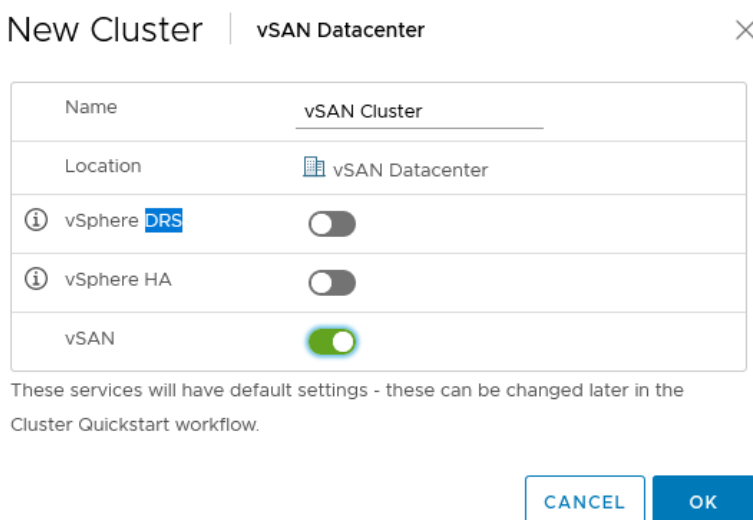

FIGURE 2. Creating a new cluster with vSAN enabled

Next, add ESXi hosts that are going to be a part of the vSAN datastore to the cluster. Once the hosts are added, the vSAN configuration wizard will open.

Configure

1. The first step is to configure a distributed switch. There is an option to create a new distributed switch or use an existing one. A new port group will be created on each host in the cluster using selected distributed switch. A new VMkernel port with the vSAN service enabled will also be automatically created on each host within the cluster to facilitate the flow of vSAN related traffic.
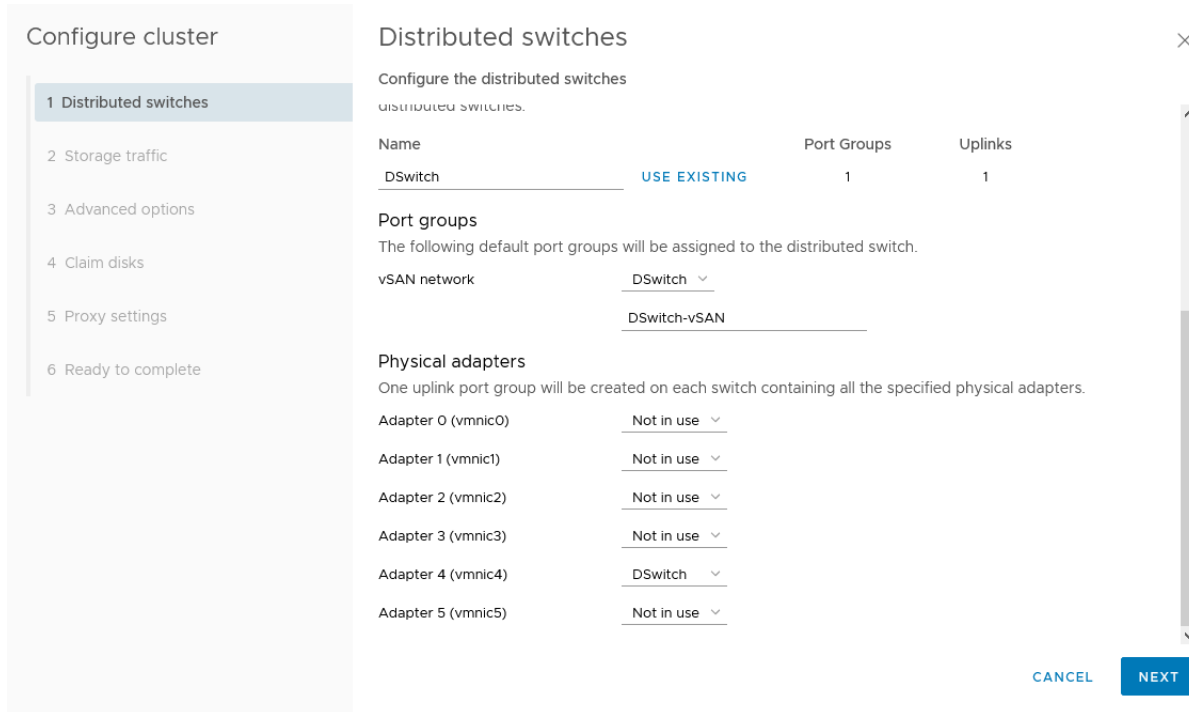

FIGURE 3. Configuring Distributed Switches

2. On the **Storage traffic** screen, check the VLAN box and specify the VLAN ID to use a separate VLAN for vSAN traffic. Otherwise, leave the box unchecked. For protocol, select IPv4 (default) and select DHCP as the IP type in IPv4 configuration and click Next.
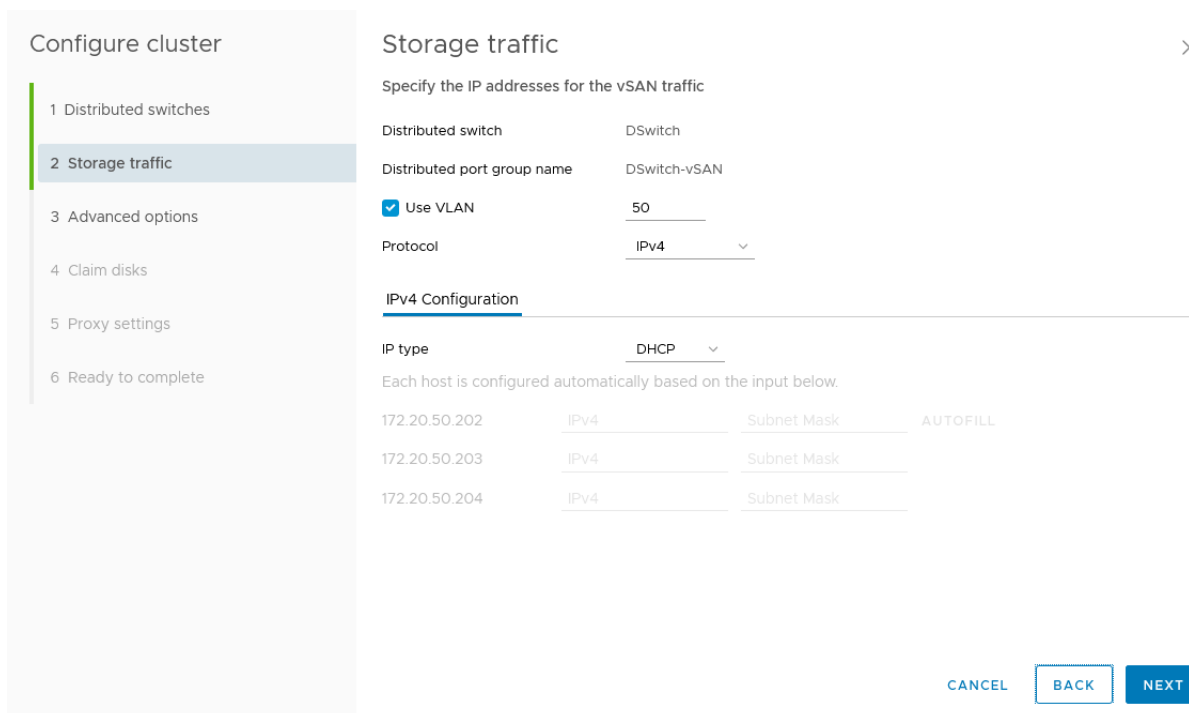
FIGURE 4. Configuring Storage Traffic

3. The **Advanced Options** screen allows customization of vSAN options, host options and has a toggle for enhanced vMotion compatibility. In vSAN options, select a deployment type (single site, stretched etc.) based on the structure of the deployment. Deduplication and compression are available on vSAN 6.2 onwards. These two space efficiency techniques are only available for all-flash configurations. They are enabled together at the cluster level and cannot be enabled individually. Deduplication occurs when data is de-staged from the caching tier to the capacity tier and is performed within each disk group. Redundant blocks across multiple disk groups will not be deduplicated. Compression is then applied after the deduplication has occurred but before the data is written to the capacity tier. For more information, see VMware's vSAN 6.2 Space Efficiency white paper.
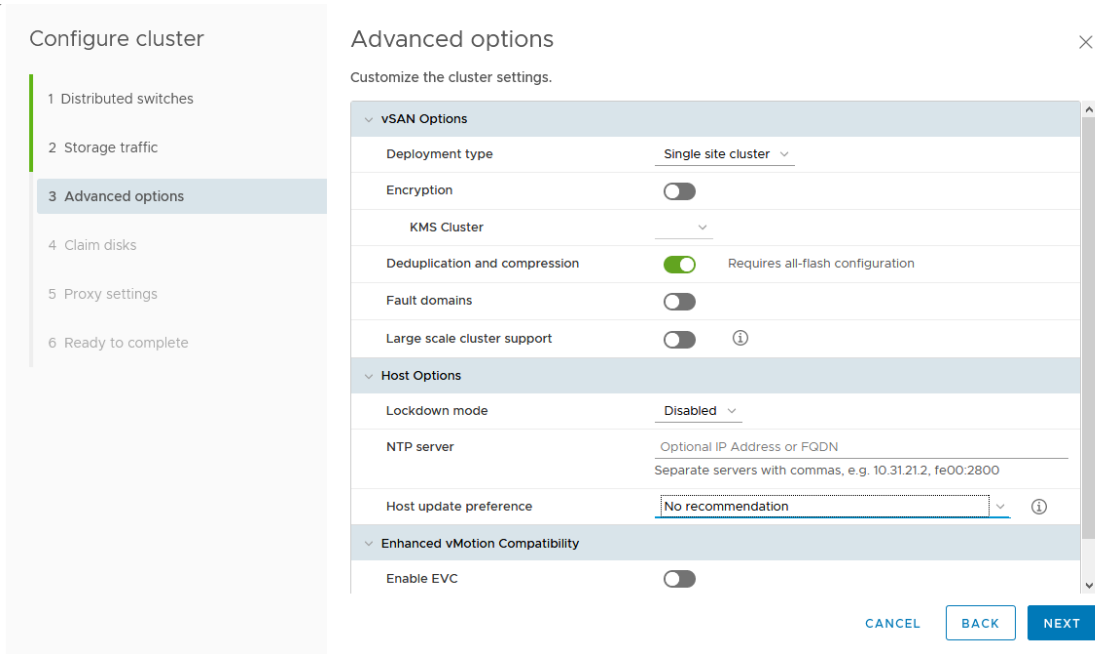
FIGURE 5. Configuring Advanced options

4. On the Claim Disks screen, claim the available disks as cache or capacity tier. To avoid warnings and errors, ensure that each host has a balanced configuration consisting of a cache tier and one or more capacity tier disks. If there are more than one disk models, choose the high speed, write intensive disks for cache tier. Once the disks are claimed correctly, a "Configuration correct" dialog box will be displayed at the bottom of the screen. Once satisfied with the configuration, click Next.
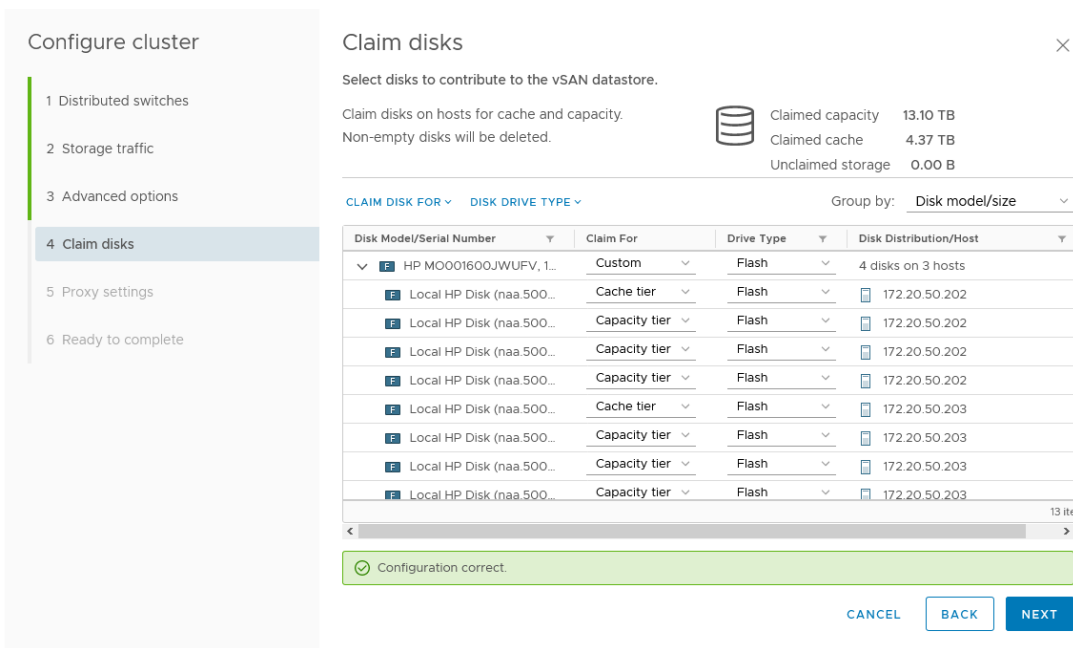
FIGURE 6. Claiming Disks

5. On the Proxy settings screen, enter requisite details pertaining to the proxy server used in the environment, if any and click Next.
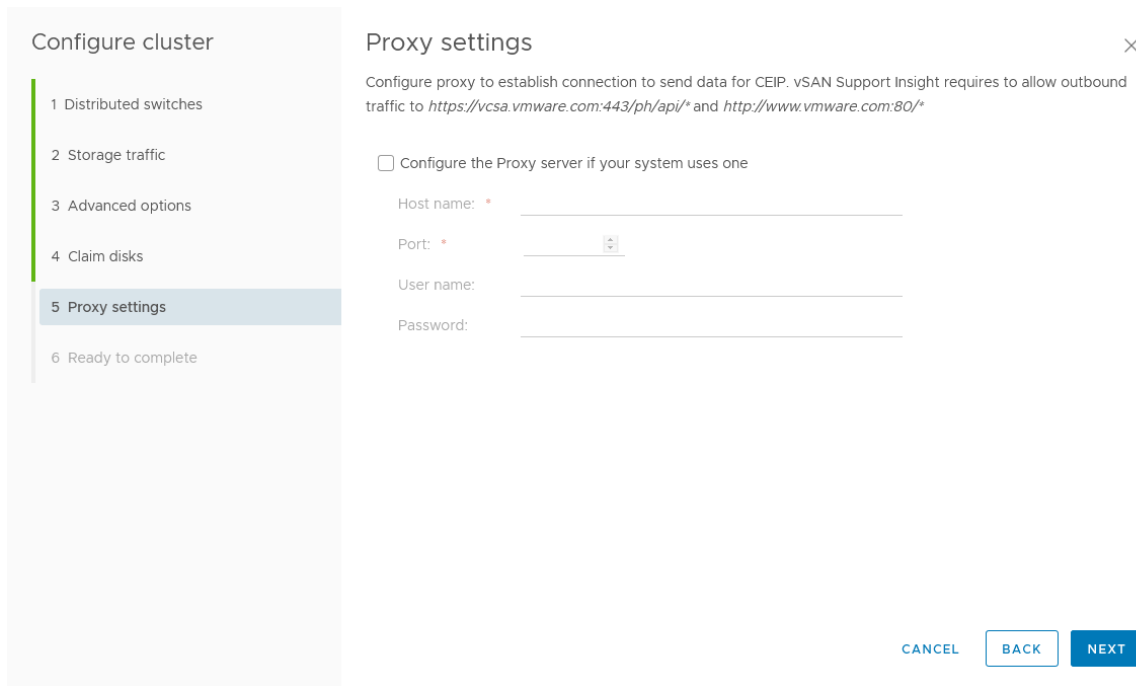
FIGURE 7. Proxy settings

6. On the ready to complete screen, review the specified options and click Finish to apply the specifies configuration settings.
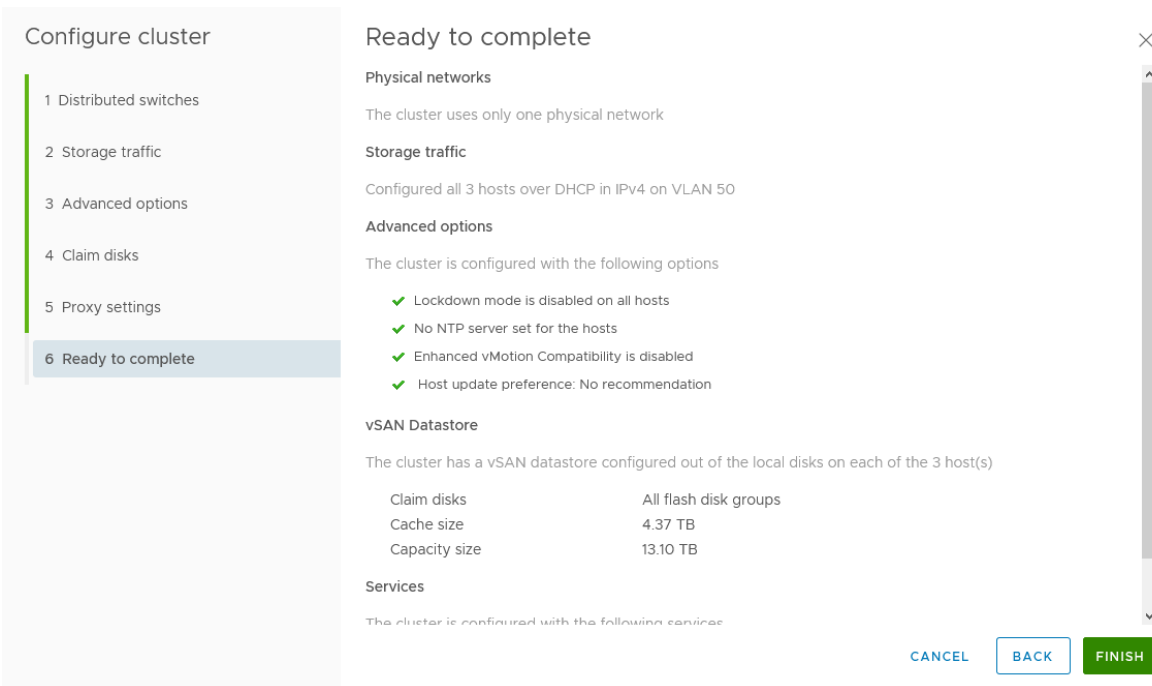

FIGURE 8. Cluster Configuration Review

STORAGE POLICIES

vSAN relies on storage policies to define Virtual Machine storage requirements. Define the storage policies based on VM performance and availability requirements. When you enable vSAN on a cluster, a default vSAN policy is created and applied to the aggregate datastore. There are several settings that can be defined within a policy including the number of failures to tolerate, which allows a virtual machine to survive a few hardware failures in a cluster.

To add a new storage policy, go to Menu > Policies and Profiles > Create VM Storage Policy. This will open a new Storage Policy Wizard.

1. On the Name and description screen, select an instance of vCenter Server and add a name and description for the storage policy.
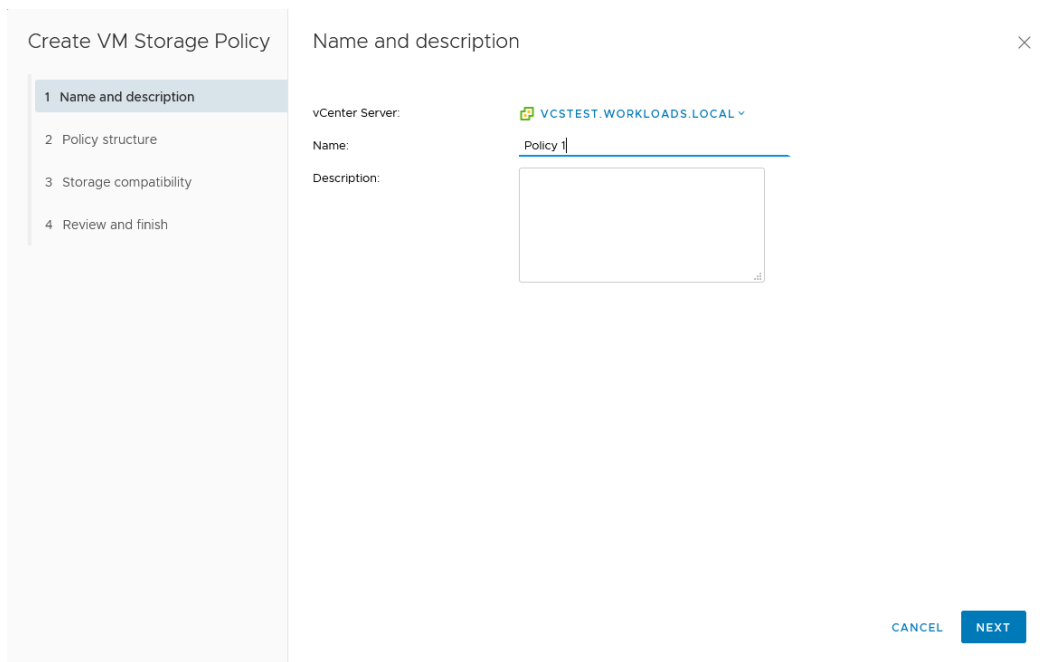
FIGURE 9. VM Storage Policy: Name and description

2. On the Policy structure screen, enable Host based rules and rules for vSAN storage.
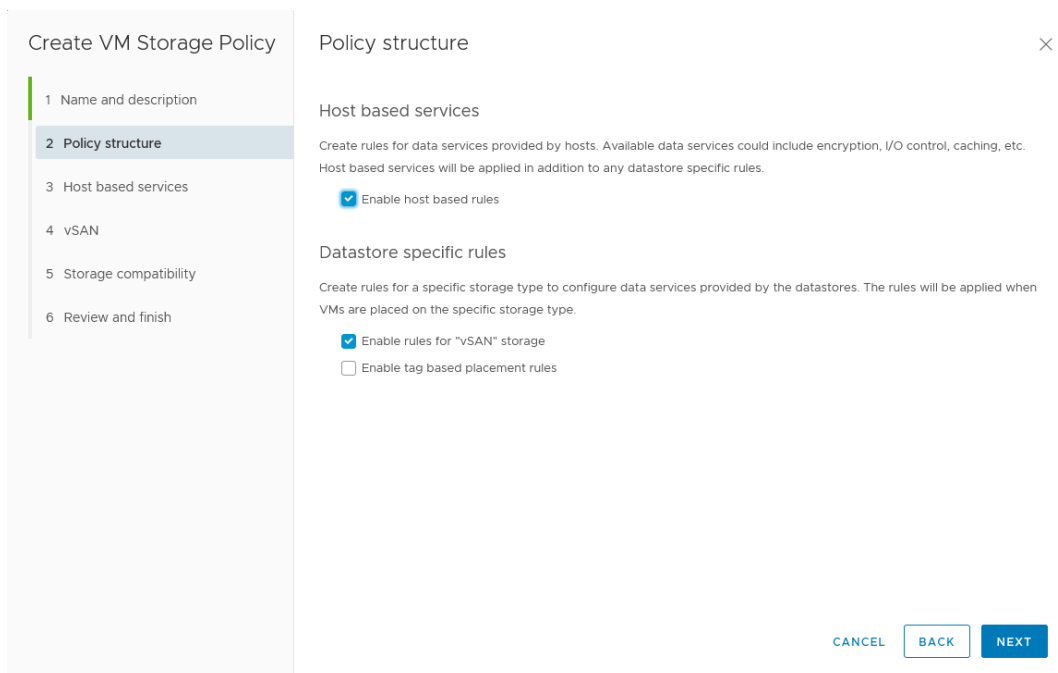
FIGURE 10. VM Storage Policy: Structure

3. On the Host based services screen, under Encryption, set use storage policy component to Default Encryption Properties.
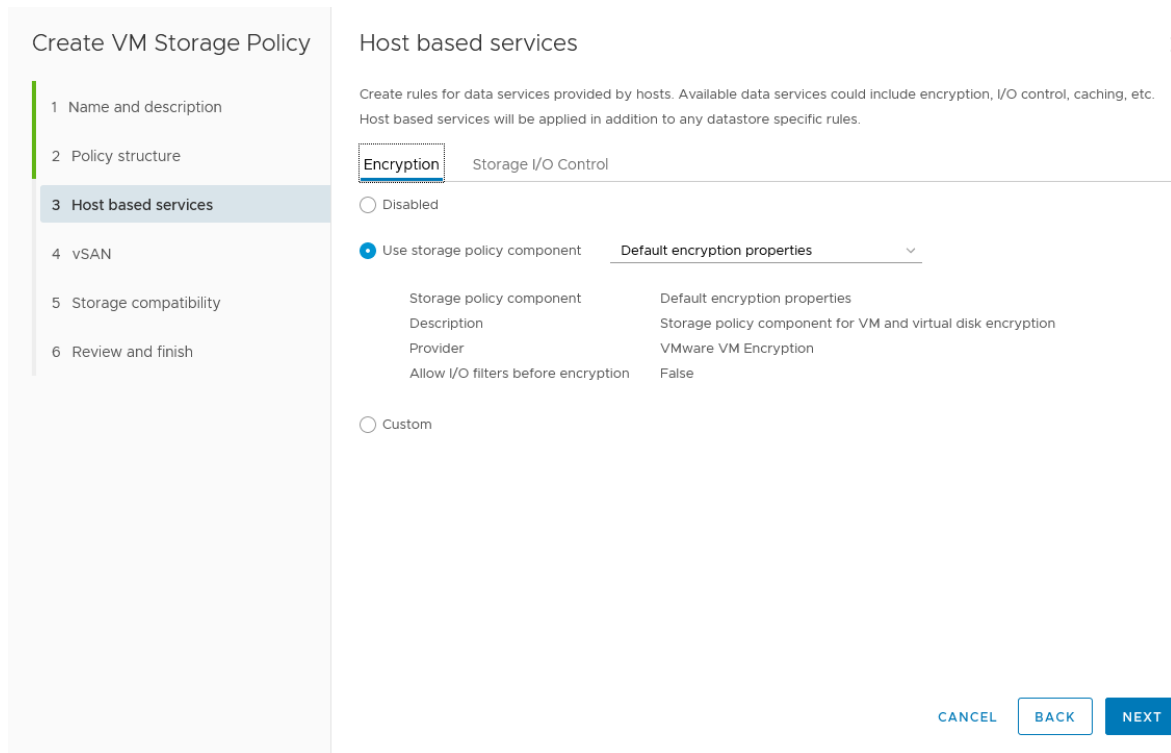
FIGURE 11. VM Storage Policy: Host based services

4. On the vSAN screen, the Availability tab offers options for Site disaster tolerance and Failures to tolerate. Set these according to requirements.
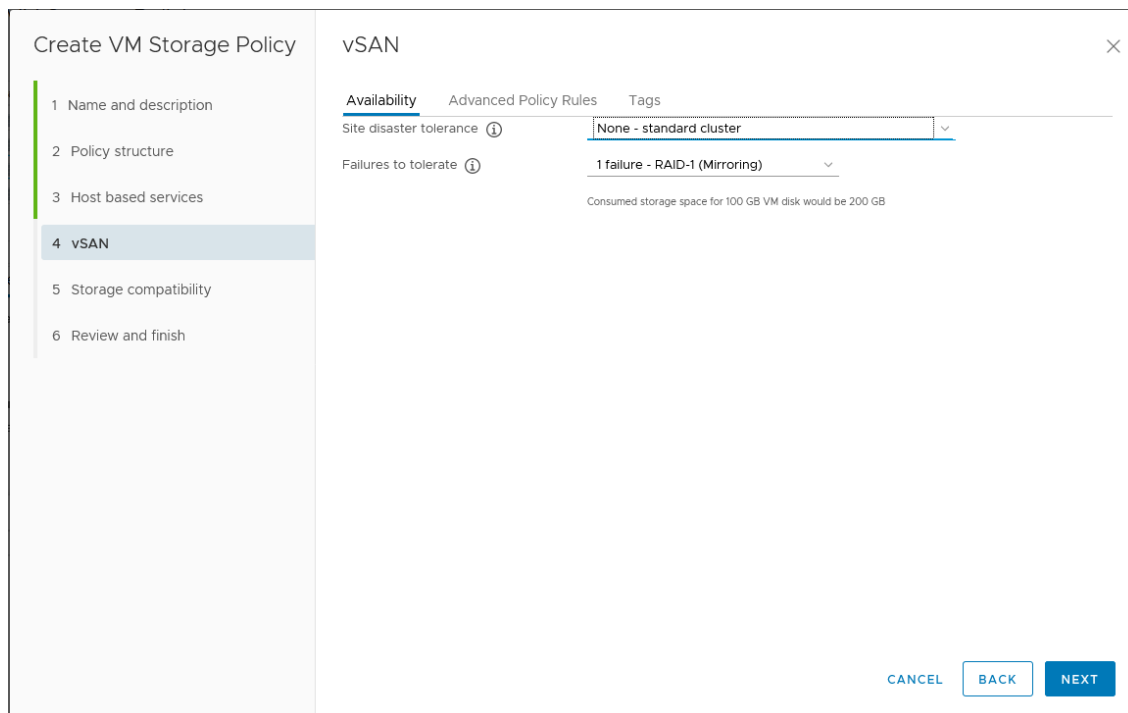

FIGURE 12. VM Storage Policy: vSAN Availability options

5. On the vSAN screen under the Advanced Policy Rules Tab provides configuration options for advanced settings like Number of disk stripes per object, IOPS limit etc. There are toggles to enable Force provisioning, which will allow the provisioning of virtual machines with a policy that cannot be met by the current Virtual SAN cluster resources and to disable object checksum. Configure these policies as required by the deployment.
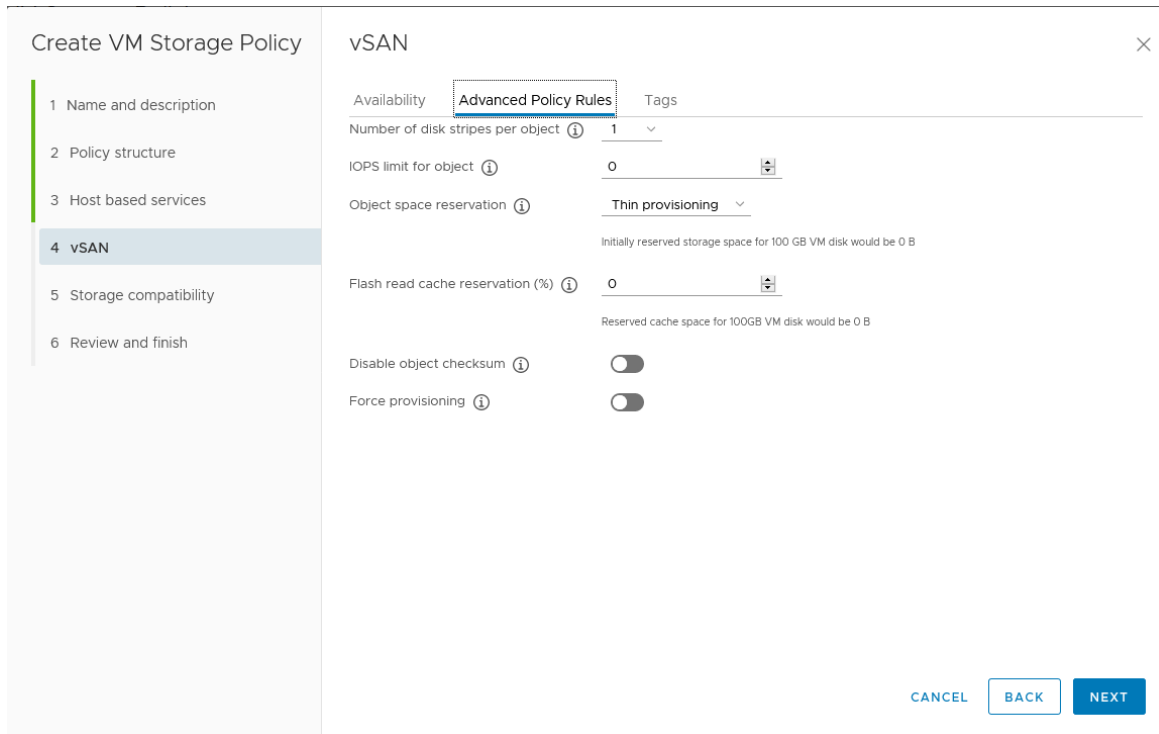
FIGURE 13. VM Storage Policy: vSAN Advanced policy rules

6. The storage compatibility screen will list all vSAN datastores that are compatible with the settings chosen in the new vSAN policy.
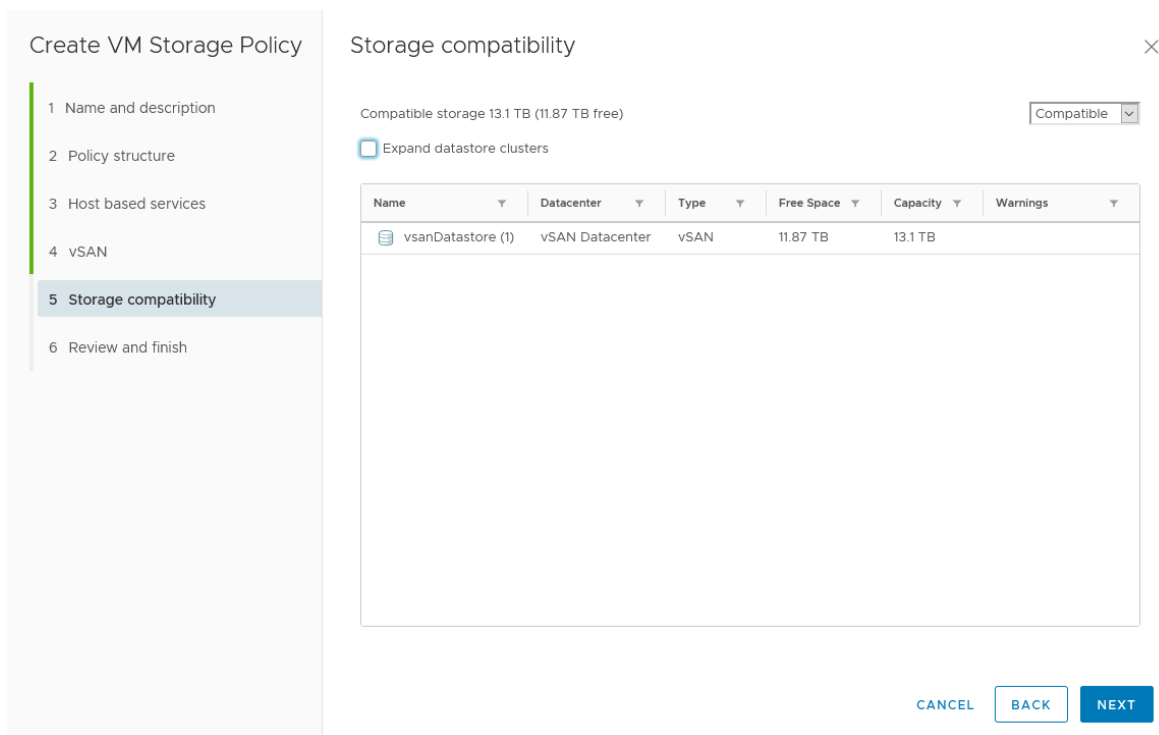
FIGURE 14. VM Storage Policy: Storage compatibility

7. Review the policy settings on the review and finish screen and click Finish to create the policy.
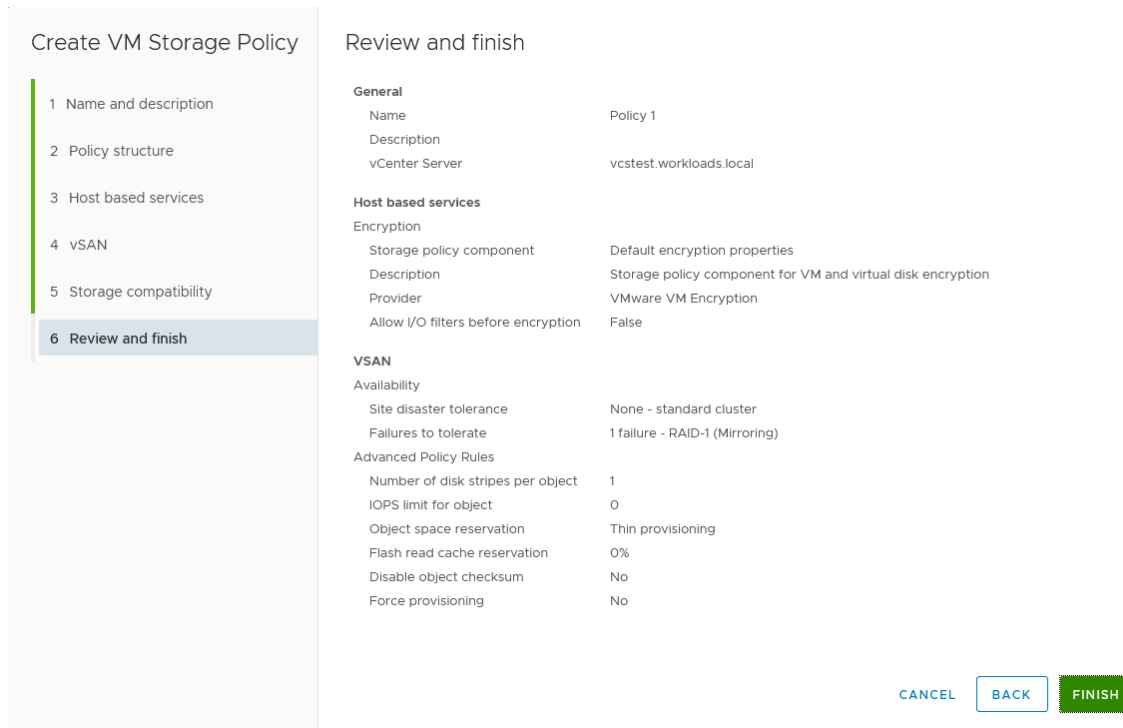

FIGURE 15. VM Storage Policy vSAN availability options

Once defined, policies can be applied when a VM is being created within the cluster. Alternatively, virtual machines can be assigned to the policy by editing their settings. If no policy is specified, the default policy will be applied with the following key parameters:

- Number of failures to tolerate = 1

- Number of disk stripes per object = 1

- Flash-read cache reservation = 0 percent

- Object space reservation = not used

- Force provisioning = disabled

Health Monitoring

VMware has introduced the ability to monitor the health of a vSAN cluster since vSAN 6.0. It runs several checks on the networking configuration, physical disks, and the vSAN cluster configuration.
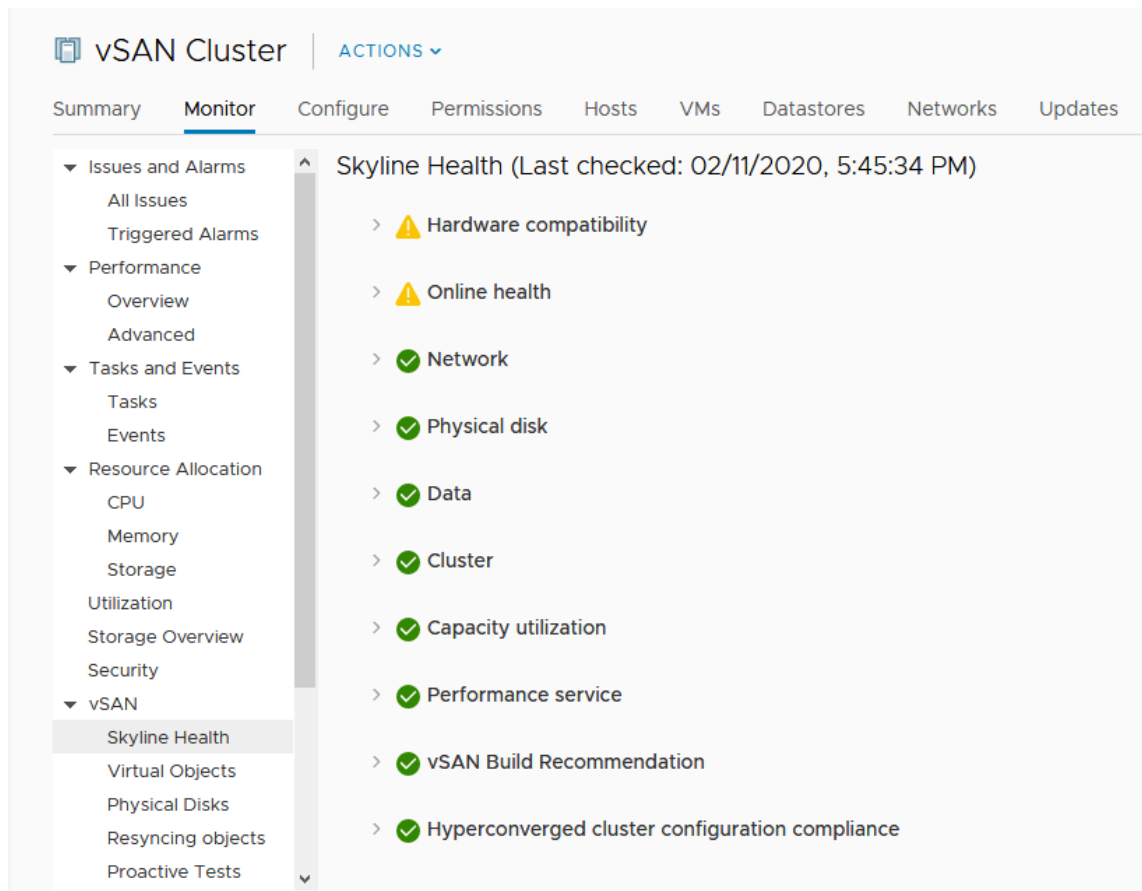

FIGURE 16. vSAN Cluster Health Monitoring

One of the most useful aspects of this feature is the ability to ensure the hardware configuration is supported. The hardware can be checked against the VMware Compatibility Guide for vSAN (either online or via uploading a copy of the VCG database). After the initial scan, the Health Monitoring service will continue monitoring the cluster on an ongoing basis and will highlight any issues found in the cluster.

ADDITIONAL RECOMMENDATIONS

- For optimal performance, enable Jumbo Frames. The benefits will need to be weighed against the added complexity involved is deploying it throughout the switching infrastructure and NICs.

- Ensure SSD or HDD are empty and not pre-formatted before adding them to the cluster.

- For better performance, create multiple smaller disk groups instead of larger disk groups. The increase in cache drives could boost performance.

- While disparate hardware configurations are supported, it is recommended to use similar (if not identical) hardware configurations (number of disks, capacity, types of disk, etc.).

- In an all flash-configuration, prefer RAID 5/6 for redundancy to save space and increase usable disk capacity instead of using RAID 1, unless RAID 1 is strictly required for the use case.

- Refer to the VMware Monitoring and Troubleshooting Guide for further help with troubleshooting.

RESOURCES AND ADDITIONAL LINKS

- HPE ProLiant Servers

- HPE Custom ESXi Images

- HPE ProLiant server and option firmware and driver support recipe for VMware

- HPE Service Pack for HPE ProLiant

- Deploying and updating VMware vSphere on HPE ProLiant servers

- VMware's technical papers portal

- VMware vSAN Planning and Deployment Guide

- vSAN ReadyNode Sizer

- VMware Compatibility Guide for vSAN

To help us improve our documents, provide feedback at hpe.com/contact/feedback

**Make the right purchase decision.
Contact our presales specialists.**

Chat   **Email**   **Call**

Share now

Get updates

**Hewlett Packard
Enterprise**