VMware Virtual SAN Proof of Concept Guide

March 2014 Edition

TECHNICAL MARKETING DOCUMENTATION

vmware®

Table of Contents

Introduction

Virtual SAN is a new software-defined storage solution that is fully integrated with vSphere. Virtual SAN aggregates locally attached disks in a vSphere cluster to create a storage solution that can be rapidly provisioned from VMware vCenter during virtual machine provisioning operations. It is an example of a hypervisor-converged platform, that is, a solution in which both storage and compute for virtual machines are combined into a single device, with the storage being provided right within the hypervisor itself (as opposed to via a storage virtual machine running alongside other virtual machines).

Virtual SAN is an object-based storage system designed to provide VM-centric storage services and capabilities through a Storage Policy Based Management (SPBM) platform. SPBM and VM Storage Policies are solutions designed to simplify virtual machine storage placement decisions for vSphere administrators.

Virtual SAN is fully integrated with the core vSphere enterprise features such as VMware vSphere High Availability (vSphere HA), VMware vSphere Distributed Resource Scheduler (vSphere DRS) and VMware vSphere vMotion®. Its goal is to provide both high availability and scale-out storage functionality. It can also be considered in the context of quality of service (QoS) because VM storage policies can be created to define the level of performance and availability required on a per virtual machine basis.

This document guides the reader through a Proof of Concept of Virtual SAN. It provides information on how to properly set up a test environment, and steps through a number of test scenarios. By completing these exercises, a customer will experience a range of capabilities that Virtual SAN has to offer.

Requirements

In order to ensure success of a Virtual SAN POC, it is critical that the hardware is supported by VSAN as per the VMware Compatibility Guide and that it is configured in the right manner. Appendix A of this document contains a checklist of pre-requisites that must be met before proceeding with the exercises.

The VMware Compatibility Guide for VSAN may be found at the following URL:
http://www.vmware.com/resources/compatibility/search.php?deviceCategory=vsan

Proof Of Concept Lab Setup

The following is a description of the lab setup that was used for creating this guide. The reader should reference this information, including names of objects such as hosts and virtual machines, when reading the subsequent sections.

Hosts

The POC requires a total of 4 ESXi hosts, all running vSphere 5.5. The hosts contain empty local disks (both HDD & SSD) that can be consumed by Virtual SAN. There is one empty SSD and two empty HDDs per host.

NOTE: This lab includes a fifth host that does not have any local storage. You may choose not to include a host without local storage for your tests. In this case, you may proceed with the exercises in this document by simply working with the 3-node cluster as a base, instead of a 4-node cluster.

Cluster

Three of the hosts are already in a cluster called Cluster Site A.

A fourth host is a stand-alone host that is not part of the cluster.

Datastores

In this guide, one NFS datastore, called **ds-site-a-nfs01**, is present and available to all hosts in the POC.

Virtual Machines

A single VM called **base-sles** resides on this NFS datastore. This is not necessary for a successful POC, but having a VM template available in your environment will speed up the POC. Optionally, you may have to deploy a Virtual Machine from an ISO image to complete the exercises.

Networking

In this setup, there is a distributed switch configured. A distributed switch is not necessary to implement a successful POC; standard virtual switches are fully supported in a Virtual SAN configuration. However a distributed switch configuration allows us to use Network I/O Control to provide Quality Of Service (QoS) on the Virtual SAN traffic.

Exercises

This guide will take customers through a number of workflows, such as configuring the Virtual SAN cluster, the creation of various VM Storage Policies and the deployment of VM with those policies. Exercises are to be performed in the following sequence:

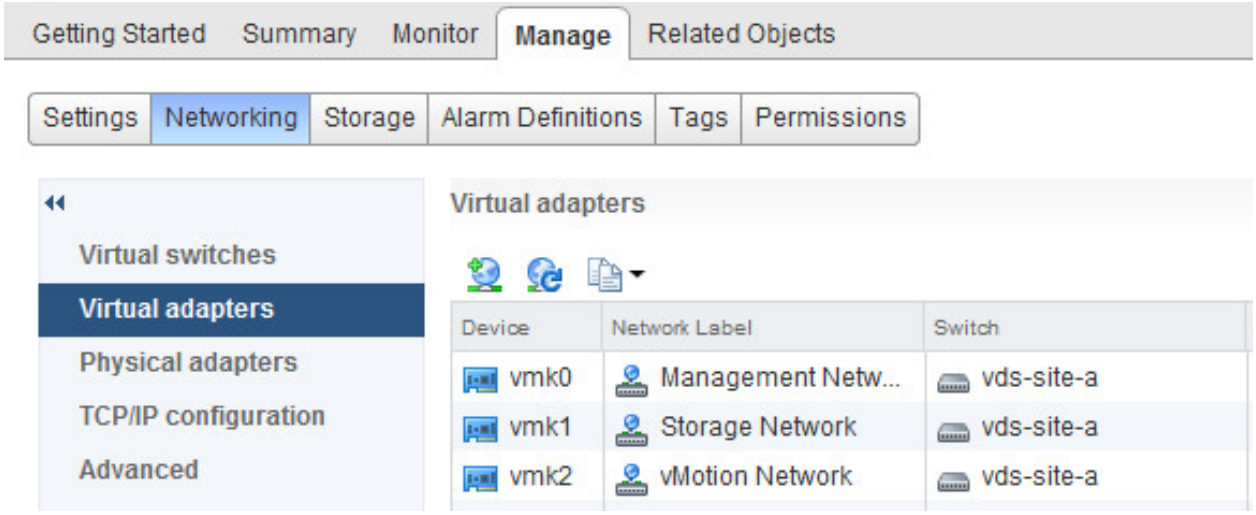
- Exercise 1 – Setup the Virtual SAN Network
- Exercise 2 – Enable Virtual SAN on the Cluster
- Exercise 3 – Scale-Out Storage
- Exercise 4 – Build VM Storage Policies
- Exercise 5 – vMotion & Storage vMotion
- Exercise 6 - vSphere HA & Virtual SAN Interoperability

Exercise 1 – Setup the Virtual SAN Network

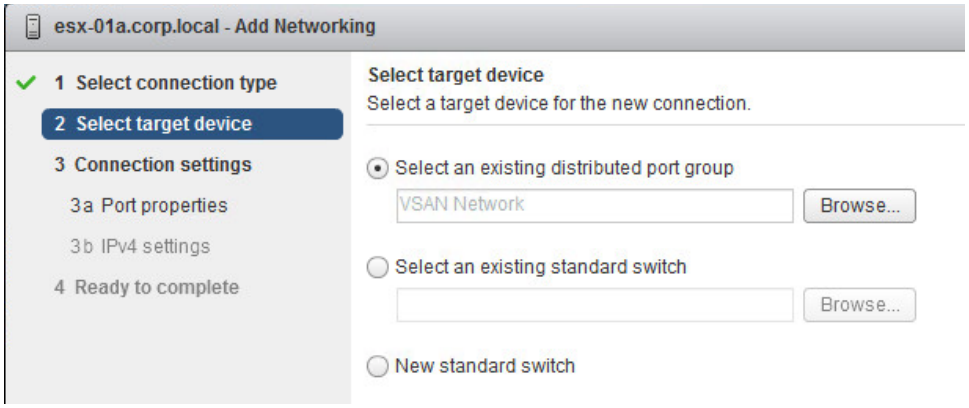
The first step is to set up communication between each of the ESXi hosts in the cluster. In this POC, there are already 3 VMkernel ports configured on each ESXi host. These are:

1. Management Network
2. vMotion Network
3. Storage Network

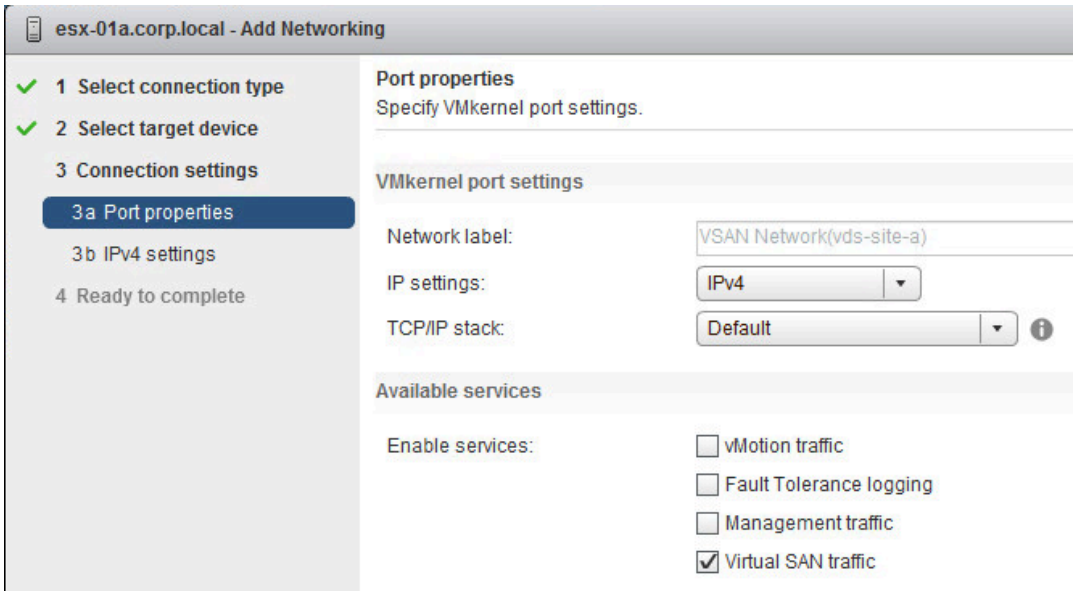
Navigate from the Home view to vCenter > Hosts & Clusters. Select the first of the four ESXi hosts and then select the Manage tab. Select **Networking** and then the **Virtual Adapters** view. It should look identical to the following:



We must now add the Virtual SAN network. Setting up the Virtual SAN network involves steps identical to setting up any other VMkernel network. Click on the icon to add a new virtual adapter (VMkernel Network Adapter), then select a distributed port group called VSAN network that has already been created:



At the point where you select the Port properties, select **Virtual SAN traffic**.



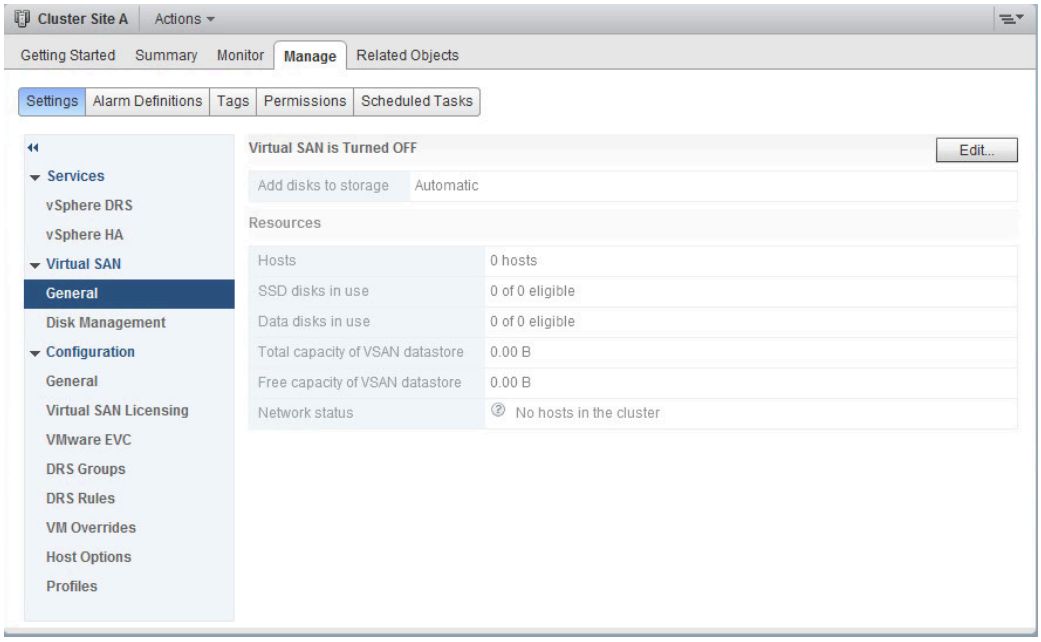
All other settings may be left at the **default** (Obtain IPv4 settings automatically). The Virtual SAN interfaces will be assigned their IP addresses via DHCP. This must be repeated for the remaining ESXi hosts in the cluster. We will deal with the standalone ESXi host separately in part 6 of the lab.

***Key Takeaway**: Virtual SAN requires a network connection between hosts for communication and I/O purposes. Virtual SAN requires multicast, but **only** in L2.*

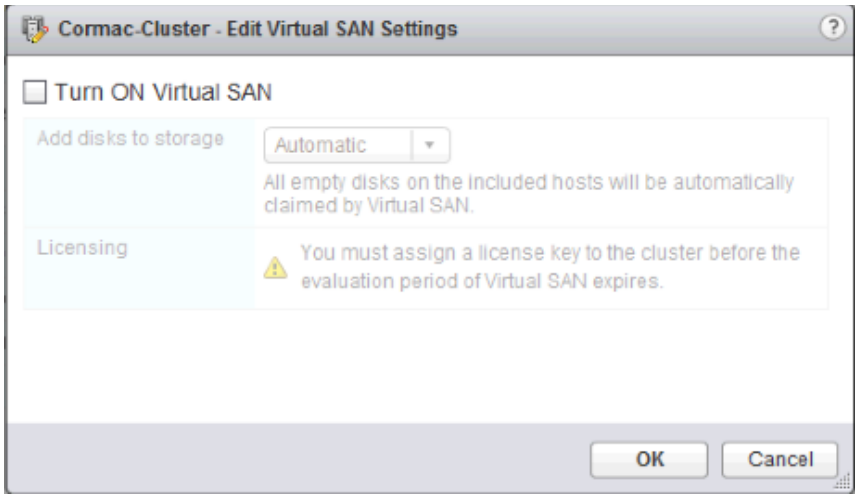
Exercise 2 – Enable Virtual SAN on the Cluster

At this point, the networking is in place on all nodes. The ESXi hosts are already in a cluster. We can now go ahead and enable the Virtual SAN cluster, and point out how easy it is to setup.

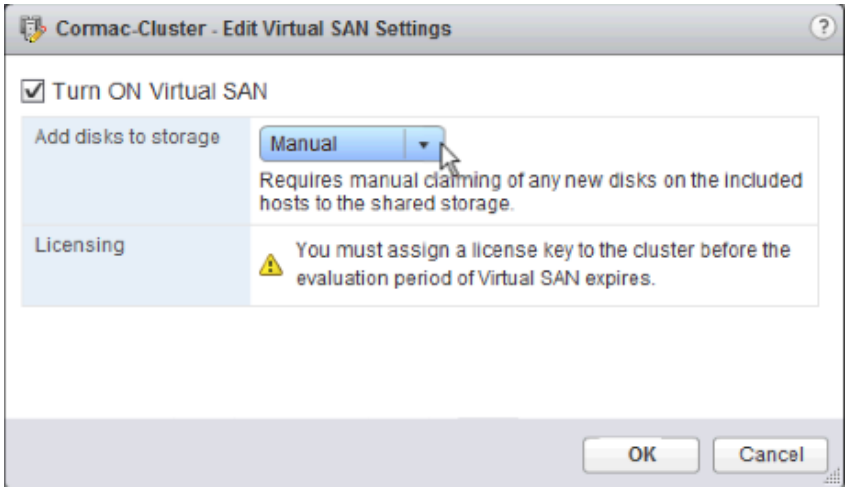
Select the cluster object, then navigate to Manage tab > Settings > Virtual SAN > General:



Next then click on the **Edit** button located to the right of the window – **Do not click OK until told to do so**:

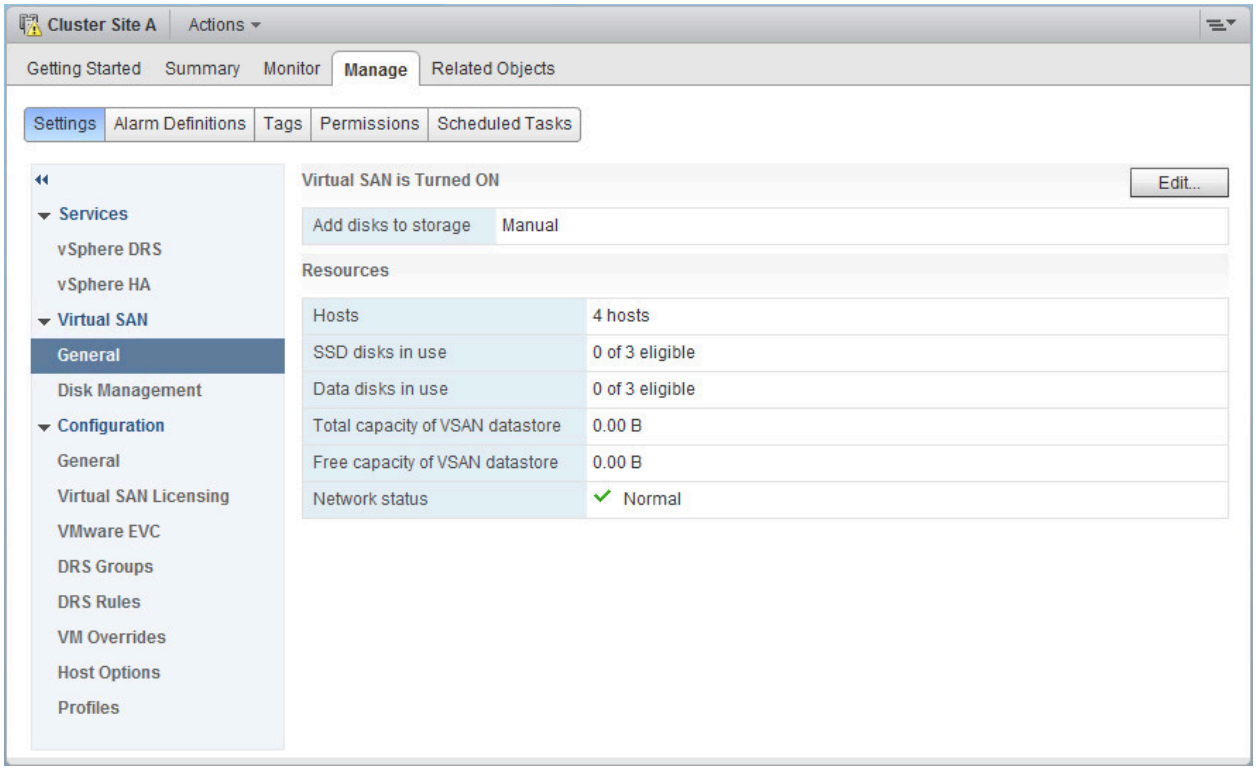


By default, Virtual SAN will automatically **add new disks** to the Virtual SAN cluster (Disk claim policy). Ensure that the setting 'Add disks to storage' is set to **manual**, as we manually add disks to the cluster.

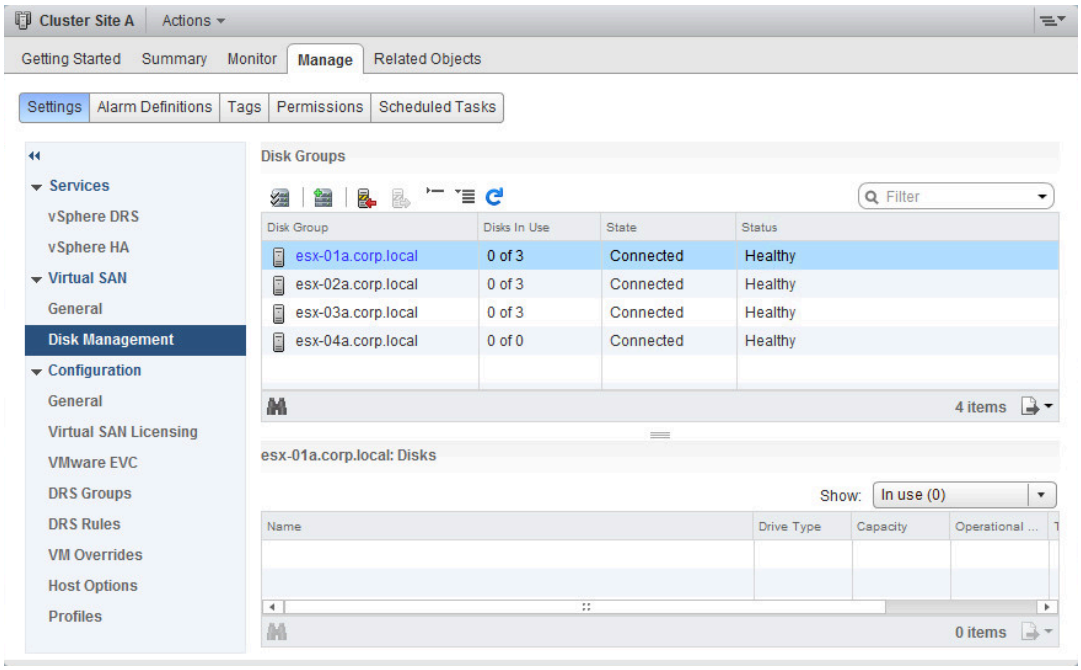


Remember that there is a requirement to have at least 1 SSD drive per host. In this lab, there is a single SSD & a two HDD on three of the four ESXi hosts in the cluster. Click **OK**.

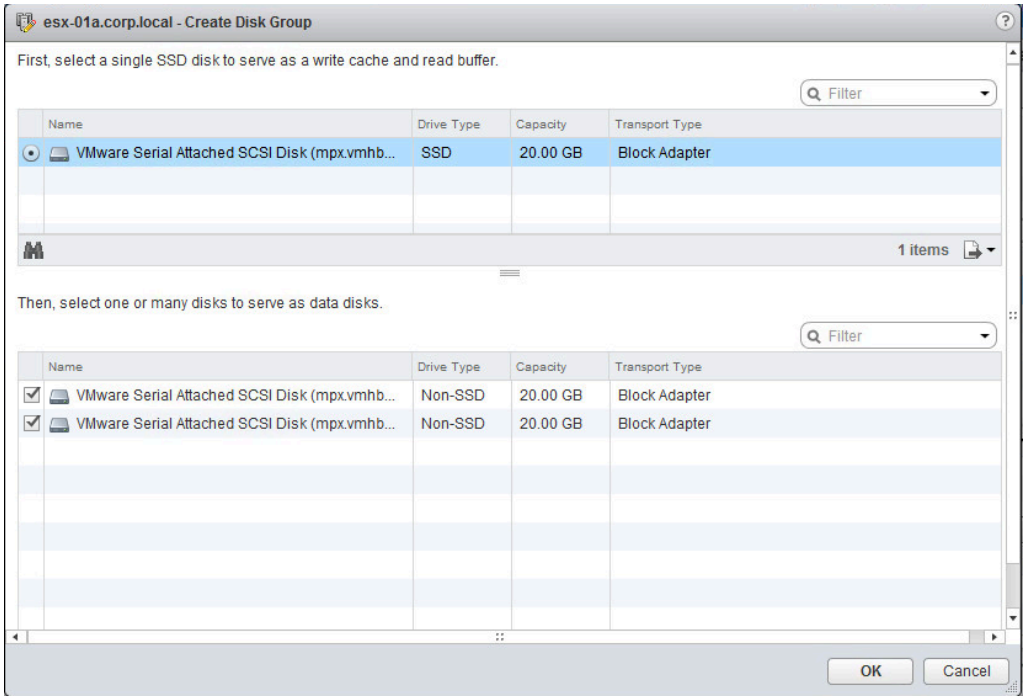
After clicking OK, you may have to refresh the vSphere web client for these changes to appear.



Go to the Virtual SAN **Disk Management** view. Here, you will see that there are 4 hosts in the cluster, but no Disk Groups are created. This is because **Manual** mode was selected when the cluster was created, so you have to create the Disk Groups manually.



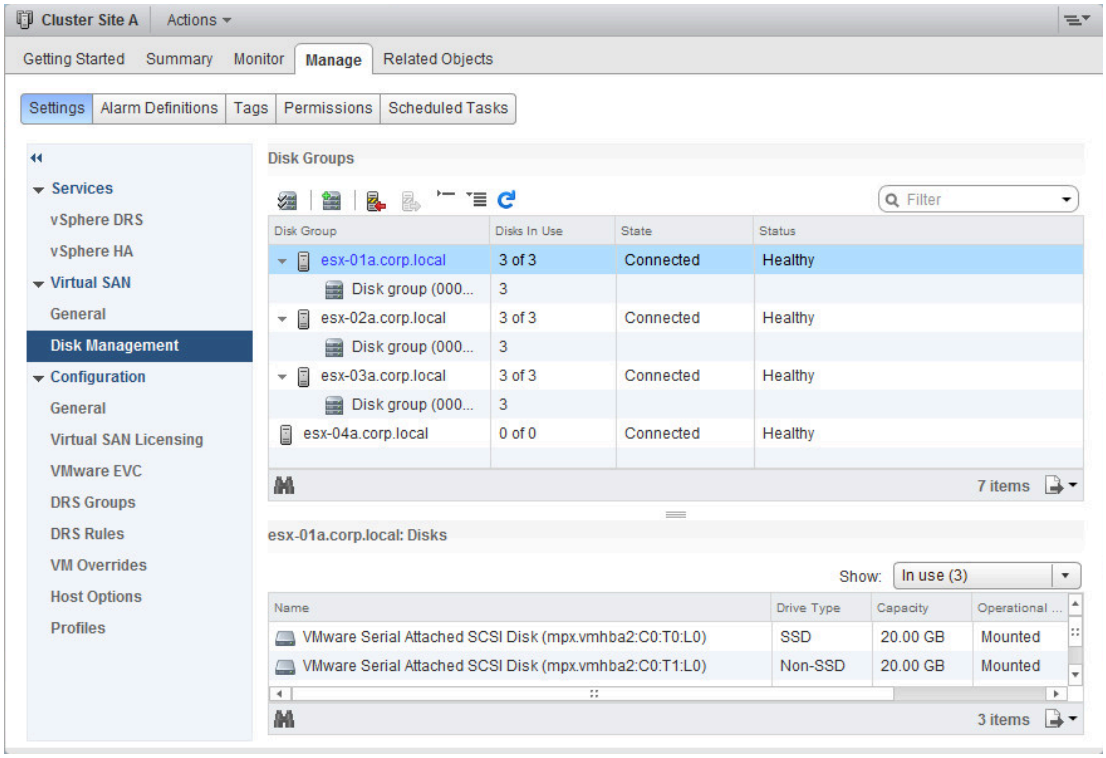
Select host esx-01a and click on the icon (with the green plus symbol) to create a new Disk Group.



Each disk group created may contain only one SSD. The SSD is used for write cache/read buffer and the HDDs are used for persistence / capacity. The SSD & both the HDDs on these hosts have been selected to be part of the Virtual SAN. Repeat this operation for all three hosts that have storage (esx-01a, esx-02a &

esx-03a).

When completed, this cluster will have three disk groups.

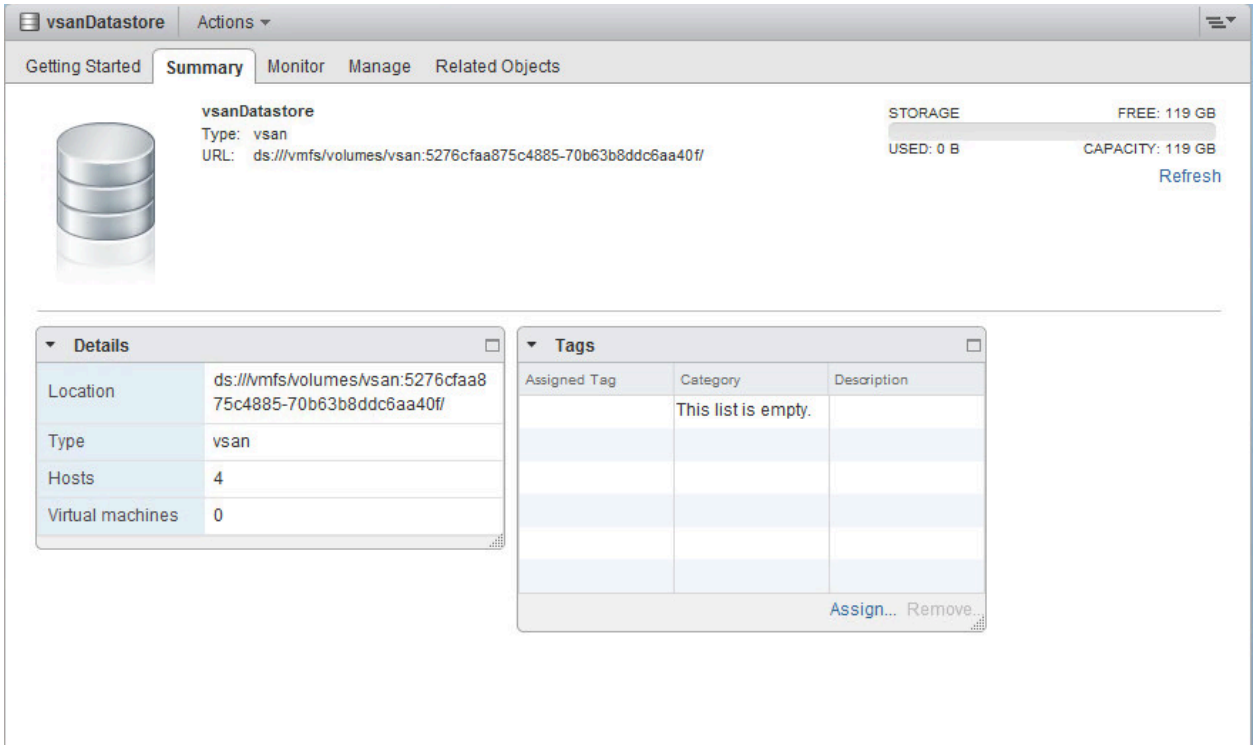


At this point, all four ESXi hosts in the cluster should be able to successfully see the Virtual SAN datastore, labeled **vsanDatastore**.

***Key Takeaway**: The steps to setup Virtual SAN are very simple, and are akin to the steps to setup vSphere HA and DRS clusters.*

2.1 Verifying the Virtual SAN datastore

At this point, we will verify that the Virtual SAN datastore has been successfully created and that its capacity correctly reflects the total local storage capacity from each of the ESXi hosts. Once the network is created, physical storage is added to disk groups and the Virtual SAN cluster created, a single Virtual SAN datastore is built. Navigate to the Storage view and check the status of the vsanDatastore.



The capacity is an aggregate of the HDDs taken from each of the 3 ESXi hosts in the cluster. That is 3 x 2 x 20GB = 120GB (less some Virtual SAN Datastore overheads). The 3 x SSD (1 on each of the 3 ESXi hosts with storage) are not considered when the capacity calculation is made.

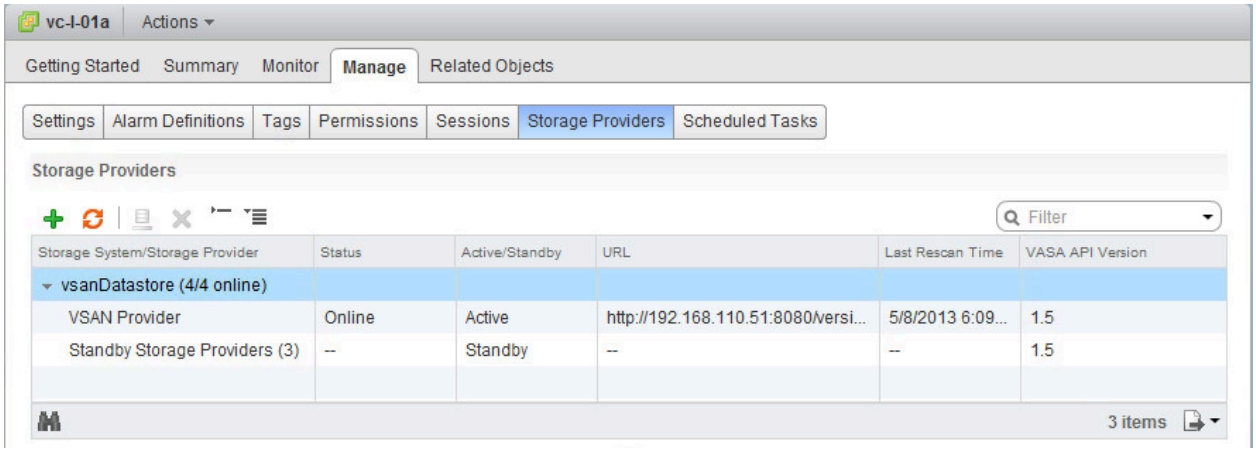
***Key Takeaway**: Virtual SAN uses the SSDs for read cache and write buffering. HDDs are used for capacity.*

2.2 Verifying the Storage Provider status

To learn about the capabilities of Virtual SAN, and to communicate between vCenter and the storage layer, a storage provider needs to be configured. Each ESXi host has a storage provider.

When a Virtual SAN cluster is formed, the storage providers will be registered automatically with SMS, the Storage Management Service (by vCenter). However, it is best to verify that the storage providers on one of the ESXi hosts has successfully registered and is active, and that the other storage providers from the remaining ESXi hosts in the cluster are registered and are in standby mode.

Navigate to the vCenter server > Manage tab > Storage Providers to check the status.



In this four-node cluster, one of the Virtual SAN providers is online and active, while the other three are in Standby. Each ESXi host participating in the Virtual SAN cluster will have a provider, but only one needs to be active to provide Virtual SAN datastore capability information.

Should the active provider fail for some reason, one of the standby storage providers will take over.

***Key Takeaway**: The storage provider, which surfaces up the Virtual SAN capabilities to vCenter, is highly available.*

2.3 Troubleshooting Installation & Configuration Issues

A number of issues may occur during the installation and configuration of Virtual SAN. This section will look at some of the common issues encountered.

- **Automatic mode does not claim disks**
  - **Issue**: Disks must be deemed local to be automatically claimed. If ESXi determines that a disk could be shared, it does not report it as "local". Many SAS controllers could (in theory) allow disks to be shared. Virtual SAN currently does not support shared disks and reports these disks as not local.
  - **Solution**: Manually create disk groups with disks that are not “local”.
- **Disk group creation fails silently**
  - **Issue**: You manually build a disk group, select SSD & HDDs, and although the operation reports success, the disk group is not created.
  - **Solution**: This is the behavior observed in beta when Virtual SAN is not licensed. License Virtual SAN.
- **Network status reports ‘Misconfiguration detected’**
  - **Issue**: Virtual SAN cluster fails to form due to hosts in the cluster being unable to communicate and the fact that the VSAN Traffic port group is configured on each of the hosts and the hosts can successfully ping each other.
  - **Solution**: Ensure ports on network switch allow multicast, ideally using a multicast router. Multicast is only used for relatively infrequent metadata operations, like object creation, change in object status (e.g. after a failure) as well as publication of stats like significant change of free disk space. The publication of stats is throttled to only update on significant changes and so is also infrequent. In steady state multicast traffic will be very low, but it may spike in events of network hiccups or disk failures. Those spikes, while higher than

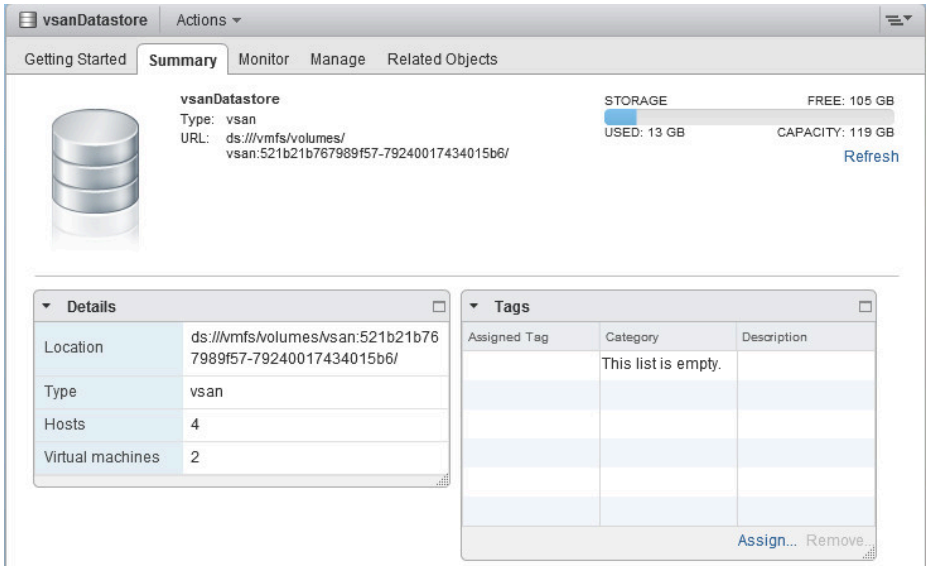
steady state, are still much lower than the TCP traffic from sustained storage IO.

Exercise 3 — Scale-Out Storage

In this section of the lab document, we are going to look at the ability to add another ESXi host with storage to the Virtual SAN cluster and observe the scale-out capabilities of the product.

At this point, we have four ESXi hosts in the cluster, although only three are contributing local storage to the Virtual SAN datastore. The VM, **base-sles**, should currently reside on a host that contributes local storage, esx-01a, esx-02a or esx-03a.

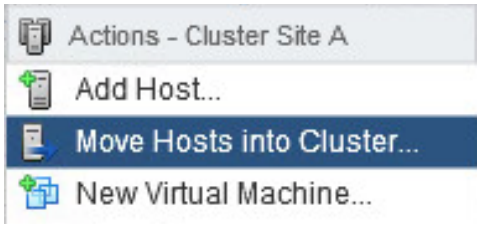
Let's check the status of the vsanDatastore. Navigate to the vsanDatastore summary tab, and it should look like this:



You have now reached the conclusion that you would like to add more compute and storage to the Virtual SAN cluster, which involves adding a new ESXi host that contains additional disks.

There is a fifth ESXi host (esxi-05a) in your inventory that has not yet been added to the cluster. We will do that now and examine how the vsanDatastore seamlessly grows to include this new capacity.

Navigate to the cluster object in the inventory, right click and select the action 'Move hosts into cluster'.

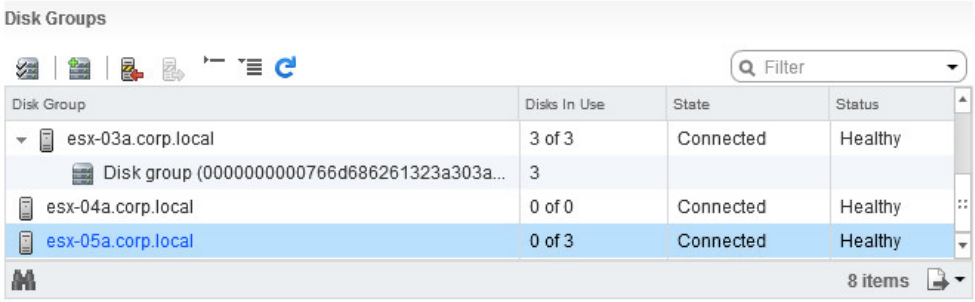


From the list of available hosts (you should only see esx-05a), select the host and click OK.

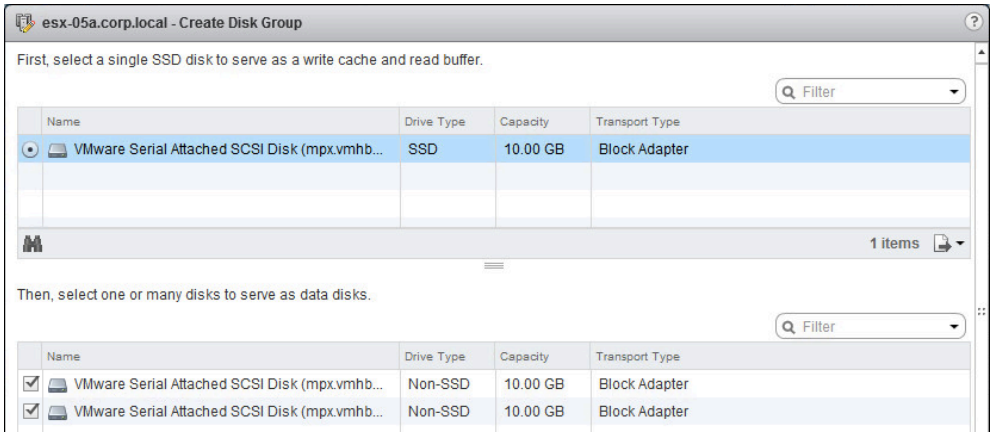
The next step is to add a Virtual SAN network to this host. As per the procedure outlined in Exercise 1 of this lab guide, create a Virtual SAN VMkernel network adapter to this host using the distributed port group called VSAN Network. Make sure you select the Virtual SAN traffic service, and let DHCP provide the IP settings.

3.1 Create New Disk Group

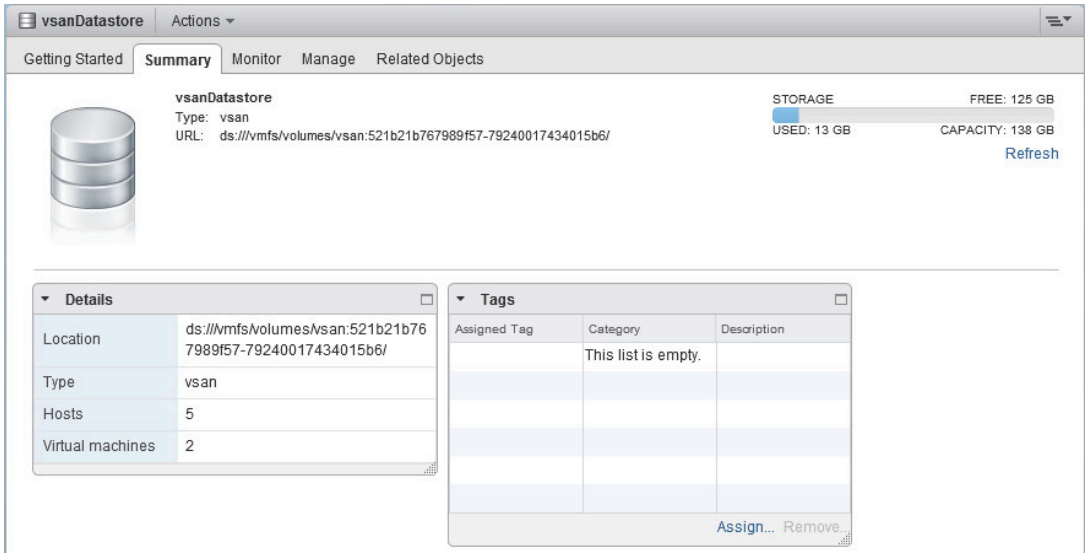
When that is completed, since the cluster was setup in Manual mode, we need to create a new disk group using the disks on host esx-05a. Select Cluster > Manage > Settings > Disk Management. Select the host just added. You will notice that 0 of 3 disks are in use.



Now create a new disk group, and add the disks (one SSD and two HDDs) to the disk group. These disks are all 10GB in size. Click OK and wait for the new disk group to be created.



When it is created, revisit the vsanDatastore summary view and check if the size has increased with the addition of the new host & disks. You should observe that the capacity of the datastore has seamlessly increased from 118GB to 138GB with the addition of two x 10GB HDDs. (remember that SSDs do not contribute towards capacity).



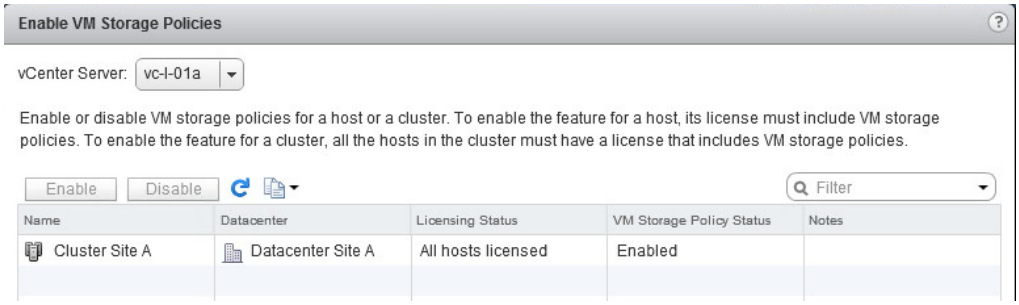
As you can clearly see, increasing storage and compute in Virtual SAN is relatively simple. Note that if Virtual SAN were setup in automatic mode, the steps to create a disk group would not be necessary.

***Key Takeaway**: Scaling out storage and compute in a Virtual SAN cluster is as simple as adding a new host to the cluster.*

Exercise 4 – Build VM Storage Policies

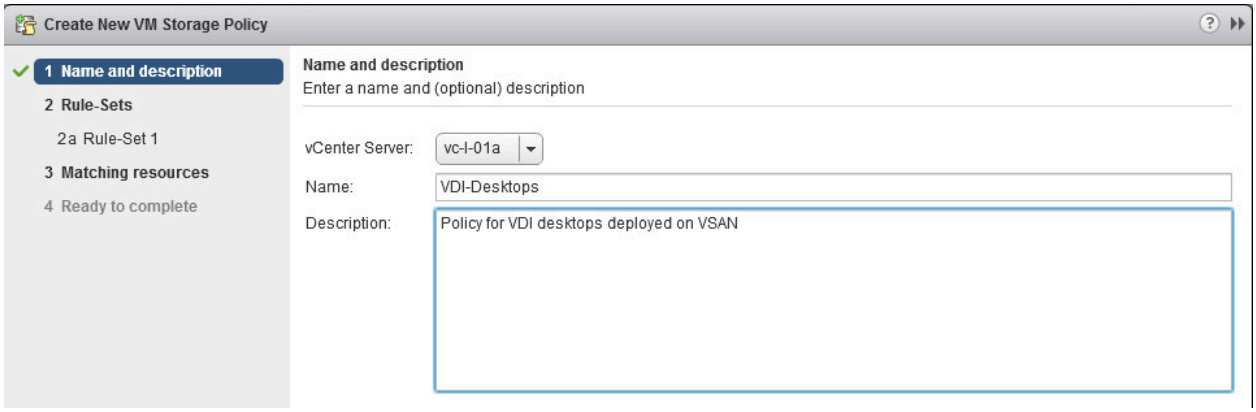
Once the storage provider is added, **capabilities** made available from Virtual SAN will be visible in the VM Storage Policies which can be found in Home > Rules & Profiles. VM Storage Policies are similar in some respects to the vSphere 5.0 & 5.1 Profile Driven Storage feature. There are two icons in this view representing 'Create VM Storage Policies' & 'Enable VM Storage Policies' respectively. The first step is to enable VM Storage Policies. *It is envisioned that Storage Policies will be automatically enabled on the Virtual SAN cluster in the final product, but for now it must still be enabled manually.*

Click on the icon with the check mark to 'Enable VM Storage Policies per compute resource', and then click **Enable**:

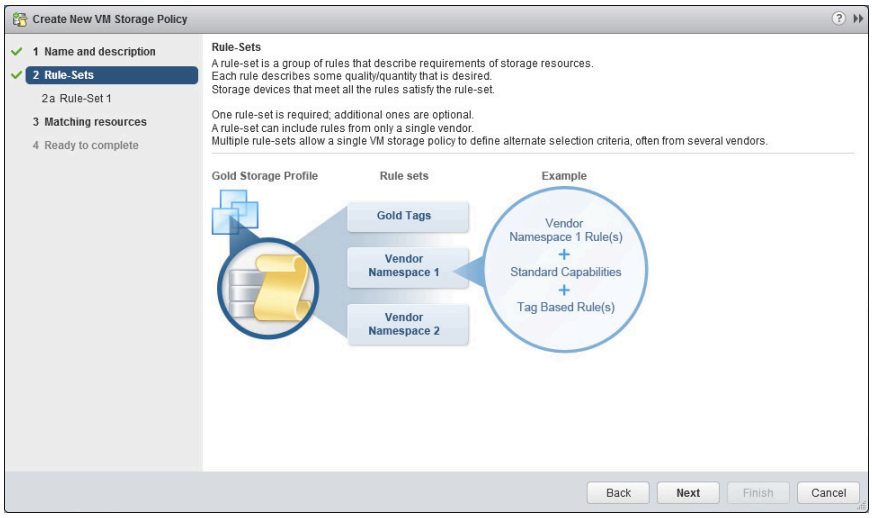


Once enabled, you may now close the window. The capabilities of the **vsanDatastore** should now be visible during VM Storage Policy creation. By using a subset of the **capabilities**, a vSphere admin will be able to create a storage policy for their VM to guarantee Quality of Service (QoS). Click on the icon with the plus sign representing 'Create New VM Storage Policy' to begin.

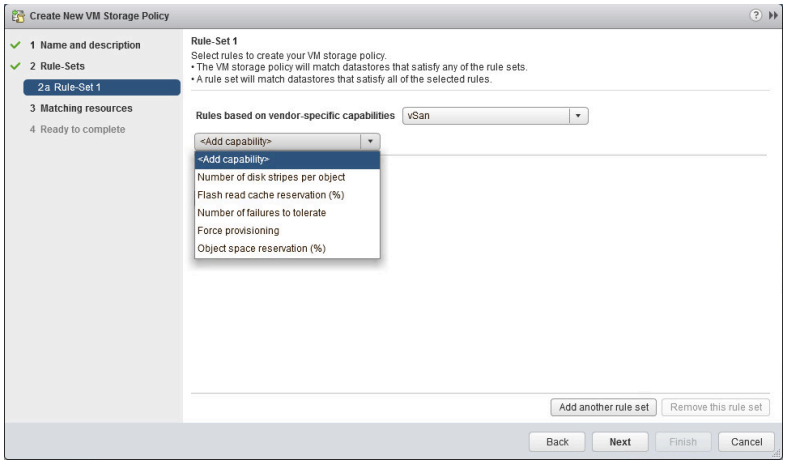
The first step is to give the VM Storage Policy a name. I will call it my **VDI-Desktops** profile for the purposes of this example:



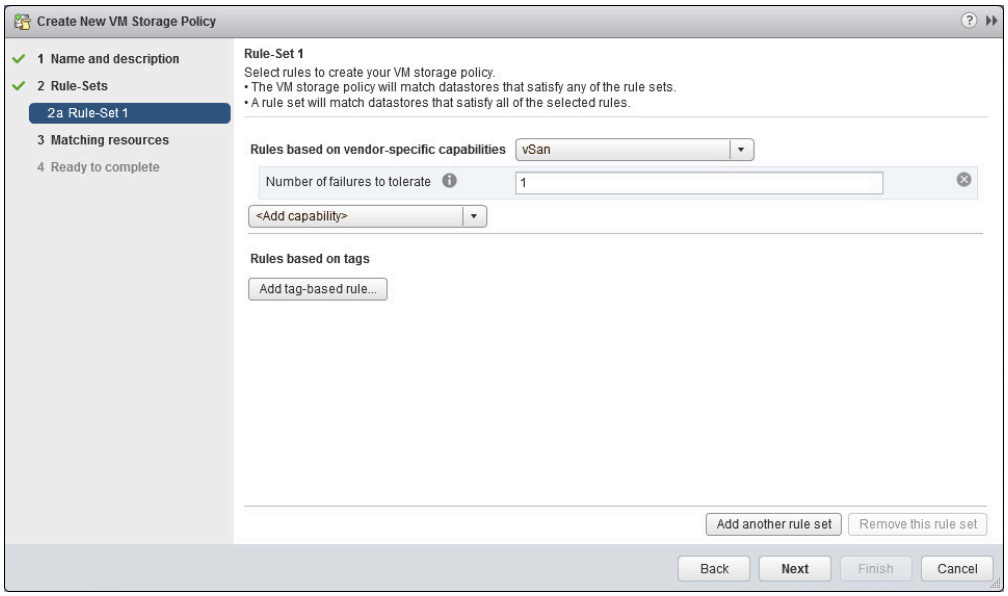
Next we get a description of rule sets. Rule-sets are a way of using storage from different vendors, e.g. for example you can have a single “bronze” policy which contains two separate rule-sets, one of which is a Virtual SAN Rule-Set and the other may be a 3rd party storage vendor Rule-Set. When “bronze” is chosen as the VM Storage Policy at VM deployment time, both Virtual SAN and the 3rd party storage are checked to see if they match the requirements in the policy.



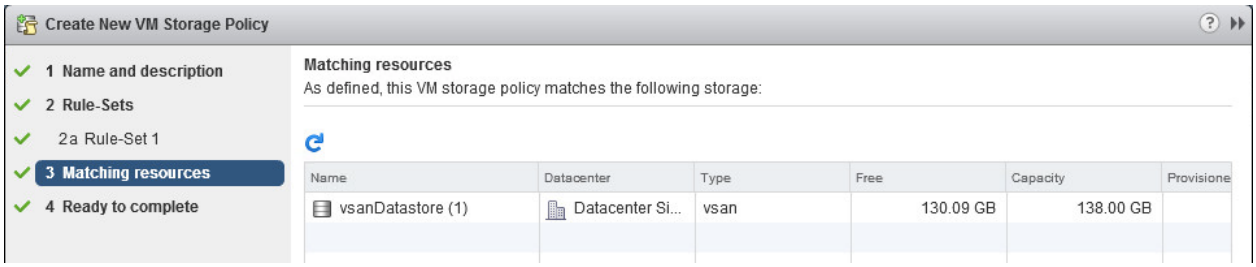
The next step is to select a subset of all of the **vsanDatastore** capabilities. Refer to official documentation for a full description of the capabilities. To begin you need to select the vendor, in this case it is called **vSan**.



The next step is to add the capabilities required for the virtual machines that you wish to deploy in your environment. In this particular example, I wish to specify an availability requirement. In this case, I want the VMs which have this policy associated with them to tolerate at least one failure (host, network or disk).



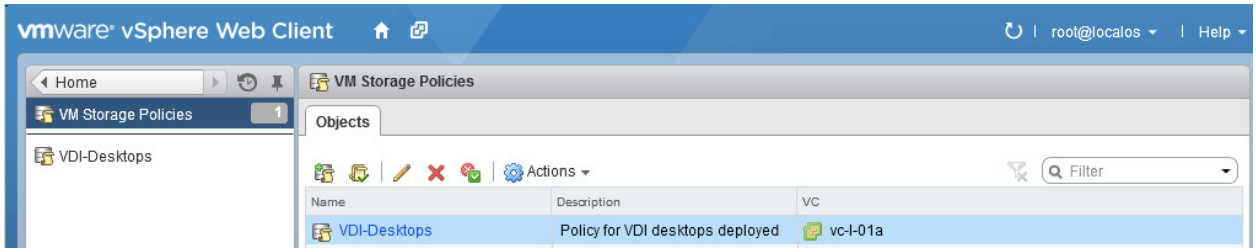
The nice thing about this is immediately I can tell whether or not any datastores are capable of understanding the requirement in the **matching resources** window. As you can see, my vsanDatastore is capable of understanding these requirements that I have placed in the VM Storage Policy:



Note that this is no guarantee that the datastore can meet the requirements in the VM Storage Policy. It simply means that the requirements in the VM Storage Policy can be understood by the datastores that show up in the matching resources.

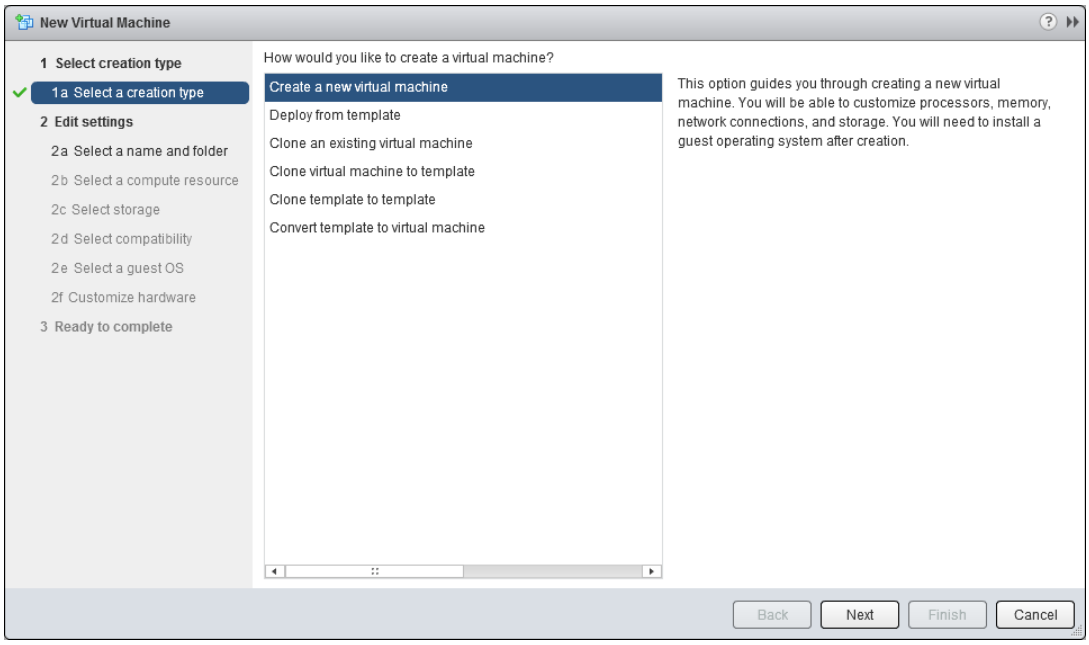
This is where we start to define the requirements for our VMs and the applications running in the VMs. Now, we simply tell the storage layer what our requirements are by selecting the appropriate VM Storage Policy during VM deployment, and the storage layer takes care of deploying the VM in such a way that it meets those requirements.

Complete the creation of the VM Storage Policy. This new 'policy' should now appear in the list of VM Storage Policies.

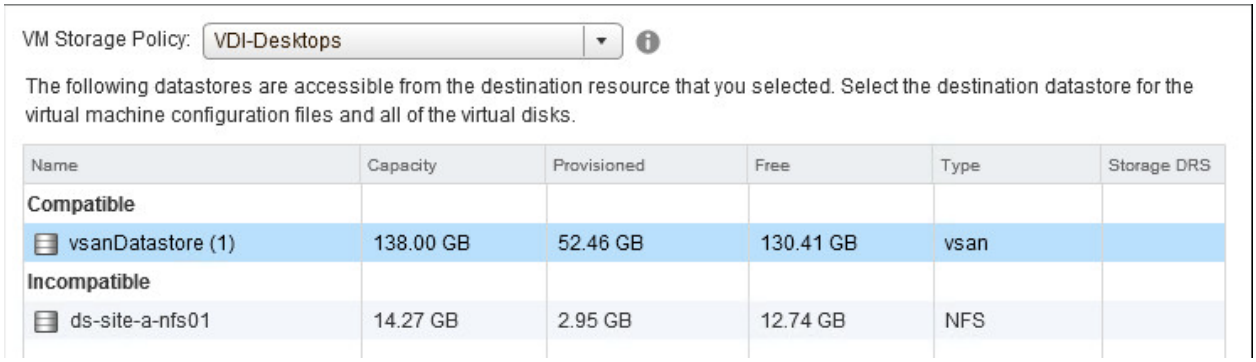


4.1 Deploy VMs

Create a virtual machine, which uses the ***VDI-Desktops*** profile created earlier.

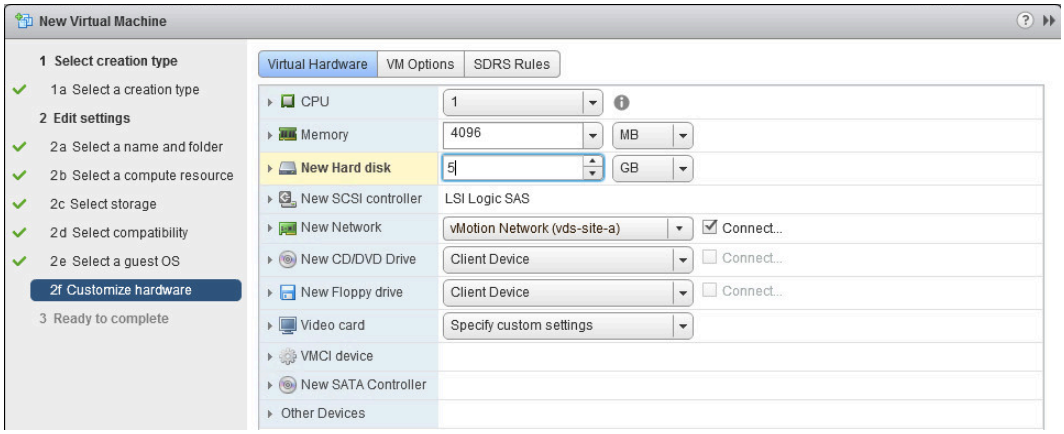


Since DRS is not enabled, you will have to choose a host for this VM. Choose host esx-01a, esx-02a or esx-03a. Do not use esx-04a at this time. When it comes to selecting storage, you can now specify a VM Storage Policy (in this case ***VDI-Desktops***). This will show that **vsanDatastore** is **Compatible** as a storage device, meaning once again that it understands the requirements placed in the storage policy. *It does not mean that the vsanDatastore will implicitly be able to accommodate the requirements – just that it understands them. This is an important point to understand about Virtual SAN.*

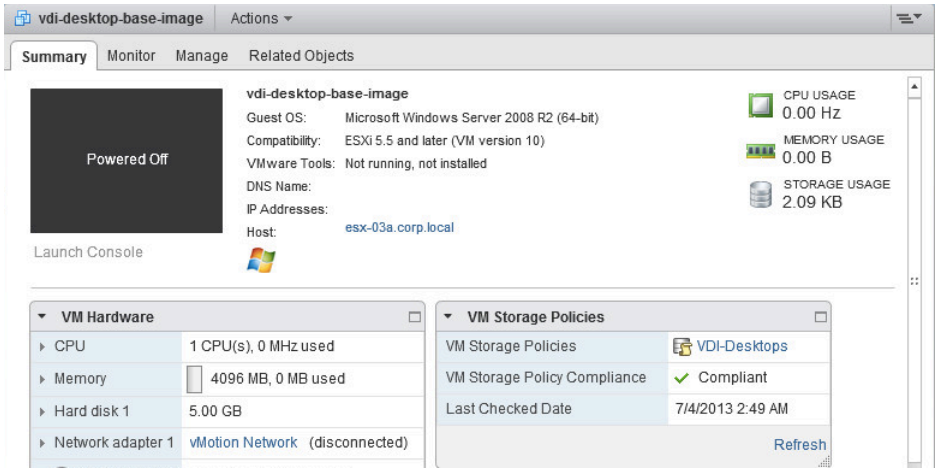


Continue with the creation of this Virtual Machine, selecting the defaults for the remaining steps, including compatibility with ESXi 5.5 and later and Windows 2008 R2 (64-bit) as the Guest OS.

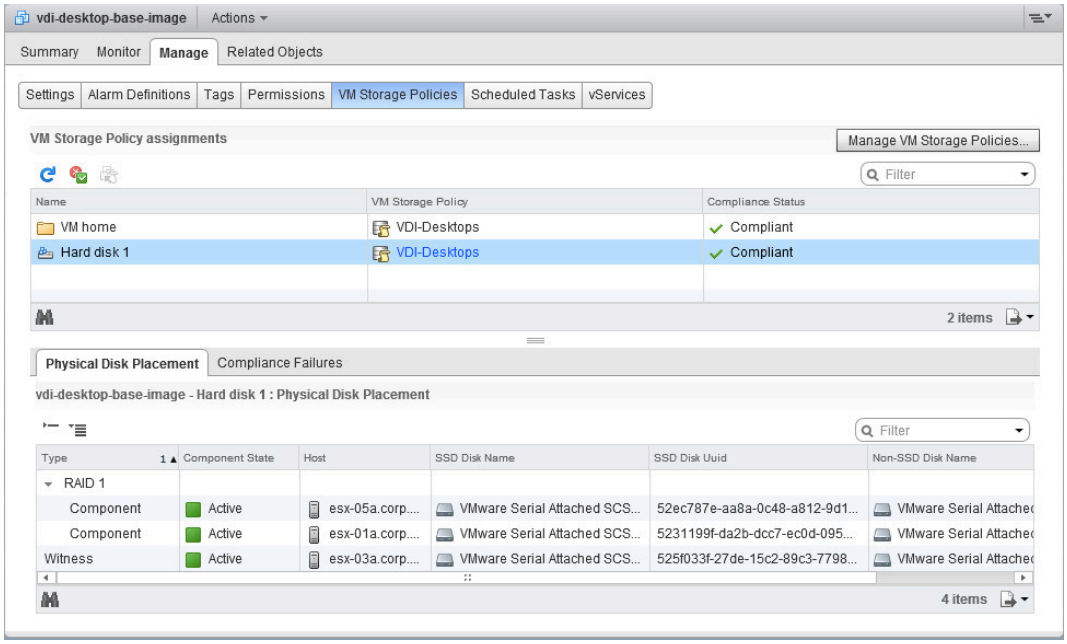
When you get to the **2f. Customize hardware** step, in the Virtual Hardware tab, expand the New Hard Disk virtual hardware and you will see VM Storage Policy set to *VDI-Desktops*. **Reduce the Hard Disk Size to 5GB in order for it to be replicated across hosts** (the default size is 40GB – we want to reduce this as this is a small lab environment)



Complete the wizard. When the VM is created, look at its **Summary** tab and check the compliance state in the VM Storage Policies window. It should say Compliant with a green check mark.



As a final step, you might be interested in seeing how your virtual machine's objects have been placed on the **vsanDatastore**. To view the placement, select your Virtual Machine > Manage tab > VM Storage Policies. If you select one of your objects, the Physical Disk Placement will show you on which host the components of your objects reside, as shown in the example below.



The RAID 1 indicates that the VMDK has a replica. This is to tolerate a failure, the value that was set to 1 in the policy. So we can continue to run if there is a single failure in the cluster. The witness is there to act as a tiebreaker. If one host fails, and one component is lost, then this witness allows a quorum of storage objects to still reside in the cluster.

Notice that all three components are on different hosts for this exact reason. At this point, we have successfully deployed a virtual machine with a level of availability that can be used as the base image for our VDI desktops.

Examining the lay out of the object above, we can see that a RAID1 configuration has been put in place by Virtual SAN, placing each replica on different hosts. This means that in the event of a host, disk or network failure on one of the hosts, the virtual machine will still be available.

If the host on which the VM does not reside fails, then no action is required. If the host on which the VM resides fails, then vSphere HA can be used to automatically bring the VM online on one of the remaining hosts in the cluster.

We will examine this interoperability with vSphere HA in a later exercise.

***Key Takeaway**: Policies enable Software Driven Storage. For the first time, administrators can communicate their storage requirements on a per VM basis to the storage layer from vCenter.*

4.2 Modifying VM Storage Policies

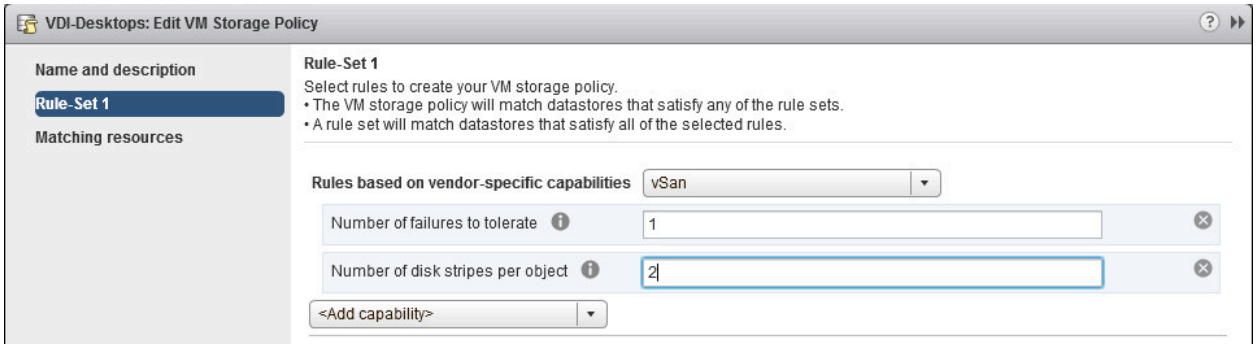
Scenario: customer notices that the VM deployed with the VDI-Desktop policy is getting a 90% read cache hit rate. This implies that 10% of reads need to be serviced from HDD. At peak time, this VM is doing 3000 IOPS. Therefore, there are 300 reads that need to be serviced from HDD. The specifications on the HDDs imply that each disk can do 150 IOPS, meaning that a single disk cannot service these additional 300 IOPS. To meet the I/O requirements of the VM implies that a stripe width of two disks should be implemented.

Edit Profile

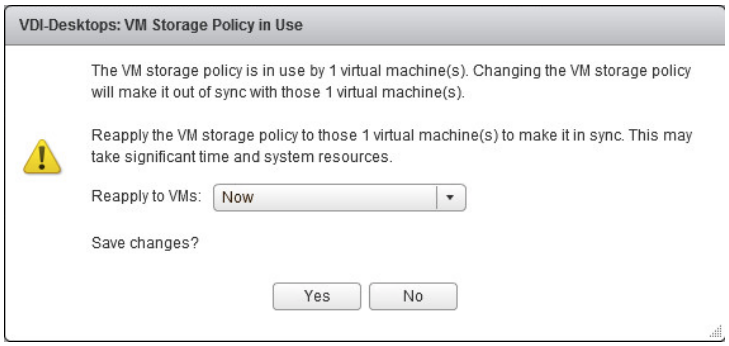
The first step is to edit the VDI-Desktops profile created earlier and add a stripe width requirement to the policy. Navigate back to Rules & Profiles, select VM Storage Policy, select the VDI-Desktop policy and click on ‘edit’.

Add Stripe Width Capability

In the Rule-Set1, add a new capability called ‘Number of disk stripes per object’ and set the value to 2. This is the number of disks that the stripe will span.

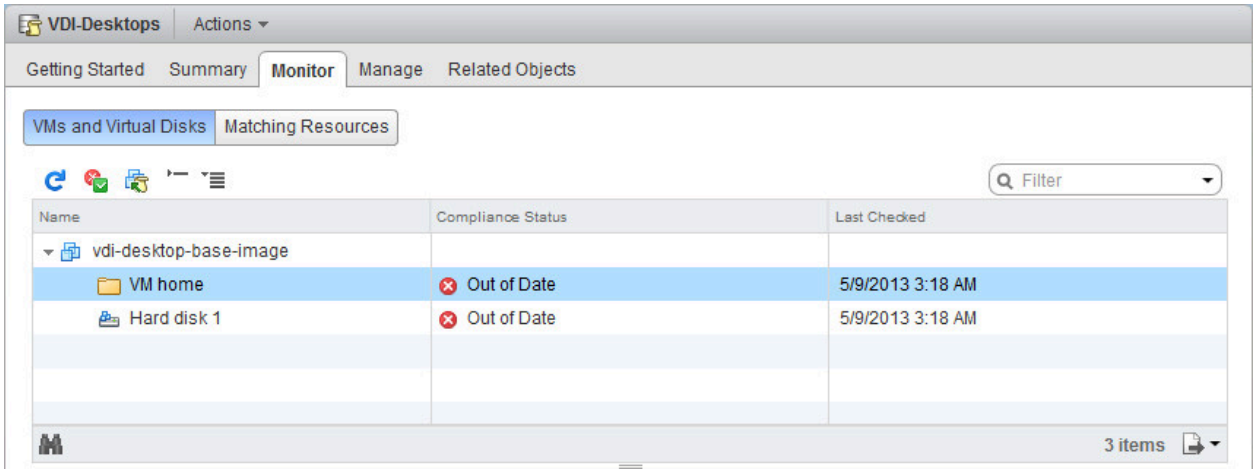


Click OK. You will observe a popup which states that the policy is already in use. We will need to synchronize the virtual machine with the policy after saving the changes. Change the ‘Reapply to VMs’ to Now and Click Yes.

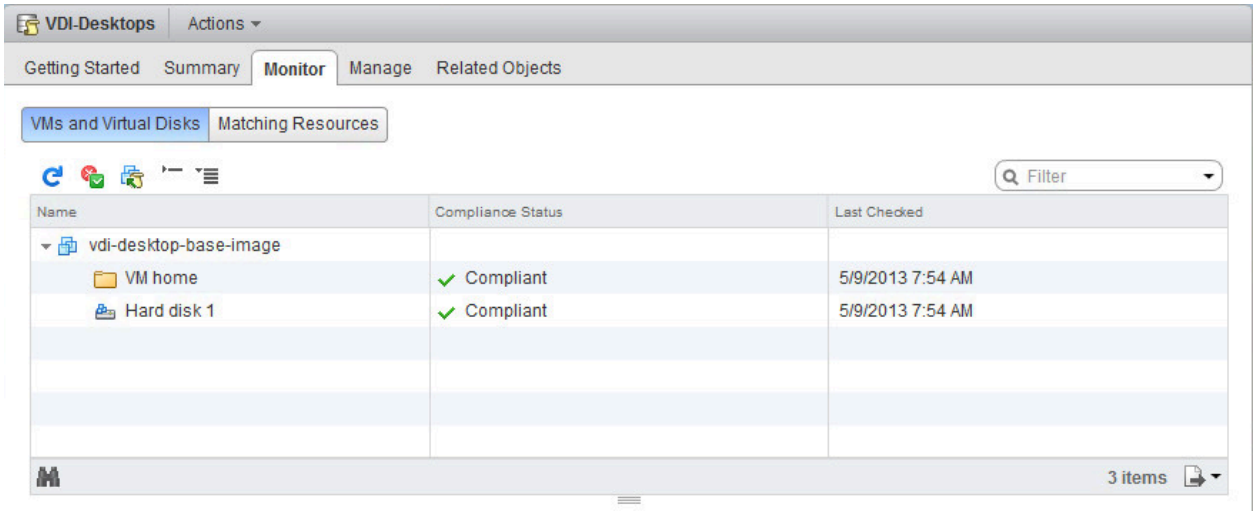


Resync Virtual Machine with Policy Changes

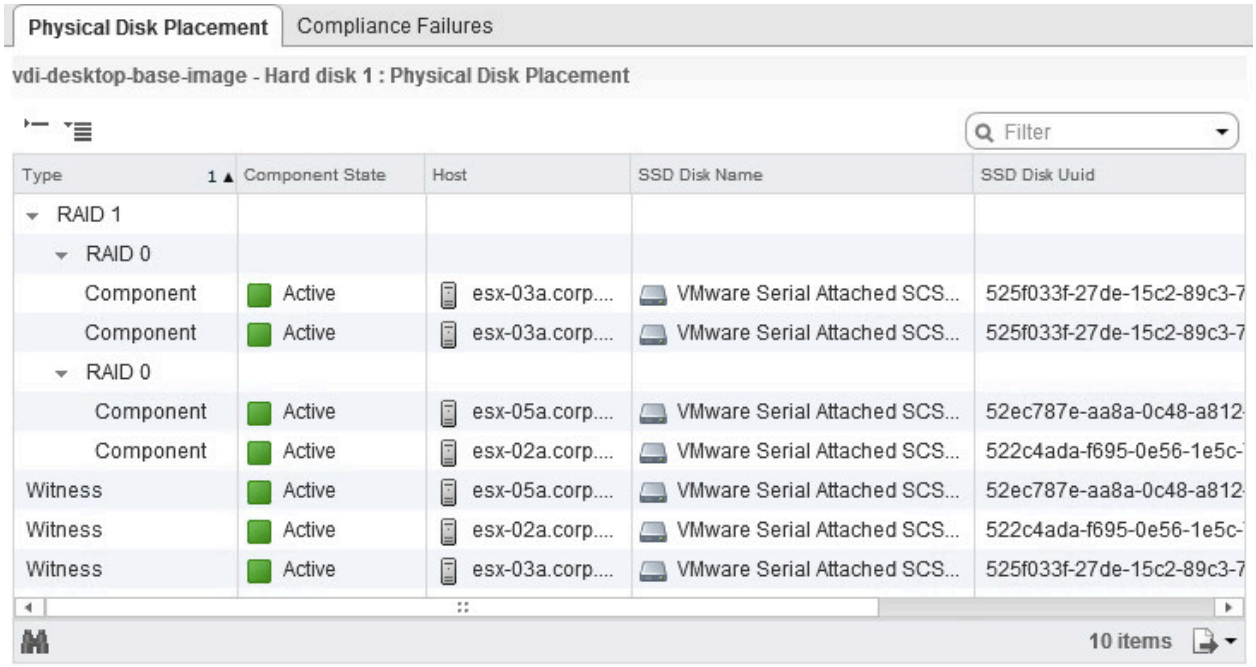
Staying on the VDI-Desktops policy, click on the Monitor tab. In the VMs & Virtual Disks view, you will see that the Compliance Status is 'Out of Date'.



Click on the reapply policy icon (3rd from left) to reapply policy to all out of date entities. Answer yes to the popup. The compliance state should now change once the updated policy is applied.



As a final step, we will now re-examine the layout of the storage object to see if the request to create a stripe width of 2 has been implemented. Return to the Virtual Machine view > Manage > VM Storage Policy & select the Hard Disk 1 object:



Now we can see that the disk layout has changed significantly. Because we have requested a stripe width of two, the components that make up the stripe are placed in a RAID-0 configuration. Since we still have our failures to tolerate requirement, these RAID-0s must be mirrored by a RAID-1. Now that we have multiple components distributed across the 3 hosts, additional witnesses are needed in case of a host failure.

We are not going to install a Guest OS in this virtual machine. Instead we will focus our attention on another small virtual machine available in your environment for the remaining tests.

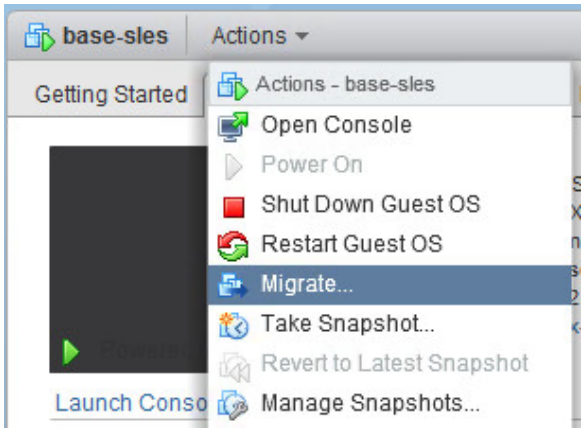
***Key Takeaway**: As virtual machine storage requirements change, administrators of Virtual SAN can simply update the policy. Compare this to a physical SAN or NAS infrastructure where a new datastore would have to be provisioned to satisfy changing virtual machine I/O requirements.*

Exercise 5 – vMotion & Storage vMotion

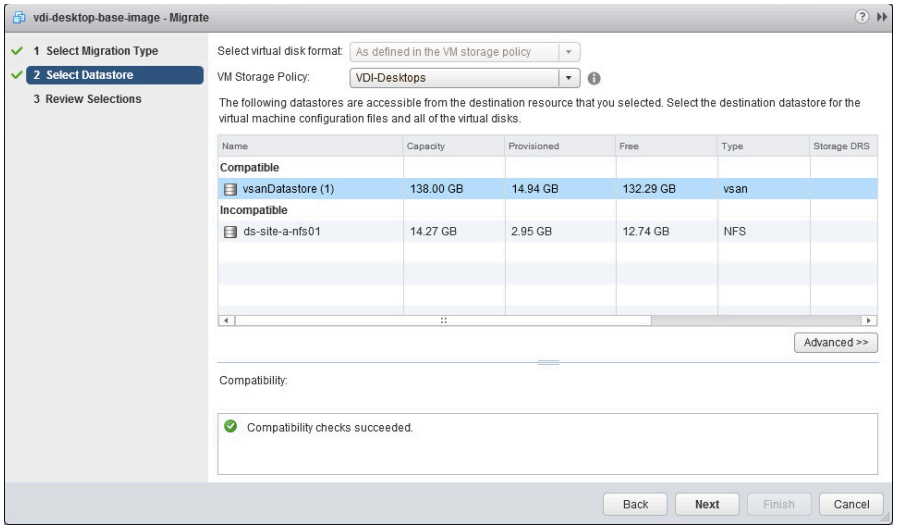
In this Exercise, we will examine the interoperability of Virtual SAN with core vSphere features such as vMotion & Storage vMotion. Power on the virtual machine called **base-sles** which resides on host esx-01a. This is a very small virtual machine, but will be sufficient for the purposes of this lab. Wait until the VMware Tools show as running before continuing. This should only take a moment or two.

5.1 Storage vMotion from NFS to vsanDatastore

This VM currently resides on an NFS datastore called **ds-site-a-nfs01**. We will migrate this virtual machine to the **vsanDatastore**. With the virtual machine **base-sles** selected in the inventory, from the actions list, choose the option to **migrate**.



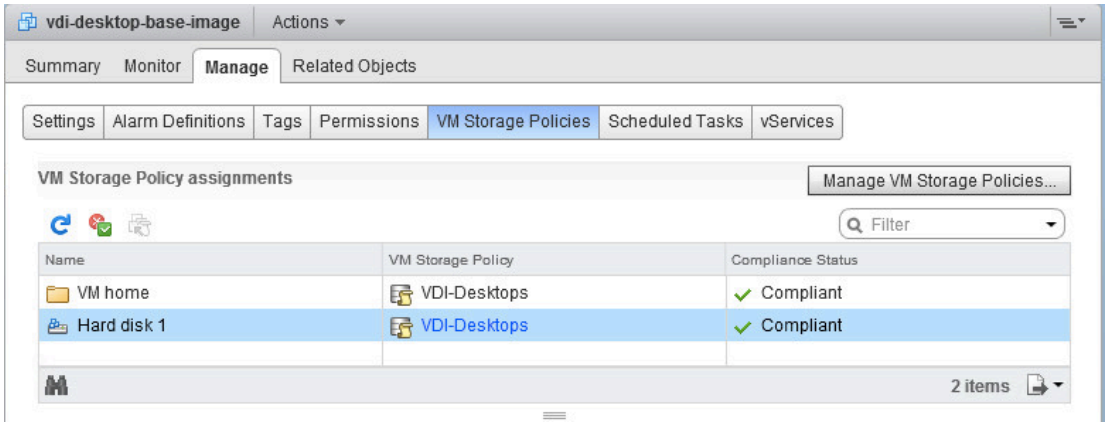
Choose the option to '**Change datastore**'. At the Select Datastore window, change the VM Storage Profile to **VDI-Desktops**. This will show the vsanDatastore as **Compatible**.



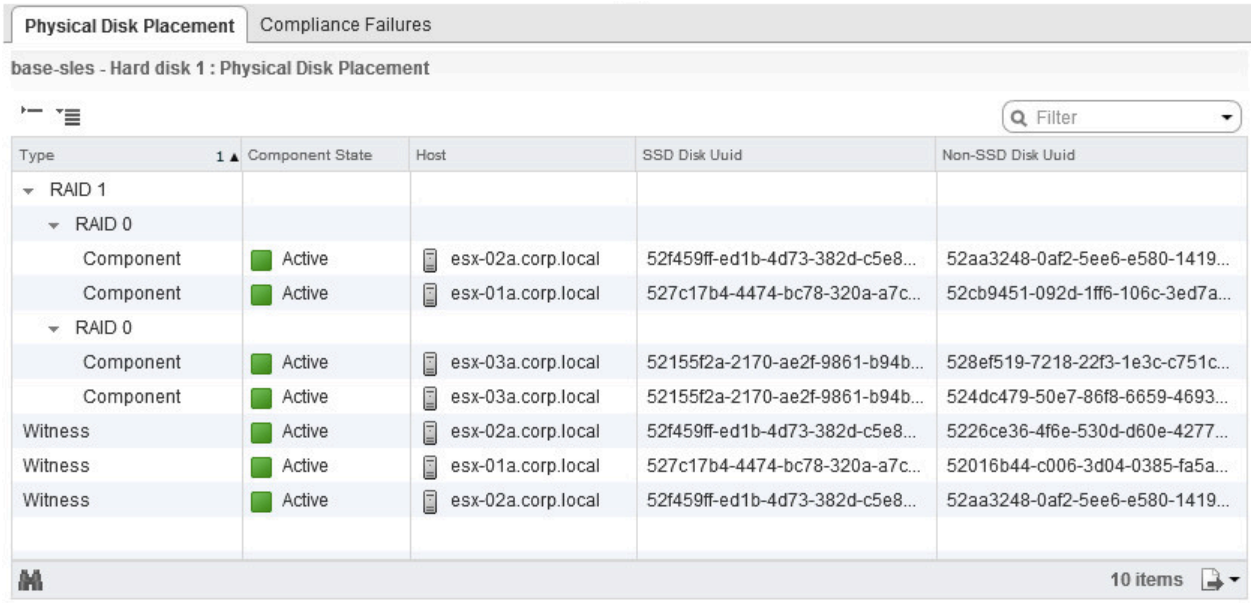
Finish the migration process and wait for the VM to migrate (a few minutes). This demonstrates that you can migrate from traditional datastore formats such as NFS & VMFS to the new vsanDatastore format.

Once the VM has been successfully migrated to the vsanDatastore, examine the layout of the VM. It should have the same layout as the vdi-desktop VM that we created earlier, i.e. a mirror and stripe configuration.

Navigate to the Virtual Machine view > Manage > VM Storage Policies. On first observation, you may see the VM home and Hard disk 1 objects state that the compliance state is **not applicable**. Simply click on the check compliance state icon (middle icon), and this should make them compliant.



Now select the Hard Disk 1 object and have a look at the physical disk placement:



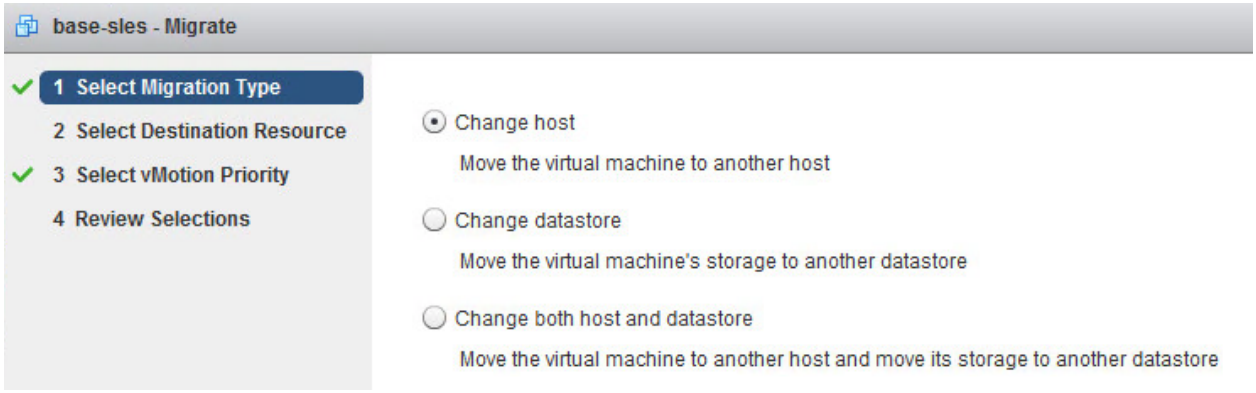
You should be able to see the stripe width of 2 (RAID-0) and the replication/mirror of each (RAID-1). And of course, we also have the tiebreaker witness disks.

5.2 (OPTIONAL) vMotion from host with local storage to host without local storage

Now we will show how hosts which are in the Virtual SAN cluster, but do not have any local storage, can still use the vsanDatastore to run VMs. This exercise can only be performed if a host without local storage has been included in the cluster, and may be skipped otherwise.

At this point, the virtual machine **base-sles** resides on the vsanDatastore. The VM is currently on a host that contributes local storage to the vsanDatastore (esx-01a.corp.local). We will now move this to a host (esx-04a.corp.local) that does not have any local storage.

Once again select the **base-sles** virtual machine from the inventory. From the Actions drop down menu, once again select **Migrate**. This time we choose the option to 'Change host'.



At Exercise 3 where a host selection needs to be made, select host esx04a.corp.local. Complete the migration by selecting all of the defaults from the remaining Exercises in the migration wizard.

When the migration has completed, you will see how hosts that do not contribute any local storage to the vsanDatastore can still run virtual machines. This means that Virtual SAN can be scaled out on a compute basis.

To complete this section, migrate the VM back to a host that has local storage making up the Virtual SAN datastore, e.g. esx-01a, esx-02a or esx-03a. Leave the VM residing on the vsanDatastore.

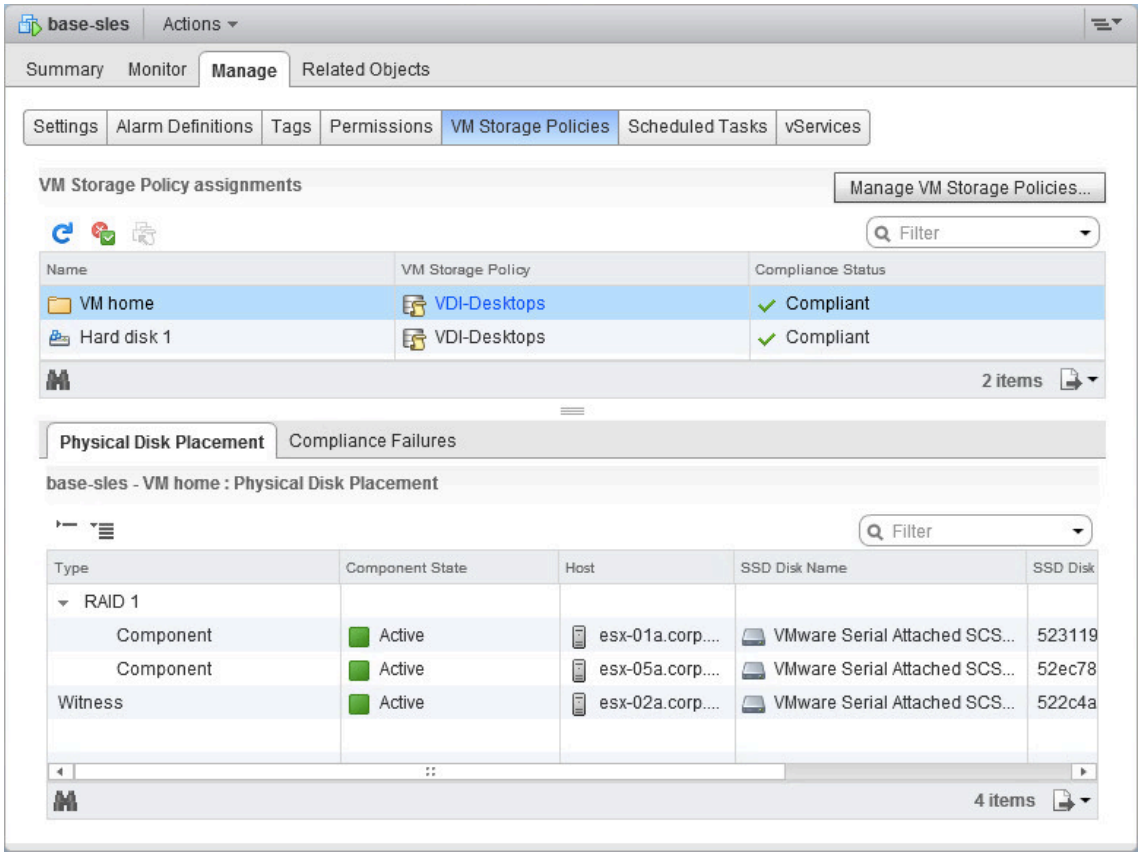
***Key Takeaway**: VMs can be migrated between a vsanDatastore, traditional VMFS and NFS datastores.*

Exercise 6 - vSphere HA & Virtual SAN Interoperability

This final exercise will provide details on how to evaluate Virtual SAN with vSphere HA.

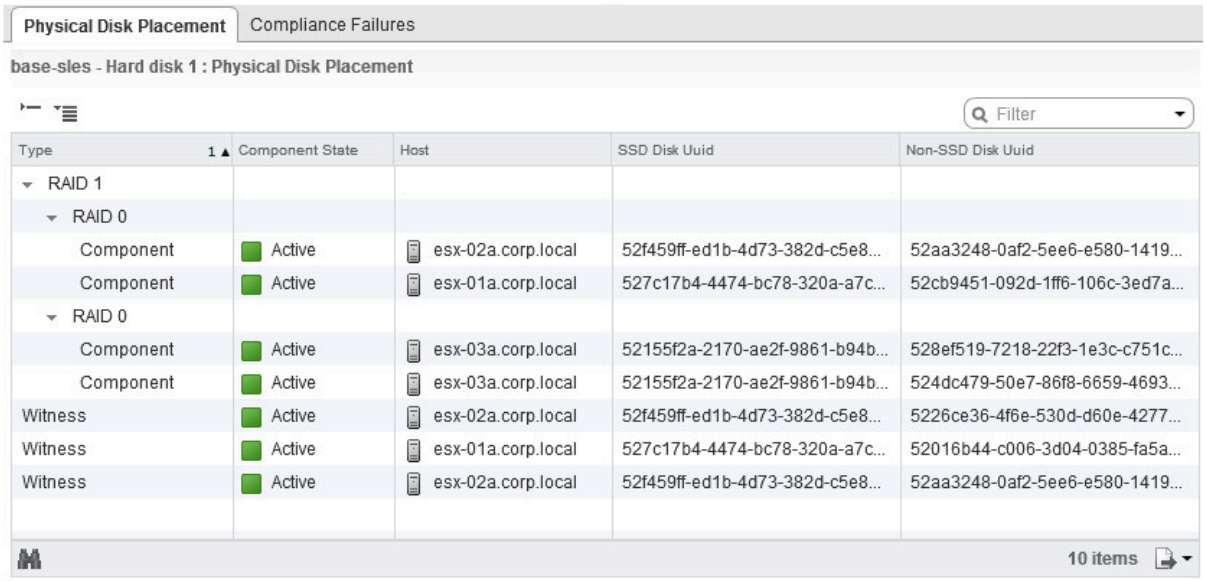
6.1 Check base-sles object layout

First, let's examine the object layout of the virtual machines. Let's first look at VM Home:



This storage object has 3 components, two of which are replicas making up a RAID-1 mirror. The third is a witness disk that is used for tie breaking.

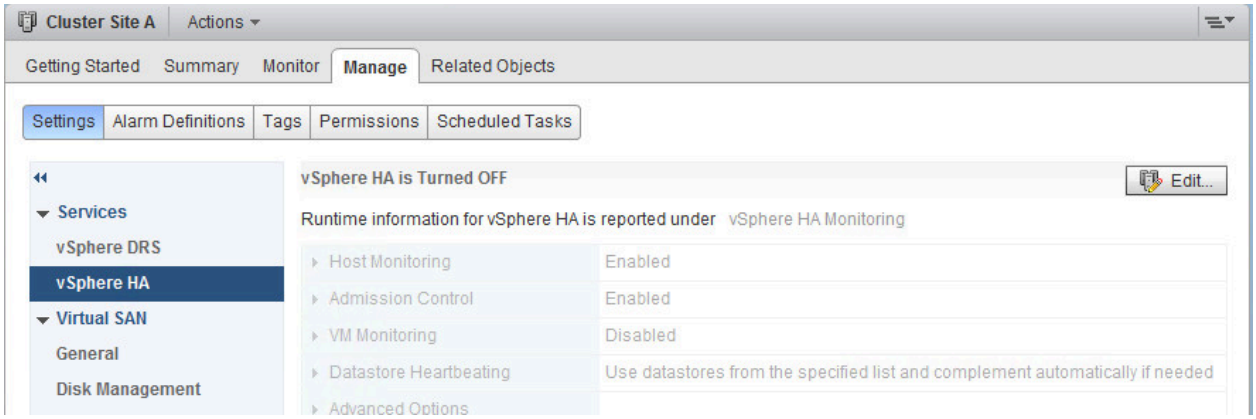
The next object is the disk, which we have looked at a number of times already. Just to recap, this has a StripeWidth set to 2; therefore there is a RAID-0 stripe component across two disks. There is no magic here – to mirror an object with a striped width of 2, 4 disks are required. Again, since ComponentFailuresToTolerate is set to 1, there is also a RAID-1 configuration to replicate the stripe. So we have two RAID-0 (stripe) configurations, and a RAID-1 to mirror the stripes. The witnesses are once again used for tie-breaking functionality in the event of failures.



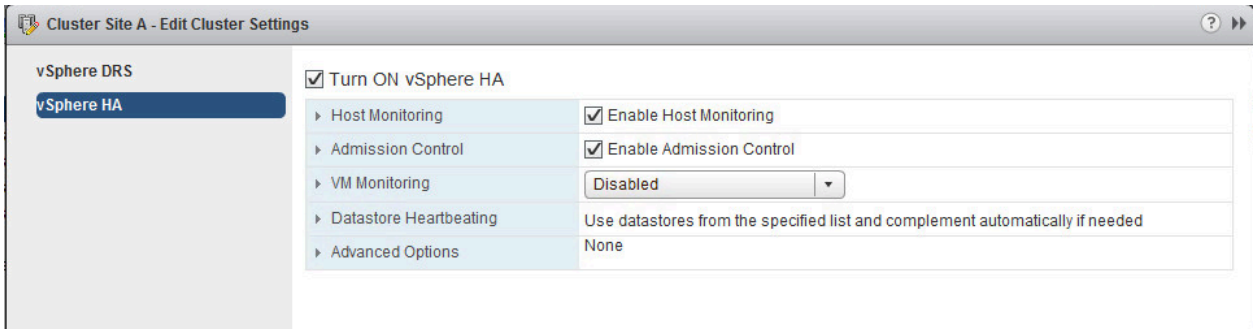
The next Exercise is to invoke some failures in the cluster to see how this impacts the components that make up our virtual machine storage objects, but also how Virtual SAN & vSphere HA interoperate to enable availability.

6.2 Enable HA on the cluster

Navigate to the Cluster and select the Manage tab > Settings. Select the vSphere HA service. vSphere HA is currently Turned OFF:



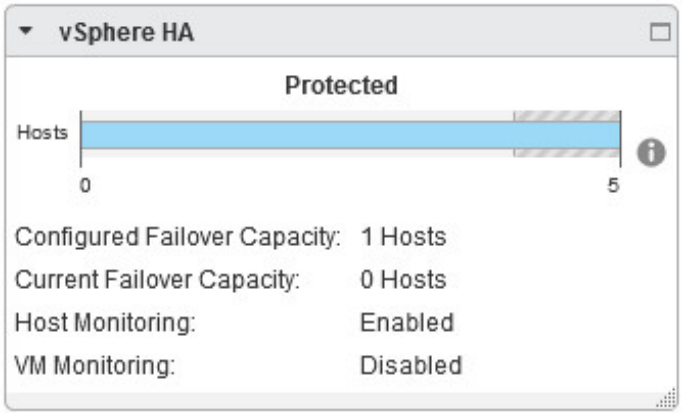
Click on the Edit button, and click on the checkbox to Turn ON vSphere HA.



By default, the vSphere HA Admission Controls have been set to tolerate a single host failure. You can examine this if you wish by breaking open the Admission Control settings to verify. When satisfied, click on the OK button to enable HA.

After enabling HA, you will see a warning about insufficient resources to satisfy vSphere HA failover level. This is a transient warning and will eventually go away after a few moments, once the HA cluster has finished configuring. You can try refreshing from time to time to remove it.

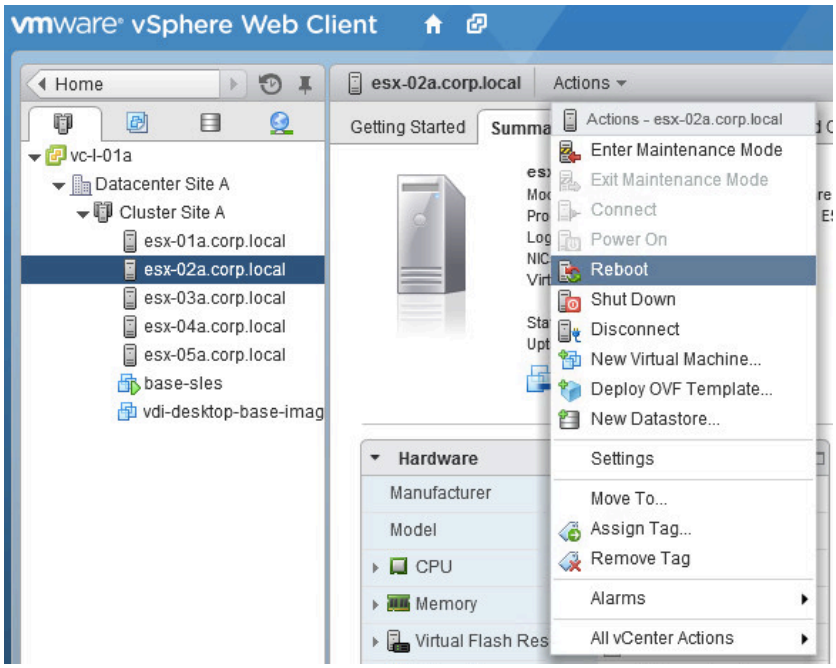
The cluster summary tab should show a vSphere HA overview as follows:



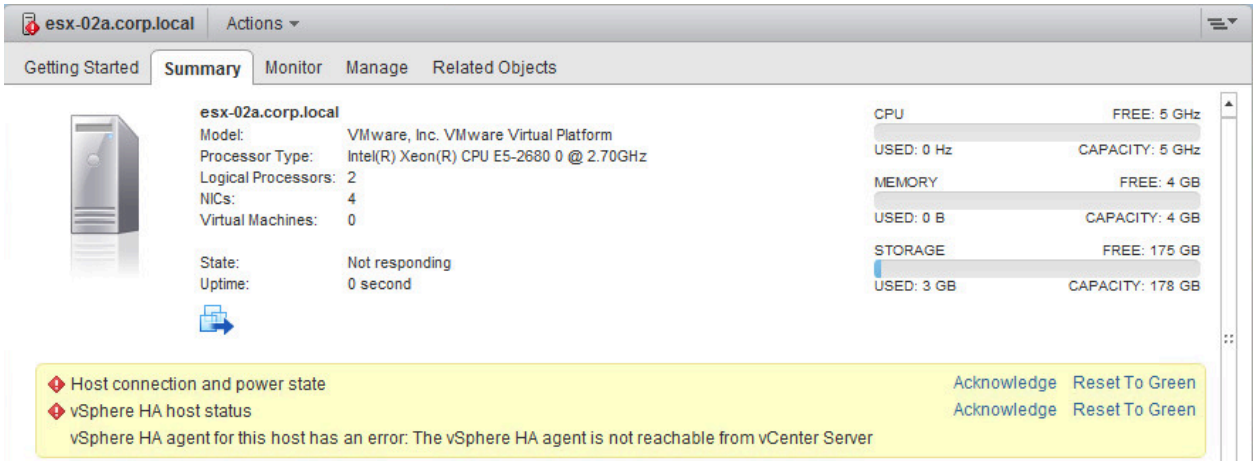
6.3 Host Failure – No running VMs

In this first failure scenario, we will take one of the hosts out of the cluster. This host does not have any running VMs, but we will use it to examine how the Virtual SAN replicas provide continuous availability for the VM, and how the Admission Control setting in vSphere HA and the ComponentFailuresToTolerate are met.

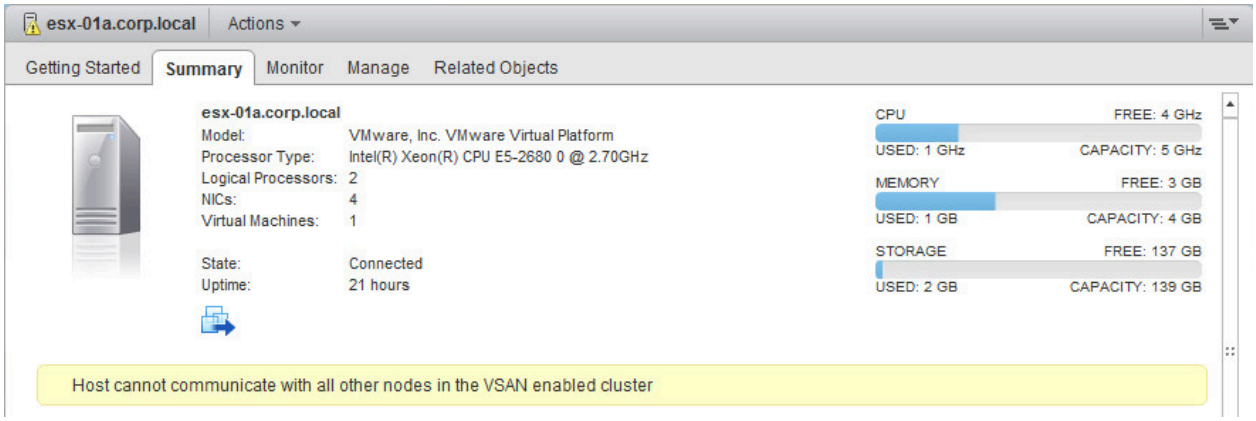
In this Exercise, host esx-02a is rebooted. Select the **Reboot** option from the ESXi host actions:



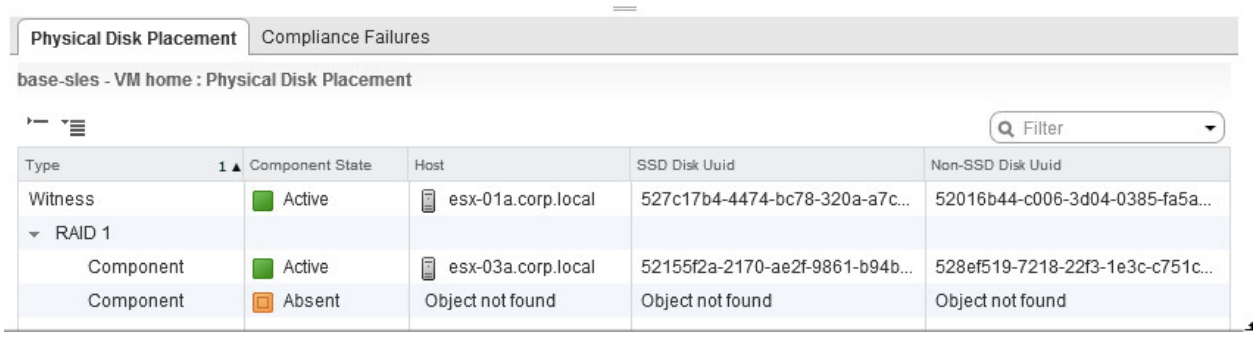
In a short time, we see warnings and errors related to the fact that vCenter can no longer reach the HA Agent and then we see errors related to host connection and power status.



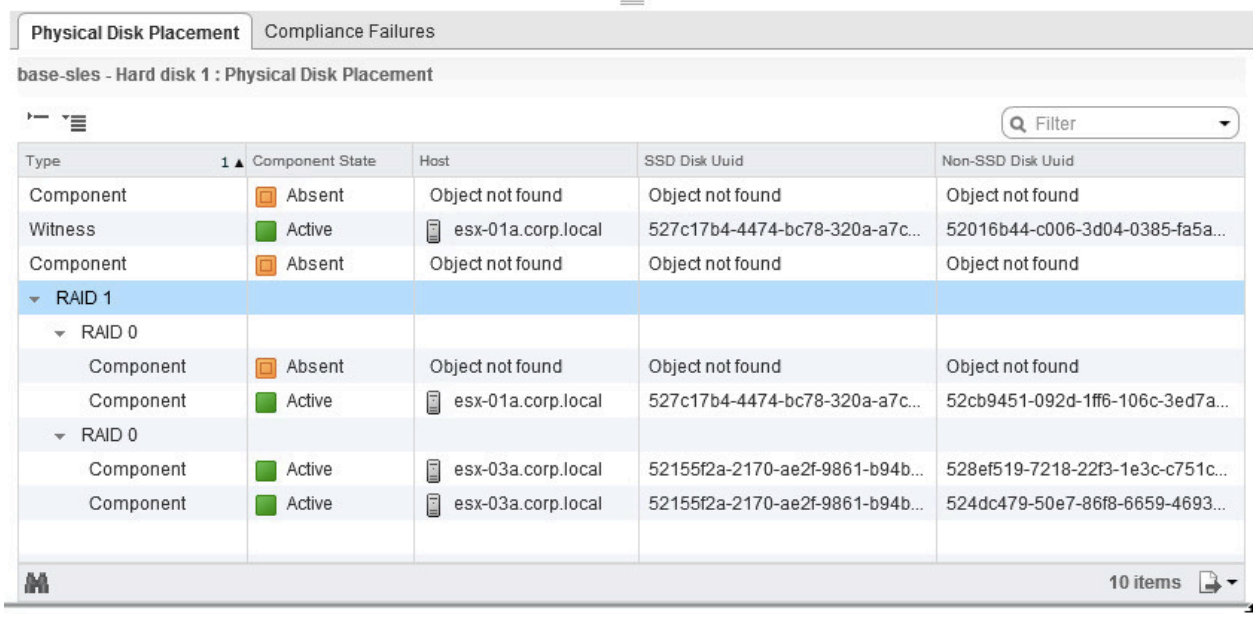
If we check on other hosts in the cluster, we see Virtual SAN communication issues.



With one host out of the cluster, and object components that were held on that host are displayed as **Absent** – Object not found. First we will look at the **VM home**:



Next we can look at the **Hard Disk**:



Basically any components on the rebooted host show up as **Absent**. When the host rejoins the cluster, all components are put back in an Active state. A bitmap of blocks that have changed since the component went absent is maintained. The resync process only needs to resync changed blocks. Now we can see one

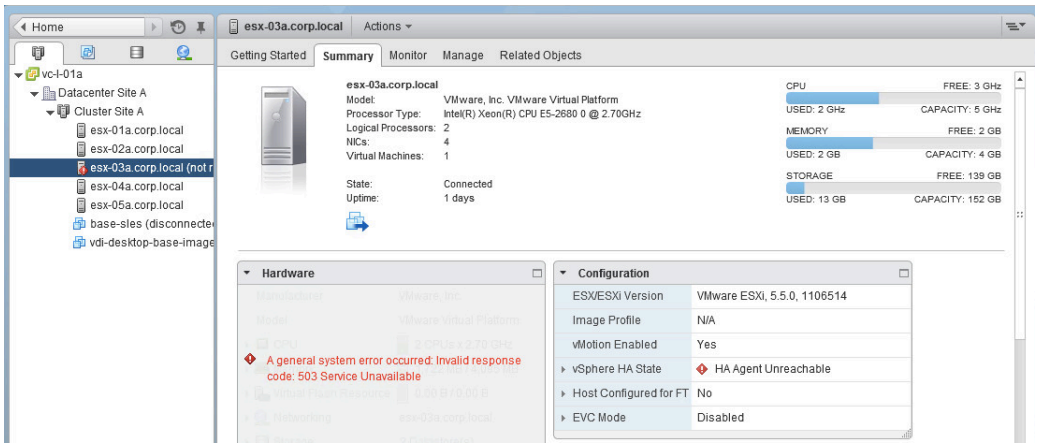
part the availability aspect of Virtual SAN, and how virtual machines continue to run even if components go absent. If the host remains absent for more than 30 minutes, the missing components are rebuilt (reconfigured) on the remaining hosts and disks in the cluster.

6.4 Host Failure – Running VMs

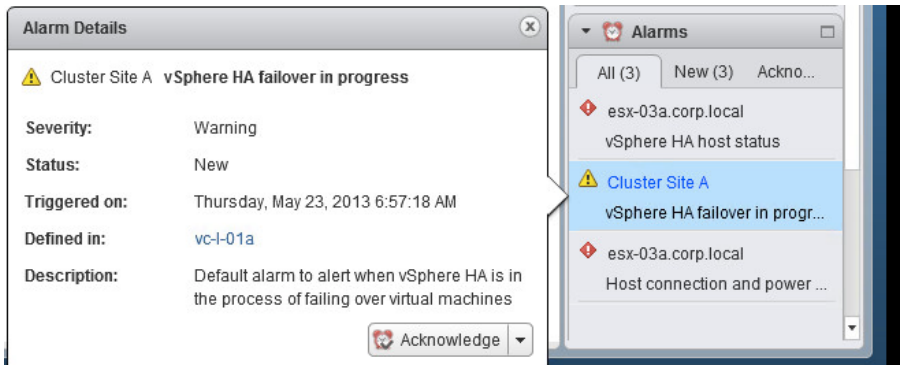
Wait for the host to reboot from the previous test before continuing. Remember that we have only set *ComponentFailuresToTolerate* to 1. In this next example, we will halt the ESXi host (in this example, esx-03a), which contains a running VM **base-sles**. Here we will see interoperability between HA and Virtual SAN.

From the Control Center desktop Start button, navigate to All Programs > PuTTY > PuTTY. Select the host esx-04 from the list and launch an SSH session. Login with the credentials *root* and give the *VMware1!* password. Type the command **halt** in the shell.

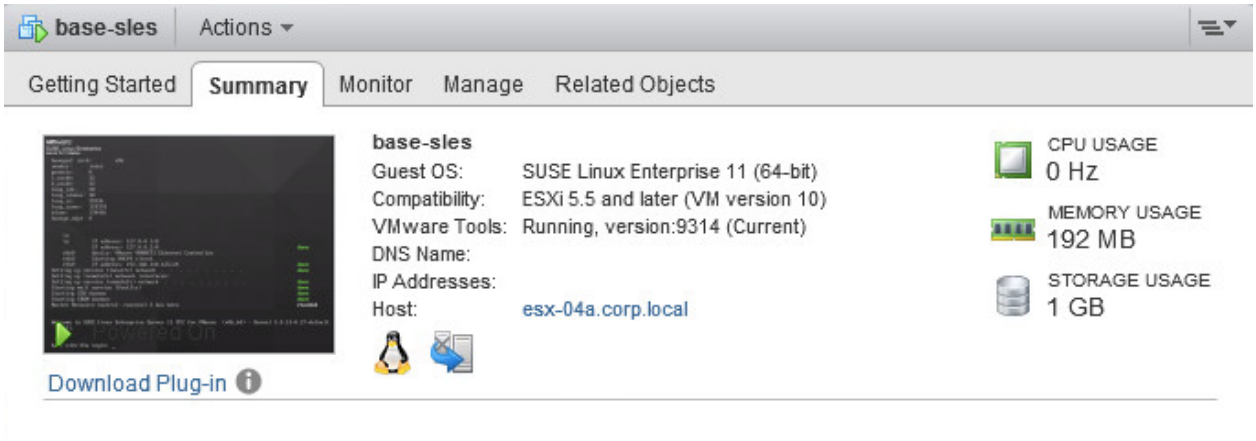
Once again, you will see vSphere HA detect the error. And as before, vCenter reports on the host connection and power state:



If you go immediately to look at the VM's storage object layout, you might find that you can no longer query the component state since the host that owned the object has gone down. Once the VM has successfully failed over to an alternate host, you will once again be able to query the object layout. You should see vSphere HA kicking in, and failing over the VM from the failed host to another host in the cluster.



Finally, check the status of the VM. In my test, the VM is successfully started on host esx-04a (although it may be restarted on a different host in your environment). Note that you should refresh the UI periodically to see these changes occur.



The VM is running, now examining the object layout. It should reveal that not all components are present. However there is still a quorum of objects available, enabling the VM to tolerate a host failure. If the failure persists for longer than 30 minutes, the components will be rebuilt on the remaining disks in the cluster.

***Key Takeaway**: While the policy can provide highly available virtual machine storage objects, interoperability between Virtual SAN & vSphere HA provides high availability at both the storage AND compute layers.*

Appendix A: Pre-requisites Checklist

	REQUIREMENT
☐	4 (optionally 5) hosts as per the VMware Compatibility Guide with vSphere 5.5 Update 1 installed.
☐	4 hosts must contain local storage. Local storage must comprise of at least one empty HDD & one empty SSD, and must be supported by VSAN as per the VMware Compatibility Guide. In this POC, each of the 4 hosts has 1 x SSD & 2 x HDD. The fifth optional host does not contain any local storage; if you choose to omit this host, you can still perform nearly all the exercises in this POC.
☐	The 4 hosts with storage must have a separate boot device. For hosts with less than 512 GB of memory, the boot device can be an SD card, a USB drive, or a dedicated hard drive. For hosts with 512 GB of memory or greater, the boot device must be a separate magnetic disk or solid-state drive.
☐	Each host must contain a RAID Controller that is capable of running in pass-thru mode (aka HBA mode) or RAID0 mode. The RAID Controller must be supported by VSAN as per the VMware Compatibility Guide. If using RAID0 mode, each drive needs to be presented as an individual RAID0 set; consult the vendor guidelines for how to do this for a specific RAID controller. NOTE: if the RAID0 set for an SSD drive is not recognized as an SSD by vSphere, follow the steps in this KB article: http://kb.vmware.com/kb/2013188.
☐	1 x vCenter Server with vSphere Enterprise Plus licensing (an evaluation license is used if only vSphere Enterprise is available) 3 ESXi Hosts added to a Cluster. No services should be enabled on the cluster initially (no HA, no DRS and no Virtual SAN)
☐	10Gb network is preferable and is highly recommended for production environments, but Virtual SAN can & will work with a 1Gb network. The network also needs to support Layer-2 multicast (e.g. IGMP Snooping). Virtual SAN does NOT require Layer-3 multicast.
☐	A vMotion VMkernel port must be configured for each host. If IP storage is available for virtual machines templates or ISOs, this must also be configured.
☐	A virtual machine with an installed Guest OS should be available. If this is not available, a new virtual machine will have to be created and a Guest OS installed.

Appendix B: Troubleshooting Virtual SAN

In this appendix, a number of utilities that can help you to troubleshoot Virtual SAN are covered.

B1. Troubleshooting Virtual SAN using the ESXCLI

vSphere 5.5 introduces a new **esxcli** namespace specifically for Virtual SAN.

```
~ # esxcli vsan
Usage: esxcli vsan {cmd} [cmd options]

Available Namespaces:
  datastore             Commands for VSAN datastore configuration
  network               Commands for VSAN host network configuration
  storage               Commands for VSAN physical storage configuration
  cluster               Commands for VSAN host cluster configuration
  maintenancemode       Commands for VSAN maintenance mode operation
  policy                Commands for VSAN storage policy configuration
  trace                 Commands for VSAN trace configuration

~ #
```

B2. Troubleshooting Virtual SAN using the RVC

The vCenter Appliance now ships with RVC, the Ruby Virtual Console. This tool is an excellent troubleshooting aid for monitoring and troubleshooting issues on a Virtual SAN. The tool is available in vCenter 5.5 and is pre-installed on both the Windows version of vCenter and the appliance version.

To get started with the RVC, from your vCenter command line type **rvc <hostname>** where <hostname> can be an ESXi host or your vCenter server. Navigation of the inventory is done using **ls** and **cd** commands. Each inventory object is associated with a numeric value.

RVC comes with an extensive help subsystem. Type **help vsan** for a list of Virtual SAN supported commands. Type **help vsan.<command>** for help on a specific command:

```
> help vsan
Commands:
enable_vsan_on_cluster: Enable VSAN on a cluster
disable_vsan_on_cluster: Disable VSAN on a cluster
host_consume_disks: Consumes all eligible disks on a host
host_info: Print VSAN info about a host
cluster_info: Print VSAN info about a cluster
disks_info: Print physical disk info about a host
cluster_set_default_policy: Set default policy on a cluster
object_info: Fetch information about a VSAN object
disk_object_info: Fetch information about all VSAN objects on a given physical disk
cmmds_find: CMMDS Find
vm_object_info: Fetch VSAN object information about a VM
disks_stats: Show stats on all disks in VSAN

To see commands in a namespace: help namespace_name
To see detailed help for a command: help namespace_name.command_name
>
```

There are many other commands that can assist you in getting additional details around Virtual SAN. It is worthwhile spending some time evaluating the RVC and the various commands that you have at your disposal.

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

B3. Troubleshooting Virtual SAN using the VSAN Observer

VSAN Observer is a performance-monitoring tool that is invoked through the RVC of vCenter. It is web based and provides granular performance statistics on each of the components in a Virtual SAN. It provides live statistics every 60 seconds and allows an administrator to identify bottlenecks anywhere within the Virtual SAN environment.

As well as looking a the Virtual SAN cluster as a whole, performance information can be retrieved on SSD IOPS, Latency and Read Cache Hit Ratio, Magnetic disk performance, Virtual SAN CPU utilization & Virtual Machine VMDK performance.

Data is refreshed every 60 seconds to give the “Average” for that 60 second period.

There are three ways that VSAN Observer can be ran

- Live Monitoring
    - Can be cancelled
    - Automatically terminates after two hours
- Offline Monitoring
    - Data can be replayed on another system through a web browser
- RAW stats bundle
    - For deeper “typically” engineering analysis

From an RVC command window navigate to your Virtual SAN cluster and run the following command:

vsan.observer ~/computers/VSAN <options>

where the <options> can be:

- Live Monitoring: --run-webserver --force
- Offline Monitoring: --generate-html-bundle
- RAW Stats Bundle: --filename

By default, VSAN Observer will output to port 8010, but this can be changed in the options.

To see the Virtual SAN performance, point a browser to the IP Address of your vCenter server and the port specified when starting the observer tool.

VSAN Observer – VSAN client (per host) view:

This is the VSAN Observer landing page and gives a good overview of the Virtual SAN cluster as a whole. This cluster has three ESXi hosts. Essentially what you are looking for is that each of the graphs that are displaying information is ‘green’. Green means that everything is operating within its expected parameters. Each of the graphs can be clicked on for a full-sized view.