

VMware Virtual SAN 6.1 Proof Of Concept Guide

January 2016 Edition

Cormac Hogan David Boone Paudie O'Riordan Brad Garvey



Contents

1. INTRODUCTION	6
2. BEFORE YOU START	6
2.1 ALL FLASH OR HYBRID	6
2.2 THREE-NODE VERSUS FOUR-NODE OR GREATER	6
2.3 FOLLOW THE VSPHERE COMPATIBILITY GUIDE PRECISELY	7
2.3.1 Why Is This Important?	7
2.3.2 Hardware, Drivers, and Firmware	7
2.3.3 Out-of-box Drivers versus Inbox (Shipped with ESXi and Listed on VCG)	7
2.3.4 RAID-0 versus Pass-Through for Disks	7
2.3.5 Controller Configuration	8
2.4 USE SUPPORTED VSPHERE SOFTWARE VERSIONS	8
3. VIRTUAL SAN POC SETUP ASSUMPTIONS AND PREREQUISITES	9
4. VIRTUAL SAN NETWORK SETUP	12
4.1 Advantages of Distributed Switch versus Standard Switch	12
4.2 CREATING A VMKERNEL PORT FOR VIRTUAL SAN	12
5. ENABLING VIRTUAL SAN ON THE CLUSTER	17
5.1 MANUAL DISK CLAIMING—CREATE DISK GROUPS	19
6. ENABLE THE VIRTUAL SAN HEALTH CHECK PLUGIN	22
6.1 CHECK YOUR NETWORK THOROUGHLY	23
6.1.1 Why Is This Important?	23
6.1.2 Check the Network Partition Groups after Creating Cluster	23
6.1.3 Use the Health Check Plugin to Verify Virtual SAN Functionality	24
6.1.4 Use the Troubleshooting Reference Manual to Verify Network Functionality	26
7. VSPHERE FUNCTIONALITY ON VIRTUAL SAN	27
7.1 DEPLOY YOUR FIRST VM	27
7.2 Snapshot VM	36
7.3 Clone a VM	40
7.4 vMotion a VM between Hosts	43
7.5 Optional: Storage vMotion a VM between Datastores	46
7.5.1 Mount an NFS Datastore to the Hosts	
7.5.2 Storage vMotion a VM from Virtual SAN to Another Datastore Type	46
7.5.3 Storage vMotion of VM to Virtual SAN from Another Datastore Type	
8. SCALE OUT VIRTUAL SAN	49
8.1 Add the Fourth Host to Virtual SAN Cluster	50
8.2 MANUAL OPTION: CREATE DISK GROUP ON NEW HOST	52
8.3 VERIFY VIRTUAL SAN DISK GROUP CONFIGURATION ON NEW HOST	53
8.4 VERIFY NEW VIRTUAL SAN DATASTORE CAPACITY	53
9. VM STORAGE POLICIES AND VIRTUAL SAN	55
9.1 CREATE A NEW VM STORAGE POLICY	56
9.2 DEPLOY A NEW VM WITH THE NEW VM STORAGE POLICY	60

9.3 ADD A NEW VM STORAGE POLICY TO AN EXISTING VM	63
9.4 Modify a VM Storage Policy	65
10. VIRTUAL SAN MONITORING	70
10.1 MONITOR THE VIRTUAL SAN CLUSTER	
10.2 Monitor Virtual Devices in the Virtual SAN Cluster	
10.3 MONITOR PHYSICAL DEVICES IN THE VIRTUAL SAN CLUSTER	
10.4 MONITOR RESYNCHRONIZATION AND REBALANCE OPERATIONS	
10.5 DEFAULT VIRTUAL SAN ALARMS	
10.7 MONITOR VIRTUAL SAN WITH VIRTUAL SAN OBSERVER	75
11. PERFORMANCE TESTING	76
11.1 USE VIRTUAL SAN OBSERVER	
11.2 Performance Considerations	
11.2.1 Single vs. Multiple Workers	
11.2.2 Working Set	77
11.2.3 Sequential Workloads versus Random Workloads	77
11.2.4 Outstanding IOs	77
11.2.5 Block Size	77
11.2.6 Cache Warm up Considerations	77
11.2.7 Number of Magnetic Disk Drives in Hybrid Configurations	
11.2.8 Striping Considerations	
11.2.9 Guest File Systems Considerations	
11.2.10 Performance during Failure and Rebuild	79
11.3 PERFORMANCE TESTING OPTION 1: VIRTUAL SAN HEALTH CHECK	80
11.4 PERFORMANCE TESTING OPTION 2: HCIBENCH	
11.4.1 Where to Get HCIbench	82
11.4.2 Deploying HCIbench	82
11.4.3 Considerations for Defining Test Workloads	87
Results	89
12. TESTING HARDWARE FAILURES	90
12.1 Understanding Expected Behavior	90
12.2 IMPORTANT: TEST ONE THING AT A TIME	90
12.3 VM BEHAVIOR WHEN MULTIPLE FAILURES ENCOUNTERED	90
12.3.1 VM Powered on and VM Home Namespace Object Goes Inaccessible	
12.3.2 VM Powered on and Disk Object Goes Inaccessible	
12.4 WHAT HAPPENS WHEN A SERVER FAILS OR IS REBOOTED?	91
12.5 SIMULATE HOST FAILURE WITHOUT VSPHERE HA	92
12.6 SIMULATE HOST FAILURE WITH VSPHERE HA	95
12.7 DISK IS PULLED UNEXPECTEDLY FROM ESXI HOST	
12.7.1 Expected Behaviors	99
12.8 SSD is Pulled Unexpectedly from ESXi Host	100
12.8.1 Expected Behaviors	100
12.9 WHAT HAPPENS WHEN A DISK FAILS?	101
12.9.1 Expected Behaviors	101
12.10 WHAT HAPPENS WHEN AN SSD FAILS?	102
12.10.1 Expected Behaviors	102
12.11 VIRTUAL SAN DISK FAULT INJECTION SCRIPT FOR POC FAILURE TESTING	103
12.12 Pull Magnetic Disk/Capacity Tier SSD and Replace before Timeout Expires	103

12.13 PULL MAGNETIC DISK/CAPACITY TIER SSD AND DO NOT REPLACE BEFORE TIMEOUT	EXPIRES
	106
12.14 PULL CACHE TIER SSD AND DO NOT REINSERT/REPLACE	
12.15 CHECKING REBUILD/RESYNC STATUS	
12.16 INJECTING A DISK ERROR	
12.16.2 Clear a Permanent Error	115
12.17 WHEN MIGHT A REBUILD OF COMPONENTS NOT OCCUR?	117
12.17.1 Lack of Resources	117
12.17.2 Underlying Failures	117
13. VIRTUAL SAN MANAGEMENT	
13.1 Put a Host into Maintenance Mode	118
13.2 REMOVE AND EVACUATE A DISK	
13.3 EVACUATE A DISK GROUP	
13.4 Add Disk Groups Back Again	
13.5 TURNING ON AND OFF DISK LEDS	
14. VIRTUAL SAN 6.1 STRETCHED CLUSTER CONFIGURATION	
14.1 VIRTUAL SAN 6.1 STRETCHED CLUSTER NETWORK TOPOLOGY	
14.2 VIRTUAL SAN 6.1 STRETCHED CLUSTER HOSTS	
14.3 VIRTUAL SAN 6.1 STRETCHED CLUSTER DIAGRAM	
14.4 Preferred Site Details	
14.4.1 Commands to Add Static Routes	
14.5 Secondary Site Details	
14.5.1 Commands to Add Static Routes	
14.6 A NOTE ON IGMP v3	
14.7 WITNESS SITE DETAILS	
14.7.1 Commands to Add Static Routes	136
14.8 vSphere HA Settings	
14.8.1 Response to Host Isolation	137
14.8.2 Admission Control	137
14.8.3 Advanced Settings	138
14.9 VM Host Affinity Groups	
14.10 DRS SETTINGS	141
15. VIRTUAL SAN STRETCHED CLUSTER NETWORK FAILOVER SCENARIOS	
15.1 NETWORK FAILURE BETWEEN SECONDARY SITE AND WITNESS	
15.1.1 Trigger the Event	142
15.1.2 Cluster Behavior on Failure	142
15.1.3 Conclusion	145
15.1.4 Repair the Failure	146
15.2 NETWORK FAILURE BETWEEN PREFERRED SITE AND WITNESS	147
15.2.1 Trigger the Event	147
15.2.2 Cluster Behavior on Failure	148
15.2.3 Conclusion	
15.2.4 Repair the Failure	
15.3 NETWORK FAILURE BETWEEN WITNESS AND BOTH DATA SITES	
15.3.1 Trigger the Event	151
15.3.2 Cluster Behavior on Failure	151
15.3.3 Conclusion	152
15.3.4 Repair the Failure	

16. FURTHER INFORMATION	152
16.1 VMware Virtual SAN Community	152
16.2 LINKS TO EXISTING DOCUMENTATION	152
16.3 VMware Support	152
APPENDIX A—FAULT DOMAINS	153
A1. SETTING UP FAULT DOMAINS	153
A2. CREATE A POLICY TO LEVERAGE FAULT DOMAINS	155
A3. CREATE A VM AND CHECK THE FAULT DOMAINS	158
APPENDIX B-MIGRATING FROM STANDARD VSWITCH TO DISTRIBUTED	161
B.1 CREATE DISTRIBUTED SWITCH	161
B.2 CREATE PORT GROUPS	163
B.3 MIGRATE MANAGEMENT NETWORK	166
B.4 MIGRATE VMOTION	171
B.5 MIGRATE VIRTUAL SAN NETWORK	171
	1 / 1

1. Introduction

VMware customers love the simplicity, performance and integration of VMware[®] Virtual SAN[™] since its launch.

Most customers choose to evaluate Virtual SAN before using it for production – always a good idea. We've made a list of issues occasionally encountered as people go through this process.

Follow this guide, and you'll have a great evaluation.

2. Before You Start

Plan on testing a reasonable hardware configuration that resembles what you plan to use in production. Refer to the <u>VMware Virtual SAN 6.0 Design and Sizing Guide</u> for information on supported hardware configurations, and consideration when deploying Virtual SAN.

2.1 All Flash or Hybrid

There are a number of additional considerations if you plan to deploy an all-flash Virtual SAN solution:

- All-flash is available in Virtual SAN since version 6.0.
- It requires a 10Gb Ethernet network; it is not supported with 1Gb NICs.
- The maximum number of all-flash hosts is 64.
- Flash devices are used for both cache and capacity.
- Flash read cache reservation is not used with all-flash configurations.
- There is a need to mark a flash device so it can be used for capacity this is covered in the Virtual SAN Administrators Guide.
- Endurance now becomes an important consideration both for cache and capacity layers.

2.2 Three-node versus Four-node or Greater

While Virtual SAN fully supports 3-node configurations, they can behave differently than configurations with 4 or greater nodes. In particular, in the event of a failure you do not have the ability to rebuild components on another host in the cluster to tolerate another failure. Also with 3-node configurations, you will not have the ability to migrate all data from a node during maintenance. This is because virtual machines on a 3 node cluster cannot be configured to tolerate more than one failure.

If you plan to deploy a 3-node cluster, then that is what you should test. But if you plan on deploying larger clusters, we strongly recommend testing 4 or more nodes.

Further considerations with three-node clusters are covered in the failure testing section of this document.

For the purposes of this proof-of-concept guide, a 4-node configuration is used. However, the cluster size is 3-nodes to begin with, and the fourth node will be added during the course of the proof-of-concept.

2.3 Follow the vSphere Compatibility Guide Precisely

2.3.1 Why Is This Important?

We cannot overstate the importance of following the vSphere Compatibility Guide (VCG) for Virtual SAN to the letter. A significant number of our support requests are ultimately traced back to customers failing to follow these very specific recommendations. This on-line tool is regularly updated to ensure customers always have the latest guidance from VMware available to them.

2.3.2 Hardware, Drivers, and Firmware

The VCG makes very specific recommendations on hardware models for Storage I/O controllers, SSDs, PCI-E flash cards and disk drives. It also specifies which drivers have been fully tested, and – in many cases – identifies the firmware level required. The most direct way to check the controller's firmware version is by interrupting the boot process and looking into the controller's BIOS settings. The <u>VMware Virtual SAN</u> <u>Diagnostics and Troubleshooting Reference Manual</u> contains information about using 'esxcli hardware pci list' and 'vmkload_mod -s' to find the I/O controller's driver version.

2.3.3 Out-of-box Drivers versus Inbox (Shipped with ESXi and Listed on VCG)

Storage controller drivers provided as part of a server vendor's vSphere distribution may or may not be certified for use with Virtual SAN. When in question, go with the driver version specified in the VCG.

Some SSD and flash vendors are revising their firmware frequently, often with significant performance enhancements resulting. Check the VCG regularly for driver and firmware updates.

Although it's well documented, people sometimes forget that Virtual SAN can't claim a disk that already has a partition on it. So make sure to check that your disks are clean before trying to use them with Virtual SAN.

2.3.4 RAID-0 versus Pass-Through for Disks

The VCG will tell you if a controller supports RAID-0 or pass-through when presenting disks to ESXi hosts. RAID-0 is only supported when pass-through is not possible.

Check that you are using the correct configuration and that the configuration is uniform across all nodes.

2.3.5 Controller Configuration

Keep the controller configuration relatively simple. For controller with cache, either disable it, or – if that is not possible - set it to 100% read. For other vendor specific controller features such as HP SSD Smart Path, we recommend disabling them. This may only be possible from the BIOS of the controller in many cases.

2.4 Use Supported vSphere Software Versions

It is highly recommended that anyone who is considering an evaluation of Virtual SAN should pick up the latest versions of software. VMware continuously fixes issues encountered by customers, so by using the latest version of the software, you avoid issues already fixed.

In this proof-of-concept guide, Virtual SAN version 6.1 from vSphere 6.0u1 is used.

3. Virtual SAN POC Setup Assumptions and Prerequisites

The following assumptions are being made with regards to the deployment:

- Four servers are available, and are compliant with the Virtual SAN HCL.
- All servers have had ESXi 6.0u1 deployed. These steps will not be covered in this POC guide.
- A 6.0u1 vCenter Server has been deployed to manage these four ESXi hosts. These steps will not be covered in this POC guide.
- Services such as DHCP, DNS and NTP are available in the environment where the POC is taking place.
- Three out of four ESXi hosts should be placed in a cluster in vCenter.
- If using HP storage controllers, install the *hpssacli* VIB.
- The cluster must <u>not</u> have any features enabled, such as DRS, HA or Virtual SAN. These will be done throughout the course of the POC.
- Each host must have a management network and a vMotion network already configured. There is no Virtual SAN network configured. This will be done as part of the POC.
- For the purposes of testing Storage vMotion operations, an additional datastore type, such as NFS or VMFS, should be presented to all hosts. This is an *optional* POC exercise.
- A set of IP addresses, one per ESXi host will be needed for the Virtual SAN traffic VMkernel ports. The recommendation is that these are all on the same VLAN and network segment.



Figure 3.1: Initial cluster configuration example

From a network perspective, it is optimal to separate the Virtual SAN network from the management and vMotion networks. Below, management, VM and vMotion networks have their own uplinks via VSS (Virtual Standard Switch) vSwitch0.

cs.ie.h01 ie local Action	· · ·		
tting Started Summary	Monitor Manage Related Objects		
ttings Networking Storad	ge Alarm Definitions Tags Permissions]	
		J	
	Virtual switches		
Virtual switches	9 Ge 🔤 🕾 🗸 🗙 🙃		
VMkernel adapters	Switch	Discovered Issues	
Physical adapters	1 vSwitch0	_	
TCP/IP configuration	T vSwitch1	-	
Advanced			
	Standard switch: vSwitch0 (VM Networ	rk)	
	Standard switch: vSwitch0 (VM Networ	rk)	
	Standard switch: vSwitch0 (VM Networ	rk)	
	Standard switch: vSwitch0 (VM Networ	rk)	
	Standard switch: vSwitch0 (VM Networ	rk)	
	Standard switch: vSwitch0 (VM Network	rk)	
	Standard switch: vSwitch0 (VM Network	rk)	
	Standard switch: vSwitch0 (VM Network	rk)	
	Standard switch: vSwitch0 (VM Networ VIN Network VLANID: 51 VLANID: 51 VLANID: 51	rk)	
	Standard switch: vSwitch0 (VM Network	rk)	
	Standard switch: vSwitch0 (VM Network VLAVID: 51 Vitual Machines (0) VLAVID: 51 Vitual Machines (1) VLAVID: 51 VItual Port (1) vmk1 : 10.27.51.31	rk)	
	Standard switch: vSwitch0 (VM Networ X X VI Network VLAN ID: 51 Virtual Machines (0) W Mikernel-vmotion VLAN ID: 51 V Mikernel-Ports (1) vm: 11: 02: 75: 31 W Mikernel-mgmt	rk)	
	Standard switch: vSwitch0 (VM Networ V N Network VLAN ID: 51 VIrtual Machines (0) @ VM kernel-wmoton VLAN ID: 51 @ VM kernel-mgmt VLAN ID: 51	rk)	
	Standard switch: vSwitch0 (VM Network VLAN ID: 51 Virtual Machines (0) VLAN ID: 51 Virtual Machines (0) VLAN ID: 51 VI Mikernel-vmotion VLAN ID: 51 VI Mikernel-Ports (1) Vmkt: 10.27.51.31 VLAN ID: 51 VI Mikernel-Ports (1) VI Mikernel-Ports (1)	rk)	

Figure 3.2: Initial host network configuration example - non-Virtual SAN networks

In this POC example, the Virtual SAN network is on its own VSS (vSwitch1) and the VSS has a number of uplinks in this configuration. You do not need to follow this design and you may use a much simpler "single uplink" VSS in your POC if you wish. In the next section, the steps to create a Virtual SAN VMkernel network interface will be shown.

cs-ie-h01.ie.local Actions	.*		≡*
Getting Started Summary	Ionitor Manage Related Objects		
Settings Networking Storage	Alarm Definitions Tags Permissions		
	Virtual switches		
Virtual switches	2 😥 🕸 🕸 🎽 🖌 🗙 🗿		
VMkernel adapters	Switch	Discovered Issues	
Physical adapters	1 vSwitch0	**	
TCP/IP configuration	1 vSwitch1	-	
Advanced			
	Standard switch: vSwitch1 (VSAN)		
	/ ×		G
	VSAN VLAN ID:- V VIAtement Po vmi2:1723	(1) (1) (20.1) (1)	

Figure 3.3: Initial host network configuration example – Virtual SAN network

It is considered best practice to dedicate 1GbE NICs to the Virtual SAN network. When using 10GbE networks, multiple traffic types may share the same uplink.

VMkernel ad	/Mkernel adapters										
<u>9</u> 😪 🛙	🧕 🏚 🐚 -					, Filter 👻					
Device	Network Label	Switch	IP Address	TCP/IP Stack	vMotion Traffic	Provisioning	FT Logging	Managemen	vSphere Rep	vSphere Rep	Virtual SAN Traffic
🐖 vmk0	🧕 VMkernel-mgmt	T vSwitch0	10.27.51.4	Default	Disabled	Disabled	Disabled	Enabled	Disabled	Disabled	Disabled
🕅 vmk1	👳 VMkernel-vmotion	1 vSwitch0	10.27.51.34	Default	Enabled	Disabled	Disabled	Disabled	Disabled	Disabled	Disabled

Figure 3.4: Initial host network configuration example

If you plan to use distributed switches in your POC, details on how to migrate from a VSS to a distributed switch are shown in appendix B of this guide.

4. Virtual SAN Network Setup

Anyone implementing a Virtual SAN POC should be aware of the <u>VMware Virtual</u> <u>SAN 6.0 Design and Sizing Guide</u>. The guide can be found here:

All ESXi hosts in a Virtual SAN Cluster communicate over a Virtual SAN network. This is a new VMkernel port type introduced in vSphere 5.5 specifically for Virtual SAN. The following example will demonstrate how to configure a Virtual SAN network on an ESXi host.

4.1 Advantages of Distributed Switch versus Standard Switch

If the plan is to test Network I/O Control (NIOC) functionality to provide Quality of Service (QoS) on the Virtual SAN traffic, then a distributed virtual switch (DVS) will be required. If you do not plan to use NIOC, then the evaluation may be done with a standard switch (VSS).

4.2 Creating a VMkernel Port for Virtual SAN

In many deployments, Virtual SAN may be sharing the same uplinks as the management and vMotion traffic, especially when 10GbE NICs are utilized. Later on, we will look at an optional workflow that migrates the standard vSwitches to a distributed switch for the purpose of providing Quality Of Service (QoS) to the Virtual SAN traffic through a feature called Network I/O Control. This is only available on distributed switches.

The Virtual SAN license also includes entitlement to distributed switch, even on the lower editions of vSphere (for use on the Virtual SAN enabled cluster only).

However, the assumption for this POC is that there is already a standard vSwitch created which contains the uplinks that will be used for Virtual SAN traffic. In this example, a separate vSwitch (vSwitch1) with dedicated 1Gbe NICs has been created for Virtual SAN traffic, while the management and vMotion network use different uplinks on a separate standard vSwitch.

To create a Virtual SAN VMkernel port, follow these steps:

Select an ESXi host in the inventory, then navigate to Manage > Networking > VMkernel Adapters. Click on the icon for "Add host networking", as highlighted below:

VMkernel adapters											
(2) 😥 🕼 -						Filter 🔹					
Device	Network Label	Switch	IP Address	TCP/IP Stack	vMotion Traffic	Provisioning	FT Logging	Managemen	vSphere Rep	vSphere Rep	Virtual SAN Traffic
🐖 vmk0	👰 VMkernel-mgmt	T vSwitch0	10.27.51.4	Default	Disabled	Disabled	Disabled	Enabled	Disabled	Disabled	Disabled
🕅 vmk1	🧕 VMkernel-vmotion	T vSwitch0	10.27.51.34	Default	Enabled	Disabled	Disabled	Disabled	Disabled	Disabled	Disabled

Figure 4.1: Add host networking

Ensure that VMkernel Network Adapter is chosen.

cs-ie-h04.ie.local - Add Networ	cs-ie-h04.ie.local - Add Networking				
1 Select connection type 2 Select target device	Select connection type Select a connection type to create.				
3 Connection settings 3a Port properties	VMkernel Network Adapter The VMkernel TCP/IP stack handles traffic for ESXi services such as vSphere vMotion, ISCSI, NFS, FCoE, Fault Tolerance, Virtual SAN and host management.				
3b IPv4 settings 4 Ready to complete	Physical Network Adapter A physical network adapter handles the network traffic to other hosts on the network.				
	Virtual Machine Port Group for a Standard Switch A port group handles the virtual machine traffic on standard switch.				

Figure 4.2: Select VMkernel Network Adapter type

The next step gives you the opportunity to build a new standard vSwitch for the Virtual SAN network traffic. In this example, an already existing vSwitch1 contains the uplinks for the Virtual SAN traffic. If you do not have this already configured in your environment, you can use an already existing switch or select the option to create a new standard vSwitch. When you are limited to 2 x 10GbE uplinks, it makes sense to use the same VSS. When you have many uplinks, some dedicated to different traffic types (as in this example), management can be a little easier if different VSS with their own uplinks are used for the different traffic types.

As there is an existing vSwitch in our environment that contains the network uplinks for the Virtual SAN traffic, the "browse" button is used to select it as shown below.

	cs-ie-h04.ie.local - Add Networking					
~	1 Select connection type	Select target device Select a target device for the new connection.				
	2 Select target device					
	3 Connection settings	 Select an existing standard switch 				
	3a Port properties	Browse				
	3b IPv4 settings					
	4 Ready to complete	New standard switch				



Select Switch	×
Select a standard switch on which to create a VMkernel	adapter.
Switch	Discovered Issi
1 vSwitch0	
1 vSwitch1	
4	Þ
ок	Cancel

Figure 4.4: Choose a vSwitch

Ē	cs-ie-h04.ie.local - Add Networking					
~	1 Select connection type	Select target device				
	2 Select target device					
	3 Connection settings	 Select an existing standard switch 				
	3a Port properties	vSwitch1 Browse				
	3b IPv4 settings	New standard switch				
	4 Ready to complete	View standard switch				

Figure 4.5: vSwitch is displayed once selected

The next step is to setup the VMkernel port properties, and choose the services, such as Virtual SAN traffic. This is what the initial port properties window looks like.

	Dort proportion		
Select connection type	Specify VMkernel port setting	15.	
Select target device		•	
Connection settings	VMkernel port settings		
3a Port properties 3b IPv4 settings	Network label:	VMkernel	
Ready to complete	VLAN ID:	None (0) 💌	
	IP settings:	IPv4 🔹	
	TCP/IP stack:	Default 🔹 🕤	
	Available services		
	Enable services:	VMotion traffic	
		Provisioning traffic	
		Fault Tolerance logging	
		Management traffic	
		vSphere Replication traffic	
		vSphere Replication NFC traffic	
		Virtual SAN traffic	
		Back Next	Finish Can

Figure 4.6: Default port properties

Here is what it looks like when populated with Virtual SAN specific information.

cs-ie-h04.ie.local - Add Network	ing		?
 1 Select connection type 2 Select target device 3 Connection settings 3a Port properties 3b IPv4 settings 4 Ready to complete 	Port properties Specify Wikernel port settings. VMkernel port settings Network label: VLAN ID: IP settings: TCP/IP stack:	VSAN None (0) IPv4 Default	
	Avallable services	VMotion traffic Provisioning traffic Fault Tolerance logging Management traffic Vophere Replication traffic Vophere Replication NFC traffic Vophere Replication NFC traffic Vophere Replication NFC traffic	
		Back Next Finish Cano	el

Figure 4.7: Port properties configured for Virtual SAN traffic

In the above example, the network label has been designated "Virtual SAN", and the Virtual SAN traffic does not run over a VLAN. If there is a VLAN used for the Virtual SAN traffic in your POC, change this from "None (0)" to an appropriate VLAN ID.

The next step is to provide an IP address and subnet mask for the Virtual SAN VMkernel interface. As per the assumptions and pre-requisites section earlier, you should have these available before you start. At this point, you simply add them, one per host by clicking on "Use static IPv4 settings" as shown below. Alternatively, if you plan on using DHCP IP addresses, leave the default setting which is "Obtain IPv4 settings automatically".

	cs-ie-h04.ie.local - Add Networking							
> >	1 Select connection type 2 Select target device	IPv4 settings Specify VMkernel IPv4 settings.						
	3 Connection settings	Obtain IPv4 settings automatically						
~	3a Port properties	 Use static IPv4 settings 						
	3b IPv4 settings	IPv4 address:	172.32.0.4					
	4 Ready to complete	Subnet mask:	255 . 255 . 255 . 0					
		Default gateway for IPv4:	10.27.51.254					
		DNS server addresses:	10.27.51.252					

Figure 4.8: IP address and subnet mask

The final window is a review window. Here you can check that everything is as per the options selected throughout the wizard. If anything is incorrect, you can navigate back through the wizard. If everything looks like it is correct, you can click on the "Finish" button.

cs-ie-h04.ie.local - Add Networkir	ng			?
 1 Select connection type 2 Select target device 	Ready to complete Review your settings selections befo	re finishing the wizard.		
3 Connection settings	Standard switch:	vSwitch1		
✓ 3a Port properties	New port group:	VSAN		
 3b IPv4 settings 	VLAN ID:	51		
	TCP/IP stack:	Default		
4 Ready to complete	vMotion traffic:	Disabled		
	Provisioning traffic:	Disabled		
	Fault Tolerance logging:	Disabled		
	Management traffic:	Disabled		
	vSphere Replication traffic:	Disabled		
	vSphere Replication NFC traffic:	Disabled		
	Virtual SAN traffic:	Enabled		
	IPv4 settings			
	IPv4 address:	172.32.0.4 (static)		
	Subnet mask:	255.255.255.0		
			Back Next Finish Cance	а.

Figure 4.9: Review window

If the creation of the VMkernel port is successful, it will appear in the list of VMkernel ports, as shown below.

VMkernel ad	/Mkernel adapters										
👲 🔂 🛛	🙎 😥 🗎 🔹										
Device	Network Label	Switch	IP Address	TCP/IP Stack	vMotion Traffic	Provisioning	FT Logging	Managemen	vSphere Rep	vSphere Rep	Virtual SAN Traffic
뺊 vmk0	👰 VMkernel-mgmt	1 vSwitch0	10.27.51.4	Default	Disabled	Disabled	Disabled	Enabled	Disabled	Disabled	Disabled
ị vmk1	🧕 VMkernel-vmotion	1 vSwitch0	10.27.51.34	Default	Enabled	Disabled	Disabled	Disabled	Disabled	Disabled	Disabled
🐖 vmk2	👷 VSAN	TvSwitch1	172.32.0.4	Default	Disabled	Disabled	Disabled	Disabled	Disabled	Disabled	Enabled

Figure 4.10: VMkernel adapters with new Virtual SAN VMkernel adapter

That completes the Virtual SAN networking setup for that host. You must now repeat this for all other ESXi hosts, including the host that is not currently in the cluster you will use for Virtual SAN.

If you wish to use a DVS (distributed vSwitch), the steps to migrate from standard vSwitch (VSS) to DVS are documented in Appendix B—Migrating from Standard vSwitch to Distributed of this POC Guide.

5. Enabling Virtual SAN on the Cluster

Enabling Virtual SAN is quite simple, and can be done in just a few clicks in the vSphere web client. However, there is one decision that needs to be made when enabling Virtual SAN, and that is whether you want Virtual SAN to claim all of the unused local storage on the ESXi hosts, or if you (as the administrator) wish to decide which physical disks and flash devices to use for the Virtual SAN datastore.

To enable Virtual SAN, select the cluster object in the vCenter inventory, then select the Manage tab > Settings > Virtual SAN > General, as shown below.



Figure 5.1: Virtual SAN is Turned OFF

To turn on (or enable) Virtual SAN, there is an "Edit" button located in the top right as shown below.

VSAN6-Cluster Actions •	•	=				
Getting Started Summary	Monitor Manage Related Objects					
Settings Scheduled Tasks Alarm Definitions Tags Permissions						
Virtual SAN is Turned OFF						
✓ Services	Add disks to storage Manual					
vSphere DRS						
vSphere HA	Resources					
- Virtual SAN	Hosts	0 hosts				
General	Flash disks in use	0 of 0 eligible				
Disk Management	Data disks in use	0 of 0 eligible				
Fault Domains	Total capacity of Virtual SAN datastore	0.00 B				
Health	Free capacity of Virtual SAN datastore	0.00 B				
- Configuration	Network status	② No hosts in the cluster				
Gonoral						

Figure 5.2: Virtual SAN Edit button

Simply click on this "Edit" button to start the process of enabling Virtual SAN. This opens the following pop-up, which provides the option to turn on Virtual SAN, and then add disks to storage manually or automatically.

🚯 VSAN6-Cluster - Edi	t Virtual SAN Settings	?
Turn ON Virtual S	AN	
Add disks to storage	Manual 👻	
	Requires manual claiming of any new disks on the included hosts to the shared storage.	
Licensing	A license must be assigned to the cluster in order to create disk groups or consume disks automatically.	
	OK Cance	

Figure 5.3: Edit Virtual SAN Settings

When "Turn ON Virtual SAN" option is checked, the option to select a manual or automatic option is available for selection.

🚯 VSAN6-Cluster - Edit	Virtual SAN Settings	?
☑ Turn ON Virtual SA	AN	
Add disks to storage	Manual Image: Second seco	I
Licensing	A license must be assigned to the cluster in order to create disk groups or consume disks automatically.	
	OK Cance	

Figure 5.4: Add disks to storage

If networking is correctly configured, and each of the ESXi hosts can communicate, then the Virtual SAN Cluster will form. In the following example, manual disk claiming is chosen so as to provide for learning more about Virtual SAN disk groups during the POC.

5.1 Manual Disk Claiming—Create Disk Groups

At present, there are three hosts in the Virtual SAN Cluster. However, because the cluster is in manual mode, no flash devices or disks have been claimed. A General view of Virtual SAN currently looks something like this.

1	VSAN6-Cluster Actions -			=*				
G	etting Started Summary Mo	nitor Manage Related Objects						
5	Settings Scheduled Tasks Alarm Definitions Tags Permissions							
•	н	Virtual SAN is Turned ON	Edit.					
	 Services 	Add disks to storage Manual		5				
	vSphere DRS vSphere HA	Resources						
	Virtual SAN	Hosts	3 hosts					
	General	Flash disks in use	0 of 3 eligible					
	Disk Management	Data disks in use	0 of 19 eligible					
	Fault Domains	Total capacity of Virtual SAN datastore	0.00 B					
	Health	Free capacity of Virtual SAN datastore	0.00 B					
	- Configuration	Network status	✓ Normal					
	General Licensing	On-disk Format Version	Upgra	de				
	VMware EVC	Virtual SAN format version 🛕 -						
	VM/Host Groups	Disks with older version 💧 0 of 0						
	VM/Host Rules VM Overrides	Support Assistant	Upload Support Bundles to Service Request	it				
	Host Options	Last upload time -						
	Profiles							

Figure 5.5: Virtual SAN enabled, no disks or flash device claimed

Any warnings against Virtual SAN format version and Disks with older version can be ignored for the moment. These appear as a result of the Health checks, but since there are no disks in the cluster as yet, these warning are displayed.

The next step is to claim some storage and flash devices for Virtual SAN and create the disk groups.

Navigate to Disk Management, just below the General view. You should observe that there are no disk groups associated with the hosts, nor are there any disks in use.

There are a number of icons here related to the claiming of disk groups that require further explaining:

W	This allows you to build disks groups across all hosts in one step. Useful for small clusters, but cumbersome when lots of hosts and disks present
	Create a new disk group on a per host basis (visible when disk group selected)

Table 5.1: Disk group icons

For this POC, one flash device and two magnetic disks (HDD) are chosen for the disk group. This is repeated for all three hosts. Of course, you may wish to include additional devices in your POC.

As mentioned, since this is a small cluster, we are only going to create a disk group containing two physical disks for capacity. Either of the icons shown above can be chosen. In this example POC, the first icon shown above can be chosen.

This immediately pops up an option to "Select all eligible disks". We are <u>not</u> choosing this option in the POC, but it is a useful option to be aware of. If you are including all disks in all hosts, then you may certainly choose this option to speed things along.

Similarly, if one clicks on the check box next to a hostname, all disks belonging to that host will be used for creating disk groups. This is also <u>not</u> a feature we wish to use in the POC either, but once again a useful option to be aware of. If you wish to select all disks on a particular host for your POC, you may choose this option.

🚯 VSAN6-Cluster - Claim Disks for Virtual SAN Use					?
When claiming disks for Virtual SAN, the following rules apply: 1. Hosts that contribute to the Virtual SAN cluster must contribute at least 4 2. The number of HDD disks that a host contributes must be greater than More flash disks means better performance, and more HDD disks mean	1 flash disk. or equal to the more datastor	e number c e capacity.	f flash disks	it contributes.	
Select all eligible disks	Show:	Eligible I	nosts	• Q Filter	•
Name			Drive Type	Capacity	Transport Type
👻 🔲 🔋 cs-ie-h03.ie.local					
🗹 🚐 HP Serial Attached SCSI Disk (naa.600508b1001ceefc4	4213ceb9b51c	c4be4)	HDD	136.70 GB	Block Adapt
🗹 🚐 HP Serial Attached SCSI Disk (naa.600508b1001cd259	ab7ef213c87	eaad7)	HDD	136.70 GB	Block Adapt ::
🗹 📧 HP Serial Attached SCSI Disk (naa.600508b1001c9c8b	5f6f0d7a2be4	4433)	Flash	186.28 GB	Block Adapt
🗌 🚐 HP Serial Attached SCSI Disk (naa.600508b1001c2b7a	3d39534ac6b	eb92d)	HDD	136.70 GB	Block Adapt
🔲 🚐 HP Serial Attached SCSI Disk (naa.600508b1001cb11f.	3292fe743a0f	d2e7)	HDD	136.70 GB	Block Adapt
🗌 🚐 HP Serial Attached SCSI Disk (naa.600508b1001c1a7f.	310269ccd51a	a4e83)	HDD	136.70 GB	Block Adapt
🔲 🚐 HP Serial Attached SCSI Disk (naa.600508b1001c9b93	053e6dc3ea9	bf3ef)	HDD	136.70 GB	Block Adapt
👻 🔲 📋 cs-ie-h01.ie.local					
🗹 🚐 HP Serial Attached SCSI Disk (naa.600508b1001c16be	6e256767284	4eaf88)	HDD	136.70 GB	Block Adapt
🗹 🚐 HP Serial Attached SCSI Disk (naa.600508b1001c2aec	3f0f34d30235	fba8)	HDD	136.70 GB	Block Adapt
🔲 🚐 HP Serial Attached SCSI Disk (naa.600508b1001c6481	6271482a56a	a48c3c)	HDD	136.70 GB	Block Adapt 👻
A6					25 items 📑 🗸
=					
Configuration validation:					
✓ Configuration correct.					
				C	K Cancel

Figure 5.6: Claiming disks for disk groups

Note that there will be warnings posted if a flash device and magnetic disk devices are not chosen, since a disk group requires one flash device and at least one magnetic disk in hybrid configurations, which is what we are working on here. Click on the OK button to complete the configuration.

Once the configuration task completes, the Disk Management view should now be updated with a disk group per host added, as well as the "Disks in Use" column populated with the number of disks in use in the disk groups, which should be three (one flash device and two magnetic disks). The view should look similar to the following, although the total number of disks will vary depending on your POC.

VSAN6-Cluster Actions 👻								=*
Getting Started Summary Mo	nitor Manage Related Objects							
Settings Scheduled Tasks Ala	rm Definitions Tags Permissions							
44	Disk Groups							
- Services	a 📷 🕵 🐁 🥍 🍹 🖼						Q Filter	•
vSphere DRS	Disk Group	Disks in Use	State	Virtual SAN	Fault Domain	Network Parti	Disk Format Version	
vSphere HA	✓ a cs-ie-h03.ie.local	3 of 7	Connecte	ed Healthy		Group 1		
Virtual SAN	Disk group (0200080000600508b1001c9c8b5f6f0d7a2be	3		Healthy			2	
General	👻 🗐 cs-ie-h01.ie.local	3 of 8	Connecte	ed Healthy		Group 1		
Disk Management	Disk group (0200080000600508b1001c61cedd42b0c3fbf5	3		Healthy			2	
Fault Domains	✓ □ cs-ie-h02.ie.local	3 of 7	Connecte	ed Healthy		Group 1		
Health	Disk group (0200080000600508b1001c64b76c8ceb56e8	3		Healthy			2	
- Configuration								
General								
Licensing								
VMware EVC	A4.						6 items	- -
VM/Host Groups			_					
VM/Host Rules	cs-ie-h03.ie.local: Disks							
VM Overrides							Show: In use (3)	•
Host Options	Name	Drive Type	Capacity	Virtual SAN Health Status	Operational	Transport Type	1	
Profiles	E HP Serial Attached SCSI Disk (naa.600508b1001c9c8b5f6f0d7a	Flash	186.28 GB	Healthy	Mounted	Block Adapte	er	
	HP Serial Attached SCSI Disk (naa.600508b1001ceefc4213ceb9	HDD	136.70 GB	Healthy	Mounted	Block Adapte	er	
	HP Serial Attached SCSI Disk (naa.600508b1001cd259ab7ef213	HDD	136.70 GB	Healthy	Mounted	Block Adapte	er	

Figure 5.7: Disk groups created

Returning to the General view (and possibly refreshing the screen) should now show the number of flash disks in use (three, one per host) and data disks (six, two per host) that are now in use. It should also show the total capacity of the Virtual SAN datastore, which in this case is \sim 812GB. That is 6 x 136GB, less some overhead. Remember that flash devices do not contribute towards capacity, only the magnetic disk devices (in the case of hybrid configurations).

VSAN6-Cluster	Actions 👻				
Getting Started Su	ummary Moni	itor Manage	Related O	bjects	
Settings Schedule	ed Tasks Alarr	m Definitions	ags Perm	nissions	
44		Virtual SAN is T	urned ON		
 Services 		Add disks to s	torage N	/anual	
vSphere DRS vSphere HA		Resources			
▼ Virtual SAN		Hosts			0 hosts
General		Flash disks in	use		0 of 0 eligible
Disk Manageme	ent	Data disks in u	ise		0 of 0 eligible
Fault Domains		Total capacity	of Virtual SA	N datastore	0.00 B
Health		Free capacity (of Virtual SA	N datastore	0.00 B
 Configuration 		Network status	6		🗸 Normal
General Licensing		On-disk Format	Version		
VMware EVC		Virtual SAN for	mat versior	Version 2	(latest)
VM/Host Groups	i	Disks with old	erversion	🥑 0 of 9	
VM/Host Rules VM Overrides		Support Assist	ant		
Host Options		Last upload tir	ne -		
D					

Figure 5.8: On-disk Format Version

6. Enable the Virtual SAN Health Check Plugin

Following on from the Virtual SAN 6.0 GA release, a new feature called Health Check plugin was released. This gives administrators valuable information regarding the state of the Virtual SAN Cluster, and is also extremely useful for POC activities as it quickly discovers issues.

There is an in-depth description of health checks, including how to install and configure it, as well as detailed information on the various checks that it carries out. Refer to the <u>VMware Virtual SAN Health Check Plugin Guide</u>.

Starting with vSphere 6.0 update 1 and vCenter 6.0 update 1, the Health Check plugin is pre-installed both in vCenter and as a VIB on each ESXi host. All that's required is to enable the health check services once Virtual SAN is enabled. This is done on a cluster-by-cluster basis at the cluster's Manage tab > Settings > Virtual SAN > Health. Once enabled, the new health check service in Virtual SAN 6.1 runs hourly by default and on-demand when you visit the cluster's Monitor tab > Virtual SAN > Health.

ug⊮ C1 Actions ₹		=
Getting Started Summary Mo	nitor Manage Related Objects	
Settings Scheduled Tasks Ala	rm Definitions Tags Permissions	
	Health Service Edit set	tings
✓ Services	Health service status 0 Enabled	
vSphere DRS	Health service version 6.1.0	
VSphere HA	Health check interval 60 minutes	
General	HCL Database Update from file Get latest version	online
Disk Management	Last updated Today	
Fault Domains	Support Assistant Upload Support Bundles to Sonice Bo	quest
Health	Opioad Support Bundles to Service Re-	quest
- Configuration	Last upload time	
General	External Proxy Settings	Edit
Licensing VMware EVC VM/Host Groups	Configure the proxy to access the internet when you use Virtual SAN CEIP(Customer Experience Improvement Program), Virtual SAN Support Assistant and get latest HCI database online.	
VM/Host Rules	Host name	
VM Overrides	Port -	
Host Options	User name	
Profiles	Proxy setting type -	

Figure 6.1: Managing Virtual SAN health check service

With the Virtual SAN -enabled cluster object selected in the inventory, navigate to the Monitor tab > Virtual SAN > Health. This will display the list of health check, and their status. Hopefully everything will show up as passed as per figure 6.2 below.

VSAN6-Cluster Actions -			E.
Getting Started Summary Mo	onitor Manage Re	elated Objects	
Issues Profile Compliance Pe	erformance Tasks I	Events Resource Reservation Vinual SAN vSphere DRS Utilization	
	Virtual SAN Health	(Last checked: 07 April 2015 13:16:43)	Retest
Health	Test Result	Test Name	
Proactive Tests	Passed	Cluster health	
Physical Disks	Passed	➤ Data health	
Virtual Disks	Passed	 Limits health 	
Resyncing Components	Passed	 Network health 	
	Passed	 Physical disk health 	
	Passed	 VSAN HCL health 	
	A4		6 items 🕞 🗧

Figure 6.2: Top level list of health checks

6.1 Check Your Network Thoroughly

Once the Virtual SAN network has been created, and Virtual SAN is enabled, you should check that each ESXi host in the Virtual SAN Cluster is able communicate to all other ESXi hosts in the cluster. The easiest way to achieve this is via the Health Check Plugin.

6.1.1 Why Is This Important?

Virtual SAN is entirely dependent on the network: its configuration, reliability, performance, etc. One of the most frequent causes of requesting support is either an incorrect network configuration, or the network not performing as expected.

6.1.2 Check the Network Partition Groups after Creating Cluster

A network partition is when a subset of hosts (one or more) in unable to communicate to another subset of hosts. The Disk Management view (found under Virtual SAN Cluster > Manage tab > Settings) provides immediate information about whether or not there is a network partition in your cluster. If the network is functioning properly, all hosts will be in Group 1. Only if multicast routing is properly configured would Virtual SAN still function with multiple partition groups. Refer to the Network health tests under Cluster > Monitor > Virtual SAN > Health.

Ę	VSAN6-Cluster Actions -							
1	Getting Started Summary Mor	nitor Manage Related Objects						
	Settings Scheduled Tasks Alar	m Definitions Tags Permissions						
	••	Disk Groups						
	✓ Services	@)- TE C						
	vSphere DRS	Disk Group	Disks in Use	State	Virtual SAN	Fault Domain	Network Partition Group	Disk Format Version
	vSphere HA		3 of 7	Connected	Healthy		Group 1	
		Disk group (0200080000600508b100	3		Healthy			2
	General		3 of 8	Connected	Healthy		Group 1	
	Disk Management	📓 Disk group (0200080000600508b100	3		Healthy			2
	Fault Domains	✓ ☐ cs-ie-h02.ie.local	3 of 7	Connected	Healthy		Group 1	
	Health	Disk group (0200080000600508b100	3		Healthy			2

Figure 6.3: Network Partition Group info

6.1.3 Use the Health Check Plugin to Verify Virtual SAN Functionality

Running individual commands from one host to all other hosts in the cluster can be tedious and time consuming. Fortunately, since Virtual SAN 6.0 supports a new health check plugin, part of which tests the network connectivity between all hosts in the cluster. If for some reason the cluster will not form, and displays a "Network misconfiguration" in the General view, you should proceed with enabling the health check plugin, outlined in the previous section. This will reduce the time to detect and resolve the networking issue, or any other Virtual SAN misconfiguration issues in the cluster.

In the screenshot below, one can see that each of the health checks for networking has successfully passed.

VSAN6-Cluster Actions -			
etting Started Summary	Nonitor Manage F	Related Objects	
sues Profile Compliance F	Performance Tasks	Events Resource Reservation Virtual SAN vSphere DRS Utilization	
	Virtual SAN Health	h (Last checked: 09 April 2015 07:59:42)	Retest
Health	Test Result	Test Name	
Proactive Tests	Passed	Cluster health	
Physical Disks	Passed	 Data health 	
Virtual Disks	Passed	Limits health	
esyncing Components	📀 Passed	 Network health 	
	Passed	All hosts have a VSAN vmknic configured	
	Passed	All hosts have matching multicast settings	
	Passed	All hosts have matching subnets	
	Passed	Basic (unicast) connectivity check (normal ping)	
	Passed	Hosts disconnected from VC	
	Passed	Hosts with connectivity issues	
	Passed	Hosts with VSAN disabled	
	Passed	MTU check (ping with large packet size)	
	Passed	Multicast assessment based on other checks	
	Passed	Unexpected VSAN cluster members	
	Passed	VSAN cluster partition	
	Passed	 Physical disk health 	
	44		17 items 🔒

Figure 6.4: Network health checks all passed

If any of the network health checks fail, select the appropriate check and examine the details screen below for details on how to resolve the issue. Each details view also contains an *AskVMware* button where appropriate, which will take you to a VMware Knowledge Base article detailing the issue, and how to troubleshoot and resolve it.

For example, in this case where one host does not have a Virtual SAN vmknic configured, this is what is displayed.

Virtual SAN Health (La	ist checked: 09 April 2015 09:12:09)	Retest
Test Result	Test Name	
🙆 Failed	▶ Data health	<u>ـ</u>
😣 Failed	✓ Network health	
🔕 Failed	All hosts have a VSAN vmknic configured	
😣 Failed	All hosts have matching multicast settings	
😣 Failed	VSAN cluster partition	::
🔥 Warning	All hosts have matching subnets	
Passed	Basic (unicast) connectivity check (normal ping)	
📀 Passed	Hosts disconnected from VC	
Passed	Hosts with connectivity issues	
Passed	Hosts with VSAN disabled	
Passed	MTU check (ping with large packet size)	
Passed	Multicast assessment based on other checks	Ŧ
M		17 items 🔒 🗸
All hosts have a VSAN	I vmknic configured	Ask VMware
Check if all hosts in a	VSAN cluster have configured VSAN traffic 🚯	
Hosts with no VSAN vi	nknic present	
Host		
cs-ie-h01.ie.local		

Figure 6.5: Network health failure example

Before going any further with this POC, it is worth downloading the latest version of the HCL database and running a "Retest" on the Health check screen. This will ensure everything in the cluster is optimal. It will also check the hardware against the VMware Compatibility Guide (VCG) for Virtual SAN, verify that the networking is functional, and that there are no underlying disk problems. All going well, after the Retest, everything should still display a "Passed" status.

Content Actions -	Actions *					
Getting Started Summary Me	onitor Manage Re	vlated Objects				
Issues Profile Compliance Pe	erformance Tasks E	Events Resource Reservation Virtual SAN vSphere DRS Utilization				
44	Virtual SAN Health (Last checked: 07 April 2015 13:16:43)	Retest			
Health	Test Result	Test Name				
Proactive Tests	Passed	Cluster health				
Physical Disks	Passed	 Data health 				
Virtual Disks	Passed	Limits health				
Resyncing Components	Passed	➤ Network health				
	Passed	Physical disk health				
	Passed	VSAN HCL health				
	M		6 items 🕒 🗧			

Figure 6.6: Virtual SAN Health checks

In particular at this juncture the Cluster health, Limits health and Physical disk health should be examined. The data health only becomes relevant once you start to deploy virtual machines to the Virtual SAN datastore.

Virtual SAN Health	(Last checked: 09 April 2015 09:22:02)	Retest
Test Result	Test Name	
📀 Passed	✓ Cluster health	A
🥑 Passed	Advanced Virtual SAN configuration in sync	
🥑 Passed	ESX VSAN Health service installation	
🥑 Passed	VSAN CLOMD liveness	
🥑 Passed	VSAN Health Service up-to-date	
🥑 Passed	 Data health 	
🥑 Passed	✓ Limits health	
🥑 Passed	After 1 additional host failure	
🥑 Passed	Current cluster situation	
🥑 Passed	 Network health 	
🥑 Passed	 Physical disk health 	
🥑 Passed	Component metadata health	
🥑 Passed	Congestion	
🥑 Passed	Disk capacity	
🥑 Passed	Memory pools (heaps)	
🥑 Passed	Memory pools (slabs)	
📀 Passed	Metadata health	
📀 Passed	Overall disks health	
🕑 Passed	Software state health	-
M		20 items 📑 🗸

Figure 6.7: Expanded Health check plugin Checks

6.1.4 Use the Troubleshooting Reference Manual to Verify Network Functionality

If you need to delve deeper into troubleshooting the network, there are a number of commands available for testing network connectivity between Virtual SAN hosts. These include *vmkping* and *tcpdump-uw*. How to use these commands, and the different parts of Virtual SAN functionality that they test, are outlined in the <u>VMware</u> <u>Virtual SAN Diagnostics and Troubleshooting Reference Manual</u>.

Virtual SAN, including the Health check monitoring, is now successfully deployed. The remainder of this POC guide will involve various tests and error injections to show how Virtual SAN will behave under these circumstances.

7. vSphere Functionality on Virtual SAN

This initial test is per VM testing, and will highlight the fact that general virtual machine operations are unchanged in Virtual SAN environments.

7.1 Deploy Your First VM

In this section, a VM is deployed to the Virtual SAN datastore using the default storage policy. This default policy is preconfigured and does not require any intervention unless you wish to change the default settings, which we do not recommend.

To examine the default policy settings, navigate to Home > VM Storage Policies.



Figure 7.1: VM Storage Policies

From there, select Virtual SAN Default Storage Policy and then select the Manage tab. Under the Manage tab, select **Rule-Set 1: Virtual SAN** to see the settings on the policy:

mware [®] vSphere Web Client	ħΞ		7		
Navigator I	F Virtual SAN Default Store	age Policy	Actions -		
VM Storage Policies	Getting Started Summary	Monitor	Manage	Related	d Objects
🛜 Virtual SAN Default Storage Policy					
🗗 Virtual Machines 🗾	Name and description	ule-Set 1: VS	AN		
VM Templates in Folders	Rule-Set 1: VSAN	Number of	failures to	tolerate	1
		Number of object	disk stripe	s per	1
		Force provi	isioning		No
		Object spa	ce reservat	ion (%)	0
		Flash read (%)	cache rese	ervation	0.0000

Figure 7.2: Rule-Set 1: Virtual SAN (default policy)

We will return to VM Storage Policies in more detail in a future chapter, but suffice to say that when a VM is deployed with the default policy, it should have a mirror copy of the VM data created. This second copy of the VM data is placed on storage on a different host to enable the VM to tolerate any single failure. Also note that object space reservation is set to 0%, meaning the object should be deployed as "thin". After we have deployed the VM, we will verify that Virtual SAN adheres to both of these capabilities.

One final item to check before we deploy the VM is the current free capacity on the Virtual SAN datastore. This can be viewed from the Virtual SAN Cluster > Manage tab > Settings > General view. In this POC, it is 811.95 GB.

Virtual SAN is Turned ON						
Add disks to storage Manual						
Resources						
Hosts	3 hosts					
Flash disks in use	3 of 3 eligible					
Data disks in use	6 of 19 eligible					
Total capacity of Virtual SAN datastore	811.95 GB					
Free capacity of Virtual SAN datastore	811.95 GB					
Network status	 Normal 					

Figure 7.3: Current free capacity of Virtual SAN datastore

Make a note of the free capacity on your POC before continuing with the deploy VM exercise.

To deploy the VM, simply follow the steps provided in the wizard.



Figure 7.4: New Virtual Machine

🔁 New Virtual Machine		(3) >>
1 Select creation type	Select a creation type How would you like to create a virtual machine?	
2 Edit settings	Create a new virtual machine	This option guides you through creating a new virtual
2a Select a name and folder	Deploy from template	machine. You will be able to customize processors, memory,
2b Select a compute resource	Clone an existing virtual machine	guest operating system after creation.
2c Select storage	Clone virtual machine to template	
2d Select compatibility	Clone template to template	
2e Select a guest OS	Convert template to virtual machine	
2f Customize hardware		
3 Ready to complete		
		Back Next Finish Cancel

Figure 7.5: Create a new virtual machine

At this point a name for the VM must be provided, and then the Virtual SAN Cluster must be selected as a compute resource.

New Virtual Machine		?
1 Select creation type 1a Select a creation type	Select a name and folder Specify a unique name and target location	
2 Edit settings	Enter a name for the virtual machine.	
2a Select a name and folder	vsan6-poc-test-vm-1	
2b Select a compute resource	Virtual machine names can contain up to 80 characters and they must be unique within each vCenter Server VM folder.	
2c Select storage	Select a location for the virtual machine.	
2d Select compatibility	Q Search	
2e Select a guest OS	- Ø ie.vrsa-09 ie Incel	
2f Customize hardware	▷ 🖥 VSAN6-DC	
3 Ready to complete	Select a datacenter or VM folder to create the new virtual machine in.	I

Figure 7.6: Select a name and folder

1 New Virtual Machine		? >>
 Select creation type 1a Select a creation type 2 Edit settings 2a Select a name and folder 2b Select a compute resource 2c Select storage 2d Select compatibility 2e Select a guest OS 2f Customize hardware 3 Ready to complete 	Select a compute resource Select the destination compute resource for this operation Q Search VSAN6-DC VSAN6-Cluster Q Cs-le-h04.ie.local Select a cluster, host, vApp or resource pool to run this virtual machine.	
	Compatibility: Compatibility checks succeeded. Back Next Finish C:	ancel

Figure 7.7: Select a compute resource

At this point, the virtual machine deployment process is almost identical to all other virtual machine deployments that you have done on other storage types. It is the next section that might be new to you. This is where a policy for the virtual machine is chosen.

From the next menu, you can either select the Virtual SAN datastore, and the "Datastore Default" policy will actually point to the "Virtual SAN Default Storage Policy" seen earlier.

🔁 New Virtual Machine							?)			
 Select creation type 1a Select a creation type 	Select storage Select the datastore	in which to sto	re the configuratior	and disk files						
2 Edit settings	VM Storage Policy:	Datastore Def	ault	• 0						
 2a Select a name and folder 	The following datas	Datastore Defa	ault	burce that	t vou selected. Selec	t the destination d	atastore for the			
 2b Select a compute resource 	virtual machine con	Wol No Requi	rements Policy							
✓ 2c Select storage	Name	Virtual SAN De	fault Storage Policy	/ ied	Free	Туре	Storage DRS			
2d Select compatibility	NFS-		50.46 TB	379.19 GB	50.19 TB	NFS v3				
2e Selecta guestOS	🗐 vsanDatastore		811.95 GB	75.95 GB	736.00 GB	vsan				
2f Customize hardware	corkisos.csl.vr	nware.com	1.77 TB	1.61 TB	170.62 GB	NFS v3				
3 Ready to complete	🗐 cs-ie-h02-scra	itch	136.50 GB	3.15 GB	133.35 GB	VMFS				
,,	🗐 cs-ie-h03-scra	itch	136.50 GB	3.21 GB	133.29 GB	VMFS				
	4						•			
	Compatibility:									
	Compatibility	checks succeed	ded.							
					Back	Next Finis	sh Cancel			

Figure 7.8: Select the Virtual SAN Default Storage Policy

Once the policy has been chosen, datastores are split into those that are compliant with the policy, and those that are non-compliant with the policy. As seen below, only the Virtual SAN datastore can understand the policy settings in the Virtual SAN Default Storage Policy so it is the only one that shows up as **Compatible** in the list of datastores.

1 Select creation type 1a Select a creation type	Select storage Select the datastore in which to store the configuration and disk files							
2 Edit settings	Mil Storago Dolicy Netwol CAN Default Storago Dolicy							
2a Select a name and folder	The following detections are as	elaun otorage i on						
2b Select a compute resource	virtual machine configuration files	and all of the virtua	il disks.	it you selected. Select	ule desultation d	atastore for the		
2c Select storage	Name	Capacity	Provisioned	Free	Type	Storage DR		
2d Select compatibility	Compatible							
2e Selecta guestOS	🗐 vsanDatastore	811.95 GB	0.00 B	811.95 GB	vsan			
2f Customize hardware	Incompatible							
3. Posdu to complete	NFS-Isilon	50.46 TB	379.14 GB	50.19 TB	NFS v3			
5 Ready to complete	corkisos.csl.vmware.com	1.77 TB	1.61 TB	170.62 GB	NFS v3			
	s-ie-h02-scratch	136.50 GB	3.06 GB	133.44 GB	VMFS			
	cs-ie-h03-scratch	136.50 GB	3.12 GB	133.38 GB	VMFS			
	4							
	Compatibility:							
	Compatibility checks succes	ded.						

Figure 7.9: vsanDatastore is compatible with Virtual SAN Default Storage Policy

The rest of the VM deployment steps in the wizard are quite straightforward, and simply entail selecting ESXi version compatibility (leave at default), a guest OS (leave at default) and customize hardware (no changes). Essentially you can click through the remaining wizard screens without making any changes.

😰 New Virtual Machine			(1)
 Select creation type 1a Select a creation type 	Select compatibilit	ty for this virtual machine depending on the hosts in your environment	
2 Edit settings	The host or cluste	r supports more than one VMware virtual machine version. Select a compatibility for the virtual machine.	
 2a Select a name and folder 			
✓ 2b Select a compute resource	Compatible with:	ESXi 6.0 and later	
✓ 2c Select storage	This virtual machin	e uses bardware version 11 and provides the best performance and latest features	
 2d Select compatibility 	available in ESXi 6	.0.	
2e Select a guest OS			
2f Customize hardware			
3 Ready to complete			
		Back Next Finish C	ancel

Figure 7.10: Select the ESXi compatibility (click next)

🔁 New Virtual Machine				? ₩
Select creation type 1a Select a creation type	Select a guest OS Choose the guest O	S that will be installed on the virtual machine		
2 Edit settings ✓ 2a Select a name and folder	Identifying the guest installation.	t operating system here allows the wizard to provide the appropriate d	efaults for the operating system	
 2b Select a compute resource 2c Select storage 2d Select compatibility 	Guest OS Family: (Guest OS Version: [Windows Microsoft Windows Server 2008 R2 (64-bit)	 ▼ ▼ 	
2e Select a guest OS 2f Customize hardware 3. Ready to complete				
5 Ready to complete				
		Comp	atibility: ESXi 6.0 and later (VM versio	on 11)
		Back	Next Finish Ca	incel

Figure 7.11: Select the guest OS (click next)

🔁 New Virtual Machine					? H
1 Select creation type ✓ 1a Select a creation type 2 Edit settings	Customize hardware Configure the virtual machine Virtual Hardware VM Opt	e hardware			
2a Select a name and folder 2b Select a compute resource 2c Select storage	GPU Memory	1 4096	• 0 • MB		^
 2d Select compatibility 2e Select a guest OS 	New Hard disk Mew SCSI controller	40 LSI Logic SAS	GB	•	
2f Customize hardware 3 Ready to complete		VM Network Client Device		▼ ✓ Connect	
	New Floppy drive Video card	Client Device	105	Connect	
	WMCI device Mew SATA Controller		.go		
	Other Devices New device:	Select		▼ Add	Ŧ
	new dense.	081601		Compatibility: ESXi 6.0 and later (VM vers	on 11)
				Back Next Finish C	ancel

Figure 7.12: Customize hardware (click next)

The final step in the wizard is to click the "Finish" button to initiate the creation of the VM.

🔁 New Virtual Machine		3) H
1 Select creation type	Provisioning type:	Create a new virtual machine	•
 1a Select a creation type 	Virtual machine name:	vsan6-poc-test-vm-1	
2 Edit settings	Folder:	VSAN6-DC	
 2a Select a name and folder 	Cluster:	VSAN6-Cluster	
 2b Select a compute resource 	Datastore:	vsanDatastore	
✓ 2c Select storage	VM storage policy:	Virtual SAN Default Storage Policy	
 2d Select compatibility 	Guest OS name:	Microsoft Windows Server 2008 R2 (64-bit)	
 2e Select a guest OS 	CPUs:	1	
 2f Customize hardware 	Memory:	4 GB	
✓ 3 Ready to complete	NICs:	1	
	NIC 1 network:	VM Network	
	NIC 1 type:	E1000	
	SCSI controller 1:	LSI Logic SAS	
	Create hard disk 1:	New virtual disk	
	Capacity:	40.00 GB	
	Datastore:	vsanDatastore	
	VM storage	Virtual SAN Default Storage Policy	Ŧ
		Compatibility: ESXi 6.0 and later (VM version 1	1)
		Back Next Finish Cance	

Figure 7.13: Finish VM creation

Once the VM is created, select the new VM in the inventory, navigate to the Manage tab, and then select "Policies". There should be two objects shown, "VM home" and "Hard disk 1". Both of these should show a compliance status of "Compliant" meaning that Virtual SAN was able to deploy these objects in accordance to the policy settings.

🗗 vsan6-poc-test-vm-1 Actions 👻							
Getting Started Summary Monitor Manage Related Objects							
Settings Alarm De	finitions Tags	Permissions	Policies	Scheduled Tasks			
Storage							
C 🍖 🚓							
Name 1	VM Storage Polic	N .		Compliance Sta	tus	Last Checked	
🛅 VM home	🔓 Virtual SAN	I Default Storag	e Policy	🗸 Complian	t	08/04/2015 12:35	
📇 Hard disk 1	F Virtual SAN	Default Storag	e Policy	🗸 Complian	t)	08/04/2015 12:35	

Figure 7.14: VM is compliant with policy settings

To verify this, navigate to the Monitor tab, and then select "Policies". Once again, both the "VM home" and "Hard disk 1" should be displayed. Select "Hard disk 1" and further down the window, select the "Physical Disk Placement" tab. This should display a RAID 1 configuration with two components, each component representing a mirrored copy of the virtual disk. It should also be noted that different components are located on different hosts. This implies that the policy setting to tolerate 1 failure is being adhered to.

😰 vsan6-poc-test-vm-1	vsan6-poc-test-vm-1 Actions *										
Getting Started Summa	etting Started Summary Monitor Manage Related Objects										
Issues Performance Policies Tasks Events Utilization											
									•		
Name		١	VM Storag	e Policy		Compliance Status			Last Checked		
🗀 VM home		1	🚰 Virtu:	al SAN Default Storage Policy		 Compliant 			08/04/2015 12:35		
📇 Hard disk 1		l	🔓 Virtu:	al SAN Default Storage Policy		 Compliant 			08/04/2015 12:35		
A4.											2 items 🔒 🗸
Compliance Failures	Physical Disk Pla	acement			=						
vsan6-poc-test-vm-1 -	Hard disk 1 : Physi	ical Disk Placem	nent								
)- `≣										Q Filter	•
Type 1	Component State	Host		Flash Disk Name	Flash Disk Uui	id	HDD Disk Name		HDD Disk Uuid		
Witness	Active	cs-ie-h01.ie	e.local	🚐 HP Serial Attached SCSI Dis	528ba019-	e369-151e-01b3-26	🚐 HP Serial Attached SCS	BI Dis	5255fd2b-83cc-911f-452b-	13b4ca74e03	le
→ RAID 1											
Component	Active	cs-ie-h02.ie	e.local	🚐 HP Serial Attached SCSI Dis	521963f0-3	3f5-eaaf-d2e1-f7a21	I HP Serial Attached SCS	BI Dis	52edaed1-2b04-b3af-ba3f-	2b03ebaa9fc	e
Component	Active	cs-ie-h03.ie	e.local	HP Serial Attached SCSI Dis	52a4acab-f	1622-6025-bee3-746	HP Serial Attached SCS	BI Dis	523de844-6b48-6bda-44a	d-c28df042c1	6e

Figure 7.15: Physical Disk Placement displays underlying layout of objects

The witness item shown above is used to maintain a quorum. For more information on the purpose of witnesses, and objects and components in general, refer to the <u>VMware Virtual SAN 6.0 Design and Sizing Guide</u>.

One final item is related to the "object space reservation" policy setting that defines how much space a VM reserves on the Virtual SAN datastore. By default, it is set to 0%, implying that the VM's storage objects are entirely "thin" and consume no unnecessary space.

If we examine Figure 7.12, we see that we requested that the VM be deployed with 40GB of disk space. However if we look at the free capacity after the VM has been deployed (as shown in figure 7.16 below), we see that the free capacity is very close to what it was before the VM was deployed, as previously captured in figure 7.3.

Virtual SAN is Turned ON							
Add disks to storage Manual							
Resources							
Hosts	3 hosts						
Flash disks in use	3 of 3 eligible						
Data disks in use	6 of 19 eligible						
Total capacity of Virtual SAN datastore	811.95 GB						
Free capacity of Virtual SAN datastore	811.57 GB						
Network status	 Normal 						

Figure 7.16: Free capacity after VM is created

Of course we have not installed anything in the VM such as a guest OS, but it shows that only a tiny portion of the Virtual SAN datastore has so far been used, verifying

that the object space reservation setting of 0% (essentially thin provisioning) is working correctly.

Do not delete this VM as we will use it for other POC tests going forward.

7.2 Snapshot VM

Using the virtual machine created previously, take a snapshot of it. The snapshot can be taken when the VM is powered on or powered off. The objective here is to see a successful snapshot delta object created, and see that the policy settings of the delta object are inherited directly from the base disk object.



Figure 7.17: Take a VM snapshot

🔞 Take VM Sn	apshot for vsan6-poc-test-vm-1	?
Name	snapshot1	1
Description		
Snapshot the	e virtual machine's memory	1
Quiesce gue	st nie system (Needs VMware Loois Installed)	
	OK Cance	D,

Figure 7.18: Provide a name for the snapshot and optional description

Once the snapshot has been requested, monitor tasks and events to ensure that it has been successfully captured. Once the snapshot creation has completed, additional actions will become available in the snapshot dropdown window. For example there
is a new action to "Revert to Latest Snapshot" and another action to "Manage Snapshots...".



If the "Manage Snapshots..." option is chosen, the following is displayed. It includes details regarding all snapshots in the chain, the ability to delete one or all of them, as well as the ability to revert to a particular snapshot.

Manage VM Snapshots for vsan6-poc-test-vm-1			?
▼ 👜 vsan6-poc-test-vm-1	Name	snapshot1	
▽ 🔞 snapshot1	Description		
⊙ You are here			
	Created	08 April 2015 14:16:26	
	Disk Usage	2.01 MB	
	Console		
	Edit		
4			
Revert to Delete Delete All		CI	ose

Figure 7.20: Manage Snapshots

There is unfortunately no way to see snapshot delta object information from the UI, like we can do for VMDKs and for VM home. Instead, the Ruby vSphere Console (RVC) must be relied on. To get familiar with RVC, see <u>VMware Ruby vSphere Console</u> <u>Command Reference for Virtual SAN</u>.

The command needed to display snapshot information is:

```
vsan.vm object info <VM>
```

Here is an output based on the snapshot created previously:

```
/ie-vcsa-09.ie.local/VSAN6-DC/vms> vsan.vm_object_info 1
 VM VSAN6-poc-test-vm-1:
   Namespace directory
     DOM Object: 95122555-8061-3328-cf10-001f29595f9f (v2, owner: cs-ie-h01.ie.local,
policy: forceProvisioning = 0, hostFailuresToTolerate = 1, spbmProfileId = aa6d5a82-1c88-
45da-85d3-3d74b91a5bad, proportionalCapacity = [0, 100], spbmProfileGenerationNumber = 0,
cacheReservation = 0, stripeWidth = 1)
       RAID 1
        Component: 96122555-80ad-3c97-dadf-001f29595f9f (state: ACTIVE (5), host: cs-ie-
h01.ie.local, md: 52fc637f-ecf9-2b53-ff31-9e8d75d2b43f, ssd: 528ba019-e369-151e-01b3-
26b103d7de0f,
                                                          votes: 1, usage: 0.3 GB)
         Component: 96122555-dc90-3e97-9c6f-001f29595f9f (state: ACTIVE (5), host: cs-ie-
h02.ie.local, md: 52edaed1-2b04-b3af-ba3f-2b03ebaa9fce, ssd: 521963f0-33f5-eaaf-d2e1-
f7a218b13be4.
                                                          votes: 1, usage: 0.3 GB)
       Witness: 96122555-fc7b-3f97-5d9a-001f29595f9f (state: ACTIVE (5), host: cs-ie-
h03.ie.local, md: 527aade4-cec7-0661-b621-6e22d69c3042, ssd: 52a4acab-f622-6025-bee3-
746d436627cf.
                                                      votes: 1, usage: 0.0 GB)
   Disk backing: [vsanDatastore] 95122555-8061-3328-cf10-001f29595f9f/VSAN6-poc-test-vm-
1-000001.vmdk
     DOM Object: 2a2a2555-946f-292b-2e23-001f29595f9f (v2, owner: cs-ie-h01.ie.local,
policy: spbmProfileGenerationNumber = 0, forceProvisioning = 0, cacheReservation = 0,
hostFailuresToTolerate = 1, stripeWidth = 1, spbmProfileId = aa6d5a82-1c88-45da-85d3-
3d74b91a5bad, proportionalCapacity = [0, 100], objectVersion = 2)
       RAID 1
         Component: 2a2a2555-8ce3-a171-fb8e-001f29595f9f (state: ACTIVE (5), host: cs-ie-
h01.ie.local, md: 5255fd2b-83cc-911f-452b-13b4ca74e03e, ssd: 528ba019-e369-151e-01b3-
26b103d7de0f,
                                                          votes: 1, usage: 0.0 GB)
         Component: 2a2a2555-78d0-a371-b90d-001f29595f9f (state: ACTIVE (5), host: cs-ie-
h02.ie.local, md: 52edaed1-2b04-b3af-ba3f-2b03ebaa9fce, ssd: 521963f0-33f5-eaaf-d2e1-
f7a218b13be4,
                                                          votes: 1, usage: 0.0 GB)
       Witness: 2a2a2555-ce29-a571-da2b-001f29595f9f (state: ACTIVE (5), host: cs-ie-
h03.ie.local, md: 527aade4-cec7-0661-b621-6e22d69c3042, ssd: 52a4acab-f622-6025-bee3-
746d436627cf,
                                                       votes: 1, usage: 0.0 GB)
     Disk backing: [vsanDatastore] 95122555-8061-3328-cf10-001f29595f9f/VSAN6-poc-test-
vm-1.vmdk
      DOM Object: 97122555-78d5-5580-bffc-001f29595f9f (v2, owner: cs-ie-h03.ie.local,
policy: forceProvisioning = 0, hostFailuresToTolerate = 1, spbmProfileId = aa6d5a82-1c88-
45da-85d3-3d74b91a5bad, proportionalCapacity = 0, spbmProfileGenerationNumber = 0,
cacheReservation = 0, stripeWidth = 1)
         RAID 1
           Component: 98122555-3ec9-d1d6-01f4-001f29595f9f (state: ACTIVE (5), host: cs-
ie-h02.ie.local, md: 52edaed1-2b04-b3af-ba3f-2b03ebaa9fce, ssd: 521963f0-33f5-eaaf-d2e1-
f7a218b13be4.
                                                             votes: 1, usage: 0.0 GB)
           Component: 98122555-5c6f-d3d6-55a7-001f29595f9f (state: ACTIVE (5), host: cs-
ie-h03.ie.local, md: 523de844-6b48-6bda-44ad-c28df042c16e, ssd: 52a4acab-f622-6025-bee3-
746d436627cf,
                                                            votes: 1, usage: 0.0 GB)
         Witness: 98122555-2028-d4d6-6ee6-001f29595f9f (state: ACTIVE (5), host: cs-ie-
h01.ie.local, md: 5255fd2b-83cc-911f-452b-13b4ca74e03e, ssd: 528ba019-e369-151e-01b3-
26b103d7de0f,
                                                        votes: 1, usage: 0.0 GB)
 /ie-vcsa-09.ie.local/VSAN6-DC/vms>
```

The three objects that are now associated with that virtual machine have a bold font in this document for clarity. There is the namespace directory (VM home), there is the disk VSAN6-poc-test-vm-1.vmdk and there is the snapshot delta VSAN6-poc-test-vm-1-000001. vmdk. The snapshot delta has been highlighted in blue above.

If you look closely, both of the disk backings have the same policy settings since every snapshot inherits its policy settings from the base disk. Both have a *stripeWidth* of 1, and *hostFailuresToTolerate* of 1 and an Object Space Reservation (shown as *proportionalCapacity* here) of 0%.

The snapshot can now be deleted from the VM. Monitor the VM's tasks and ensure that it deletes successfully. When complete, snapshot management should look similar to this.



Figure 7.21: Manage Snapshots... Snapshot deleted

This completes the snapshot section of this POC. Snapshots in a Virtual SAN datastore are very intuitive because they utilize vSphere native snapshot capabilities. Starting with Virtual SAN 6.0, they are stored efficiently using "vsansparse" technology that improves the performance of snapshots compared to Virtual SAN 5.5. In Virtual SAN 6.1, snapshot chains can be up to 16 snapshots deep.

7.3 Clone a VM

The next POC test is cloning a VM. We will continue to use the same VM as before. This time make sure the VM is powered on first. There are a number of different cloning operations available in vSphere 6. These are shown here.

Actions - vsan6-poc-test-vm-1	
Power	•
Guest OS	•
Snapshots	•
🛃 Open Console	
🖶 Migrate	
Clone	Clone to Virtual Machine
Template	Cione to Template
Fault Tolerance	 Clone to Template in Library

Figure 7.22: Clone operations

The one that we shall be running as part of this POC is the "Clone to Virtual Machine". The cloning operation is very much a "click, click, next" type activity. This next screen is the only one that requires human interaction. One simply provides the name for the newly cloned VM, and a folder if desired.

⁶ vsan6-poc-test-vm-1 - Clone Exis	ting Virtual Machine	?)
1 Edit settings 1a Select a name and folder	Select a name and folder Specify a unique name and target location	
1b Select a compute resource	Enter a name for the virtual machine.	
1c Select storage	vsan6-poc-test-vm2	
1d Select clone options	Mrtual machine names can contain up to 80 characters and they must be unique within each vCenter Server VM folder.	
2 Ready to complete	Select a location for the virtual machine.	
	Q Search Image: Search Select a datacenter or VM folder to create the new virtue machine in.	al
	Back Next Finish C	Cancel

Figure 7.23: Select a name and folder

We are going to clone the VM in the Virtual SAN Cluster, so this must be selected as the compute resource.

😤 vsan6-poc-test-vm-1 - Clone Exis	sting Virtual Machine		? •
Edit settings In Select a name and folder In Select a compute resource Ic Select storage Id Select clone options 2 Ready to complete	Select a compute resource Select the destination compute resource for this operation	Select a cluster, host, wapp or resource pool to run this virtual machine.	
	Compatibility: Compatibility checks succeeded.		
		Back Next Finish C	ancel

Figure 7.24: Select a compute resource

The storage will be the same as the source VM, namely the vsanDatastore. This will all be pre-selected for you if the VM being cloned also resides on the vsanDatastore.

😤 vsan6-poc-test-vm-1 - Clone Exist	ting Virtual Machine								(? ₩
1 Edit settings ✓ 1a Select a name and folder	Select storage Select the datastore in wh	ich to store	the configuration	n and disk files					
 1b Select a compute resource 1c Select storage 	Select virtual disk format: VM Storage Policy:	As define Virtual SA	ed in the VM stora	ge policy	v •	0			
1d Select clone options 2 Ready to complete	The following datastores virtual machine configurat	are access tion files ar	ible from the des nd all of the virtua	stination resou I disks.	rce ti	nat yo	u selected. Selec	t the destination d	atastore for the
	Name		Capacity	Provisioned			Free	Туре	Storage DR
	compatible		811.95 GB	84.55 GB			811.55 GB	vsan	
	Incompatible								
	NFS-Isilon		50.46 TB	379.15 G	в		50.19 TB	NFS v3	
	corkisos.csl.vmware	.com	1.77 TB	1.61 TB			170.62 GB	NFS v3	
	Fill cs-ie-h02-scratch		136 50 GB	3.06.GB			133.44 GB	VMES	¥
									Advanced >>
	Compatibility:								
	Compatibility checks succeeded.								
	Virtual SAN storage	consumpti	on would be 80.0	00 GB disk spa	ice a	nd 0.0	00 B reserved Fla	sh space.	
							Back	Next Finis	h Cancel

Figure 7.25: Select storage

s ^a vsan6-poc-test-vm-1 - Clone Exist	ting Virtual Machine
Seiect a name and folder 1a Select a name and folder 1b Select a compute resource 1c Select storage 1d Select clone options 2 Ready to complete	ing Virtual Machine Select clone options Select furmer clone options Customize the operating system Customize this virtual machine's hardware (Experimental) Power on virtual machine after creation
	Back Next Finish Cancel

Figure 7.26: Select options (leave unchecked - default)

This will take you to the "Ready to Complete" screen. If everything is as expected, click Finish to commence the clone operation. Monitor the VM tasks for status of the clone operation.

88	vsan6-poc-test-vm-1 - Clone Exist	ing Virtual Machine					(?)»
	1 Edit settings	Provisioning type:	Clone an existing virtual machine				
~	1a Select a name and folder	Source virtual machine:	vsan6-poc-test-vm-1				
~	1b Select a compute resource	Virtual machine name:	vsan6-poc-test-vm2				
~	1c Select storage	Folder:	VSAN6-DC				
~	1d Select clone options	Cluster:	VSAN6-Cluster				
~	2 Ready to complete	Datastore:	vsanDatastore				
		Disk storage:	As defined in the VM storage policy				
		VM storage policy:	Virtual SAN Default Storage Policy				
				Back	Next	Finish	Cancel

Figure 7.27: Ready to Complete

Do not delete the newly cloned VM. We will be using it in subsequent POC tests.

This completes the cloning section of this POC. Cloning with Virtual SAN has improved dramatically with the new on-disk (v2) format in version 6.0 and 6.1.

7.4 vMotion a VM between Hosts

The first step is to power-on the newly cloned virtual machine. We shall migrate this VM from one Virtual SAN host to another Virtual SAN host using vMotion.

Note: Take a moment to revisit the network configuration and ensure that the vMotion network is distinct from the Virtual SAN network. If these features share the same network, performance will not be optimal.

First, determine which ESXi host the VM currently resides on. Selecting the "Summary" tab of the VM does this. On this POC, the VM that we wish to migrate is on host *cs-ie-h01.ie.local*.



Figure 7.28: VM Summary tab - Host is displayed

Right click on the VM and select **Migrate**.



Figure 7.29: Migrate

Migrate allows you to migrate to a different compute resource (host), a different datastore or both at the same time. In this initial test, we are simply migrating the VM to another host in the cluster, so this initial screen can be left at the default of "Change compute resource only". The rest of the screens in the migration wizard are pretty self-explanatory.

vsano-poc-test-vinz - migrate		?
1 Select the migration type 2 Select a compute resource	Select the migration type Change the virtual machines' compute resource, storage, or both.	
 Select a compact resource Select who priority Ready to complete 	 Change compute resource only Migrate the virtual machines to another host or cluster. Change storage only Migrate the virtual machines' storage to a compatible datastore or datastore cluster. Change both compute resource and storage Migrate the virtual machines to a specific host or cluster and their storage to a specific datastore or datastore cluster. Select compute resource first Select storage first 	

Figure 7.30: Change compute resources only

🖆 vsan6-poc-test-vm2 - Migrate		(?) ₩
 1 Select the migration type 2 Select a compute resource 	Select a compute resource Select a cluster, host, vApp or resource pool to run the virtual machines.	
3 Select network 4 Select vMotion priority 5 Ready to complete	Filter	•)
	Name 1 Custer Image: Constant of the state of the s	3 Objects
	Back Next Finis	h Cancel

Figure 7.31: Select a destination host

🗗 vsan6-poc-test-vm2 - Migrate				(?) ₩					
 1 Select the migration type 2 Select a compute resource 	Select network Select the destination network for	Select network Select the destination network for the virtual machine migration.							
3 Select network 4 Select vMotion priority	Migrate VM networking by selecting	ng a new destination network for all VM network	adapters attached to the same sour	ce network.					
 4 Select vision priority 5 Deadu to complete 	Source Network	Used By	Destination Network						
5 Ready to complete	VM Network	1 VMs / 1 Network adapters	VM Network	•					
				Advanced >>					
	Compatibility:								
	Compatibility checks succe	eded.							
			Back Next Finish	Cancel					

Figure 7.32: Select a destination network



Figure 7.33: Priority can be left as high (default)

At the "Ready to Complete" window, click on Finish to initiate the migration. If the migration is successful, the summary tab of the virtual machine should show that the VM now resides on a different host.

vsan6-poc-test-vm2 Ac	tions 👻		
Getting Started Summary	Monitor	Manage	Related Objects
Advect lood Type Into 1 2000 Constants 42 2003 2004 (Annum Ten	vs	an6-poc-tes	st-vm2
Depring to the Device of the second s	G	lest OS:	Microsoft Window's Server 2008 R2 (64-bit)
PRE-MET. Dilling latel 250 JUN Aperaling Systems and Samuel	Co	mpatibility:	ESXi 6.0 and later (VM version 11)
	VI	Ww are Tools:	Not running, not installed
	D	IS Name:	
Powered On	IP	Addresses:	
	Ho	ost:	cs-ie-h02.ie.local
Launch Remote Console	_		
Download Remote Console	0 ~		

Figure 7.34: Verify VM has migrated to new host

Do not delete the migrated VM. We will be using it in subsequent POC tests.

This completes the "VM migration using vMotion" section of this POC. As you can see, vMotion works just great with Virtual SAN.

7.5 Optional: Storage vMotion a VM between Datastores

This test will only be possible if you have another datastore type available to your hosts, such as NFS/VMFS. If so, then the objective of this test is to migrate the VM from another datastore type into Virtual SAN. The VMFS datastore can even be a local VMFS disk on the host.

7.5.1 Mount an NFS Datastore to the Hosts

The steps to mount an NFS datastore to multiple ESXi hosts are described in the vSphere 6.0 Administrators Guide. See the <u>Create NFS Datastore in the vSphere</u> <u>Client</u> topic for detailed steps.

7.5.2 Storage vMotion a VM from Virtual SAN to Another Datastore Type

Currently the VM resides on the Virtual SAN datastore. Launch the migrate wizard, just like we did in the last exercise. However, on this occasion, to move the VM from the Virtual SAN datastore to the other datastore type you need to select "Change storage only".

😰 vsan6-poc-test-vm2 - Migrate		? H
 vsan6-poc-test.vm2 - Migrate 1 Select the migration type 2 Select storage 3 Ready to complete 	Select the migration type Change the virtual machines' compute resource, storage, or both. Change compute resource only Migrate the virtual machines to another host or cluster. Change storage only Migrate the virtual machines' storage to a compatible datastore or datastore cluster. Change both compute resource and storage Migrate the virtual machines to a specific host or cluster and their storage to a specific datastore or datastore cluster. Select compute resource first Select storage first	•
	Back Next Finish	Cancel

Figure 7.35: Change storage only

In this POC, we have an NFS datastore presented to each of the ESXi hosts in the Virtual SAN Cluster. This is the datastore where we are going to migrate the virtual machine to.

🖆 vsan6-poc-test-vm2 - Migrate								(4 ?)
 1 Select the migration type 2 Select storage 	Select storage Select the destination storage for the virtual machine migration.							
3 Ready to complete	Select virtual disk format:	Thin Provision 🔹						
	VM Storage Policy: The following datastores a virtual machine configurat	M Storage Policy: Datastore Default					tastore for the	
	Name		Capacity	Provisioned		Free	Type	Storage DRS
	NFS-Isilon		50.46 TB	379.15 GE	3	50.19 TB	NFS v3	
	🗐 vsanDatastore		811.95 GB	93.64 GB		802.45 GB	vsan	
	🗐 cs-ie-h02-scratch		136.50 GB	3.07 GB		133.43 GB	VMFS 5	
	4		::					•
								Advanced >>
	Compatibility:							
	Compatibility checks	succeede	ed.					
						Back	Next Finish	Cancel

Figure 7.36: Select destination storage

One other item of interest in this step is that the VM Storage Policy should also be changed to "Datastore Default" as the NFS datastore will not understand the Virtual SAN policy settings.

At the "Ready to complete" screen, click "Finish" to initiate the migration:

🖞 vsan6-poc-test-vm2 - Migrate			? ₩
 1 Select the migration type 2 Select storage 	Ready to complete The wizard is ready	. Verify that the information is correct and click finish to start the migration.	
 2 Select storage 3 Ready to complete 	Migration Type Virtual Machine Storage Disk Format	Change storage. Leave VM on the original compute resource. vsan6-poc-test-vm2 [NFS-Isilon] Thin Provision	
		Back Next Finish Ca	ncel

Figure 7.37: Ready to complete

Once the migration completes, the VM Summary tab can be used to examine the datastore on which the VM resides.

vsan6-poc-test-vm2 Actio	ons 🔻			
Getting Started Summary	Monitor Manage	Related Objects		
Marking for the first fir	vsan6-poc-te: Guest OS: Compatibility: VMw are Tools: DNS Name: IP Addresses: Host:	st-vm2 Microsoft Windows ESXi 6.0 and later (Not running, not ins cs-ie-h02.ie.local	rver 2008 R2 (64-bit) version 11) :d	
A VMware Tools is not instal	led on this virtual m	achine.		
 VM Hardware 			 Advanced Configuration 	
► CPU	1 CPU(s), 0 MHz	used	- N-4	
▶ Memory	4096 MB, 0 M	B memory active	 Notes 	
Hard disk 1				
Capacity	40.00 GB			
Location	NFS-Isilon (50.	2 TB free)		Edit

Figure 7.38: Verify VM has moved to new storage

7.5.3 Storage vMotion of VM to Virtual SAN from Another Datastore Type

Now Storage vMotion the virtual machine back to the Virtual SAN datastore to prove that Storage vMotion works in both directions. This now completes the optional "VM migration using Storage vMotion" section of this POC. Different storage policies can be chosen as part of the migration.

Storage vMotion works seamlessly with Virtual SAN.

8. Scale out Virtual SAN

One of the really nice features is the simplistic scale-out nature of Virtual SAN. If you need more compute or storage resources in the cluster, simply add another host to the cluster.

Let's remind ourselves about how our cluster currently looks. There are currently three hosts in the cluster, and there is a fourth host not in the cluster. We also created two VMs in the previous exercises.



Figure 8.1: Current inventory status

Let us also remind ourselves of how big the Virtual SAN datastore is.

VSAN6-Cluster Actions -		<u>F</u> *
Getting Started Summary Mo	onitor Manage Related Objects	
Settings Scheduled Tasks Aa	arm Definitions Tags Permissions	
	Virtual SAN is Turned ON	Edit
▼ Services	Add disks to storage Manual	
vSphere DRS vSphere HA	Resources	
✓ Virtual SAN	Hosts	3 hosts
General	Flash disks in use	3 of 3 eligible
Disk Management	Data disks in use	6 of 19 eligible
Fault Domains	Total capacity of Virtual SAN datastore	811.95 GB
Health	Free capacity of Virtual SAN datastore	811.21 GB
	Network status	V Normal
General		
Licensing	On-disk Format Version	Upgrade
VMware EVC	Virtual SAN format version Version 2	(latest)
VM/Host Groups	Disks with older version 📀 0 of 9)
VM/Host Rules	Support Accietant	
VM Overrides	support Assistant	Upload Support Bundles to Service Request
Host Options	Last upload time -	
Profiles		

Figure 8.2: Total and Free Virtual SAN datastore capacity

In this POC, the Virtual SAN datastore is 811.95GB in size with 811.21GB free.

8.1 Add the Fourth Host to Virtual SAN Cluster

We will now proceed with adding a fourth host to the Virtual SAN Cluster.

Note: Back in section 5 of this POC guide, you should have already setup a Virtual SAN network for this host. If you have not done that, revisit section 5, and setup the Virtual SAN network on this fourth host.

Having verified that the networking is configured correctly on the fourth host, select the cluster object in the inventory, right click on it and select the option "Move Hosts into Cluster..." as shown below.



Figure 8.3: Move hosts into Cluster

You will then be prompted to select which host to move into the cluster. In this POC, there is only one additional host. Select that host.

Move Hosts into Cluster			×
Filter (1) Selected Objects			
		Q Filter	•
Name	Cluster		
Cs-ie-h04.ie.local			
M			1 items
		ОК	Cancel

Figure 8.4: Select a host to move into the cluster

The next screen is related to resource pools. You can leave this at the default, which is to use the cluster's root resource pool, then click OK.

VSAN6-Cluster - Move Host into This Cluster	(?)»			
What would you like to do with the virtual machines and resource pools for cs-ie-h04.ie.loca	al?			
 Put all of this host's virtual machines in the cluster's root resource pool. Resource pools currently present on the host will be deleted. 				
Create a new resource pool for this host's virtual machines and resource pools. This preserves the host's current resource pool hierarchy.				
Resource Pool Name: Grafted from cs-ie-h04.ie.local 1				
ОК Саг	ncel			

Figure 8.5: Resource Pools

This moves the host into the cluster. Next, navigate to the Manage tab > Settings > Virtual SAN > General view and verify that the cluster now contains the new node.

VSAN6-Cluster Actions -			±
Getting Started Summary Mo	nitor Manage Related Objects		
Settings Scheduled Tasks Ala	rm Definitions Tags Permissions		
44	Virtual SAN is Turned ON		Edit
✓ Services	Add disks to storage Manual		
vSphere HA	Resources		
✓ Virtual SAN	Hosts	4 hosts	
General	Flash disks in use	3 of 4 eligible	
Disk Management	Data disks in use	6 of 25 eligible	
Fault Domains	Total capacity of Virtual SAN datastore	811.95 GB	
Health	Free capacity of Virtual SAN datastore	811.21 GB	
- Configuration	Network status	✓ Normal	
General			
Licensing	On-disk Format Version		Upgrade
VMware EVC	Virtual SAN format version Version	2 (latest)	
VM/Host Groups	Disks with older version 🛛 🥥 0 of	9	
VM/Host Rules VM Overrides	Support Assistant	[Upload Support Bundles to Service Request
Host Options	Last upload time -		
Profiles			

Figure 8.6: Resource Pools

As you can clearly see, there are now 4 hosts in the cluster. However, you will also notice that the Virtual SAN datastore has not changed with regards to total and free capacity. This is because the cluster was configured in "Manual" mode back in section 6. Therefore Virtual SAN will not claim any of the disks automatically. You will need to create a disk group for the new host and claim disks manually. At this point, it

would be good practice to re-run the health check tests. If there are any issues with the fourth host joining the cluster, use the Virtual SAN Health check to check where the issue lies. Verify that the host appears in the same network partition group as the other hosts in the cluster.

8.2 Manual Option: Create Disk Group on New Host

This process has already been covered in section 6.2. Navigate to the Disk Management section, select the new host and then click on the icon to create a new disk group:

U VSAN6-Cluster Actions -						=*
Getting Started Summary M	onitor Manage Related Objects					
Settings Scheduled Tasks A	arm Definitions Tags Permissions					
	Disk Groups					
▼ Services	🛥 🕋 🔍 🖉 🐂 👘			Q	Filter	•
vSphere DRS	Disk Group	Disks in Use	State	Virtual SAN	Fault Domain	Network Part
vSphere HA	👻 📱 cs-ie-h03.ie.local	3 of 7	Connected	Healthy		Group 1
	Bisk group (0200080000600508b1001c9c8b5f6f0d7a2be	3		Healthy		
General	👻 🗐 cs-ie-h01.ie.local	3 of 8	Connected	Healthy		Group 1
Disk Management	Disk group (0200080000600508b1001c61cedd42b0c3fbf5	3		Healthy		
Fault Domains	✓ ☐ cs-ie-h02.ie.local	3 of 7	Connected	Healthy		Group 1
Health	Disk group (0200080000600508b1001c64b76c8ceb56e8	3		Healthy		
- Configuration	cs-ie-h04.ie.local	0 of 7	Connected	Healthy		Group 1

Figure 8.7: Create a new disk group

As before, we select a flash device and two magnetic disks. This is so that all hosts in the cluster maintain a uniform configuration.

ret colocto cinglo flach diek to conco	as a write cache or	nd road buffer					
rst, select a single liash disk to serve	as a write cache ar	id read buller.					
					Q Filter		
Name		Drive Type	Capacity	Transport Type			
HP Serial Attached SCSI Disk ((naa.600508b10	Flash	186.28 GB	Block Adapter			
						1 items	E
1						T Itellio	
en selectione or more HDD disks to	serve as data disk						
					Q Filter		
Name		Drive Type	Capacity	Transport Type			
📃 🚐 HP Serial Attached SCSI Disk ((naa.600508b10	HDD	136.70 GB	Block Adapter			
HP Serial Attached SCSI Disk ((naa.600508b10 (naa.600508b10	HDD HDD	136.70 GB 136.70 GB	Block Adapter Block Adapter			
HP Serial Attached SCSI Disk ((naa.600508b10 (naa.600508b10 (naa.600508b10	HDD HDD HDD	136.70 GB 136.70 GB 136.70 GB	Block Adapter Block Adapter Block Adapter			
AP Serial Attached SCSI Disk (AF Serial A	(naa.600508b10 (naa.600508b10 (naa.600508b10 (naa.600508b10	HDD HDD HDD HDD	136.70 GB 136.70 GB 136.70 GB 136.70 GB	Block Adapter Block Adapter Block Adapter Block Adapter			
HP Serial Attached SCSI Disk (HP Serial A	(naa.600508b10 (naa.600508b10 (naa.600508b10 (naa.600508b10 (naa.600508b10	HDD HDD HDD HDD HDD	136.70 GB 136.70 GB 136.70 GB 136.70 GB 136.70 GB	Block Adapter Block Adapter Block Adapter Block Adapter Block Adapter			
Altached SCSI Disk (Altached SCSI Di	(naa.600508b10 (naa.600508b10 (naa.600508b10 (naa.600508b10 (naa.600508b10 (naa.600508b10	HDD HDD HDD HDD HDD HDD	136.70 GB 136.70 GB 136.70 GB 136.70 GB 136.70 GB 136.70 GB	Block Adapter Block Adapter Block Adapter Block Adapter Block Adapter Block Adapter			
A HP Serial Attached SCSI Disk (A HP S	(naa.600508b10 (naa.600508b10 (naa.600508b10 (naa.600508b10 (naa.600508b10 (naa.600508b10	HDD HDD HDD HDD HDD HDD	136.70 GB 136.70 GB 136.70 GB 136.70 GB 136.70 GB 136.70 GB	Block Adapter Block Adapter Block Adapter Block Adapter Block Adapter Block Adapter			
HP Serial Attached SCSI Disk ((naa.600508b10 (naa.600508b10 (naa.600508b10 (naa.600508b10 (naa.600508b10 (naa.600508b10 (naa.600508b10	HDD HDD HDD HDD HDD HDD HDD HDD	136.70 GB 136.70 GB 136.70 GB 136.70 GB 136.70 GB 136.70 GB 136.70 GB Disk 5d80ba7665b	Block Adapter Block Adapter Block Adapter Block Adapter Block Adapter Block Adapter Block Adapter			
HP Serial Attached SCSI Disk ((naa.600508b10 (naa.600508b10 (naa.600508b10 (naa.600508b10 (naa.600508b10 (naa.600508b10 (naa.600508b10	HDD HDD HDD HDD HDD HDD HDD Sached SCSI Bb1001cadff	136.70 GB 136.70 GB 136.70 GB 136.70 GB 136.70 GB 136.70 GB Disk 5d80ba7665t	Block Adapter Block Adapter Block Adapter Block Adapter Block Adapter Block Adapter Block Adapter			
HP Serial Attached SCSI Disk (HP Serial A	(naa.600508b10 (naa.600508b10 (naa.600508b10 (naa.600508b10 (naa.600508b10 (naa.600508b10 HP Serial Att (naa.60050b	HDD HDD HDD HDD HDD HDD HDD sached SCSI Bb1001cadff	136.70 GB 136.70 GB 136.70 GB 136.70 GB 136.70 GB 136.70 GB 136.70 GB Disk 5d80ba7665t	Block Adapter Block Adapter Block Adapter Block Adapter Block Adapter Block Adapter Block Adapter		6 items	

Figure 8.8: Select flash and capacity devices

8.3 Verify Virtual SAN Disk Group Configuration on New Host

Once the disk group has been created, the disk management view should be revisited to ensure that it is healthy.

U VSAN6-Cluster Actions -						=
Getting Started Summary	Monitor Manage Related Objects					
Settings Scheduled Tasks A	Varm Definitions Tags Permissions					
44	Disk Groups					
✓ Services				Q	Filter	-
vSphere DRS	Disk Group	Disks in Use	State	Virtual SAN	Fault Domain	Network Part
vSphere HA	E cs-ie-h03.ie.local	3 of 7	Connect	ed Healthy		Group 1
✓ Virtual SAN	Disk group (0200080000600508b1001c9c8b5f6f0d7a2be	3		Healthy		
General	cs-ie-h01.ie.local	3 of 8	Connect	ed Healthy		Group 1
Disk Management	Bisk group (0200080000600508b1001c61cedd42b0c3fbf5	3		Healthy		
Fault Domains	👻 📋 cs-ie-h02.ie.local	3 of 7	Connect	ed Healthy		Group 1
Health	Disk group (0200080000600508b1001c64b76c8ceb56e8	3		Healthy		
- Configuration		3 of 7	Connect	ed Healthy		Group 1
General	Disk group (0200080000600508b1001c29d8145d6cc192	3		Healthy		
Licensing						
VMware EVC	4					•
VM/Host Groups	M				8 if	ems 🔒 🕶
VM/Host Rules		-				
VM Overrides	Disk group (0200080000600508b1001c29d8145d6cc1925e9fb94c4f4)	/494341): Disl	ks			
Host Options	<u>*</u>			Show:	In use (3)	•
Profiles	Name	Drive Type	Capacity	Virtual SAN Health Status	Operational	Transport Ty
	F HP Serial Attached SCSI Disk (naa.600508b1001c29d8145d6cc	Flash	186.28 GB	Healthy	Mounted	Block Ada
	HP Serial Attached SCSI Disk (naa.600508b1001c846c000c3d9	HDD	136.70 GB	Healthy	Mounted	Block Ada
	HP Serial Attached SCSI Disk (naa.600508b1001c258181f0a088	HDD	136.70 GB	Healthy	Mounted	Block Ada

Figure 8.9: Check disk group health

8.4 Verify New Virtual SAN Datastore Capacity

The final step is to ensure that the Virtual SAN datastore has now grown in accordance to the capacity devices in the disk group that was just added on the fourth host. Return to the General tab and examine the total and free capacity field.

VSAN6-Cluster Actions -			
Getting Started Summary Mo	nitor Manage Related Obje	cts	
Settings Scheduled Tasks Ala	rm Definitions Tags Permiss	ions	
"	Virtual SAN is Turned ON		Edit
▼ Services	Add disks to storage Man	ual	
vSphere DRS	Resources		
- Virtual SAN	Hosts		4 hosts
General	Flash disks in use		4 of 4 eligible
Disk Management	Data disks in use		8 of 25 eligible
Fault Domains	Total capacity of Virtual SAN of	datastore (1.06 TB
Health	Free capacity of Virtual SAN of	iatastore	1.06 TB
- Configuration	Network status		✓ Normal
General Licensing	On-disk Format Version		Upgrade
VMware EVC	Virtual SAN format version	Version 2	2 (latest)
VM/Host Groups	Disks with older version	Ø 0 of 1:	12
VM/Host Rules VM Overrides	Support Assistant		Upload Support Bundles to Service Request.
Profiles			

Figure 8.10: Virtual SAN Datastore capacity details

As we can clearly see, the Virtual SAN datastore has now grown in size to 1.06TB. Free space is shown as 1.06TB as the amount of space used is minimal.

This completes the "Scale Out" section of this POC. As seen, scale-out on Virtual SAN is simple but very powerful.

9. VM Storage Policies and Virtual SAN

VM Storage Policies form the basis of VMware's Software Defined Storage vision. Rather than deploying VMs directly to a datastore, a VM Storage Policy is chosen during initial deployment. The policy contains characteristics and capabilities of the storage required by the virtual machine. Based on the policy contents, the correct underlying storage is chosen for the VM.

If the underlying storage meets the VM storage Policy requirements, the VM is said to be in a compatible state.

If the underlying storage fails to meet the VM storage Policy requirements, the VM is said to be in an incompatible state.

In this section of the POC Guide, we shall look at various aspects of VM Storage Policies. The virtual machines that have been deployed thus far have used the default storage policy, which has the following settings:

- *NumberOfFailuresToTolerate* = 1
- *NumberOfDiskObjectsToStripe* = 1
- *ObjectSpaceReservation* = 0%
- *FlashReadCacheReservation* = 0%
- ForceProvisioning = False

We will create some additional policies in this section of the POC.

9.1 Create a New VM Storage Policy

In this part of the POC, we will build a policy that creates a stripe width of 2 for each storage object deployed with this policy. The VM Storage Policies can be accessed from the Home page on the vSphere web client as shown below.



Figure 9.1: VM Storage Policies

There will be some existing policies already in place, such as the Virtual SAN Default Storage policy, which we've already used to deploy VMs in section 7 of this POC guide. There is another policy called "VVol No Requirements Policy", which is used for Virtual Volumes and is not applicable to Virtual SAN. There are a number of icons on this page that may need further explanation:

1 5	Create a new VM Storage Policy
	Edit an existing VM Storage Policy
×	Delete an existing VM Storage Policy
&	Check the compliance of VMs using this VM Storage Policy
Ê	Clone an existing VM Storage Policy

Table 9.1: VM Storage Policy icons

To create a new policy, click on the "Create a new VM Storage Policy" icon.

Navigator	Torage Policies		
- Home 🔊 🔊	Objects		
From VM Storage Policies			
Contract SAN Default Storage Policy	[(🔁 🌽 🛛 🍪 Actions	-	
	Name	Description	VC
To VVol No Requirements Policy	F Virtual SAN Default Storage	Storage policy used as default f	🕝 ie-vcsa-09.ie.local
	🔂 VVol No Requirements Policy	Allow the datastore to determine	🝘 ie-vcsa-09.ie.local

Figure 9.2: Create a new VM Storage Policy

The next step is to provide a name and an optional description for the new VM Storage Policy. Since this policy will contain a stripe width of 2, we have given it a name to reflect this. You may also give it a name to reflect that it is a Virtual SAN policy.

🚰 Create New VM Storage Policy		(?) ₩
 I Name and description Rule-Sets 2a Rule-Set 1 3 Storage compatibility 4 Ready to complete 	Name and description Enter a name and (optional) description vCenter Server: Ie-vcsa-09.Ie.local • Name: StripeVidth=2 Description: NumberOfDiskStripesPerObject	
	Back	Finish Cancel

Figure 9.3: VM Storage Policy Name and Description

The next section contains a description of Rule-Sets and how to use them.

😚 Create New VM Storage Policy						· (?) >>	
I Name and description Rule-Sets 2 Rule-Sets AVM storage policy consists of rules that describe requirements for the storage resources. The policy can include multiple rule-sets describing the storage resource requirements for different datastore types. 2a Rule-Set 1 The VM storage policy will match datastores that satisfy all the rules in at least one of the rule-sets.							
3 Storage compatibility 4 Ready to complete Gold Storage Policy							
	Rule-set 1 Datastore type 1	Rule-set 2 Datastore type 2	or	Rule-set 3 Gold tags			
				Back	Next Finish C	ancel	

Figure 9.4: Rule-Sets

Now we get to the point where we create a set of rules for our Rule-Set (we are only creating a single Rule-Set in this VM Storage Policy). The first step is to select "Virtual

SAN" as the "Rules based on data services". Once this is selected, the five customizable capabilities associated with Virtual SAN are exposed. Since this VM Storage Policy is going to have a requirement where the stripe width of an object is set to two, this is what we select from the list of rules. It is officially called "*Number of disk stripes per object*".

🚰 Create New VM Storage Policy		? ₩
 1 Name and description 2 Rule-Sets 	Rule-Set 1 Select rules specific for a datastore type. Rules can be based on data services provided by datastore or based on tags. The VM storage policy will match datastores that satisfy all the rules in at least one of the rule-sets.	
2a Rule-Set 1 3 Storage compatibility 4 Ready to complete	Rules based on data services VSAN Step 1 Add rule> Step 2 Add rule> Rubs tripes per object Flash read cache reservation (%) Flash read cache read cache reservation (%) Flash read cache read	
	Force provisioning Object space reservation (%)	
	Add another rule set Remove this rule Back Next Finish Car	ncel

Figure 9.5: Number of disk stripes per object

We also want to set this value to 2. Once the disk stripe rule is chosen, change the default value from 1 to 2 as shown below. Notice also the Storage Consumption Model display on the right hand side, detailing how much disk space will be consumed based on the rules placed in the policy.

😤 Create New VM Storage Policy		?₩
 1 Name and description 2 Rule-Sets 	Rule-Set 1 Select rules specific for a datastore type. Rules can be based on data services provided by data The VM storage policy will match datastores that satisfy all the rules in at least one of the rule-se	store or based on tags. ets.
2a Rule-Set 1 3 Storage compatibility 4 Ready to complete	Rules based on data services VSAN Number of disk stripes per object 2 «Add rule» Rules based on tags Add tag-based rule Add another	Storage Consumption Model A virtual disk with size 100 GB would consume: Storage space 200.00 GB Initially reserved storage space 0.00 B Reserved flash space 0.00 B
	Back	Next Finish Cancel

Figure 9.6: Setting Stripe Width to 2

Clicking next moves on to the Storage Compatibility screen. Note that this displays which storage "understands" the policy settings. In this case, the vsanDatastore is the only datastore that is compatible with the policy settings.

Note: This does not mean that the Virtual SAN datastore can successfully deploy a VM with this policy; it simply means that the Virtual SAN datastore understands the rules or requirements in the policy.

🚰 Create New VM Storage Policy									(?)
 1 Name and description 2 Rule-Sets Storage compatibility As defined, this VM storage policy is compatible with the following storage: 									
✓ 2a Rule-Set 1	C								
✓ 3 Storage compatibility	Storage Compatibility	Total Capa	al Capacity Virtual SAN Capacity 1.95 GB 811.95 GB		I Capacity	Virtual Volumes Cap	VMFS Capacit	NFS Capacity	у
✓ 4 Ready to complete	Compatible	811.95 G			в	0.00 B	0.00 B	0.00 B	
	Incompatible	52.63 TB		0.00 B		0.00 B	409.50 GB	52.23 TB	
	Compatible storage								
	I →							Q Filter	•
	Name	ne		Datacenter		Free Space	Capacity	Warnings	
	🗐 vsanDatastore		📄 VSAN	6-DC	vsan	811.57 GB	811.95 GB		
						Bac	k Next	Finish	Cancel

Figure 9.7: Storage Compatibility

At this point, you can click on next to review the settings once more, or alternatively, at this point, you can click "Finish" instead of reviewing the policy. On clicking "Finish", the policy is created.

Let's now go ahead and deploy a VM with this new policy, and let's see what effect it has on the layout of the underlying storage objects.

9.2 Deploy a New VM with the New VM Storage Policy

We have already deployed a VM back in 7.1. The steps will be identical, until we get to the point where the VM Storage Policy is chosen. This time, instead of selecting the default policy, we will select the newly created *StripeWidth=2* policy as shown below.

🔁 New Virtual Machine							(?))
1 Select creation type ✓ 1a Select a creation type	Select storage Select the datastore in which to store the configuration and disk files						
2 Edit settings	VM Storage Policy:	M Storage Policy: Datastore Default					
 2a Select a name and folder 	The following data:	Datastore Det	fault	burce that	t vou selected. Select	the destination d	atastore for the
 2b Select a compute resource 	virtual machine cor	StripeWidth=2	2		,		
✓ 2c Select storage	Name	Wol No Requ	irements Policy	ed	Free	Type	Storage DRS
2d Select compatibility	NFS-Isilon	Virtual SAN De	efault Storage Policy	GB	50.19 TB	NFS v3	
2e Select a guest OS	🗐 vsanDatastore)	1.06 TB	44.88 GB	1.06 TB	vsan	
2f Customize hardware	corkisos.csl.v	mware.com	1.77 TB	1.62 TB	159.10 GB	NFS v3	
3 Ready to complete	🗐 cs-ie-h02-scra	atch	136.50 GB	3.00 GB	133.50 GB	VMFS	
	🗐 cs-ie-h03-scra	atch	136.50 GB	3.06 GB	133.44 GB	VMFS	
	cs-ie-h04-scra	atch	136.50 GB	3.10 GB	133.40 GB	VMFS	
	4						•
	Compatibility:						
	 Compatibility 	checks succee	ded.				
					Back	Vext Finis	h Cancel

Figure 9.8: Selecting a non-default policy

And as before, the vsanDatastore should show up as the compatible datastore, and thus the one to which this VM should be provisioned if we wish to have the VM compliant with its policy.

🔁 New Virtual Machine						(?) ₩
Select creation type I a Select a creation type Edit settings	Select storage Select the datastore in which to store the configuration and disk files					
 2a Select a name and folder 2b Select a compute resource 	The following datastores are acce virtual machine configuration files	2 ssible from the dea and all of the virtua	stination resource that I disks.	t you selected. Select	the destination d	atastore for the
✓ 2c Select storage	Name	Capacity	Provisioned	Free	Туре	Storage DRS
2d Select compatibility	Compatible					
2e Select a guest OS	🗐 vsanDatastore	1.06 TB	44.88 GB	1.06 TB	vsan	
2f Customize hardware	Incompatible					
3 Ready to complete	NFS-Isilon	50.46 TB	419.21 GB	50.19 TB	NFS v3	
	corkisos.csl.vmware.com	1.77 TB	1.62 TB	159.10 GB	NFS v3	
	cs-ie-h02-scratch	136.50 GB	3.00 GB	133.50 GB	VMFS	
	🗐 cs-ie-h03-scratch	136.50 GB	3.06 GB	133.44 GB	VMFS	
	cs-ie-h04-scratch	136.50 GB	3.10 GB	133.40 GB	VMFS	*
	Compatibility:	eded.				
				Back	Next Finis	h Cancel

Figure 9.9: vsanDatastore is compatible with the policy

Let's now go ahead and examine the layout of this virtual machine, and see if the policy requirements are met; i.e. do the storage objects of this VM have a stripe width

of 2? First, ensure that the VM is compliant with the policy by navigating to VM > Manage tab > Policies, as shown here.

vmware [®] vSphere Web Client ਜ ≣				Updated at 11:35 🐧	Administrator@VSPHERE.LOCAL + Help
Navigator	Ŧ	🕆 vsan6-poc-test-vm-3 Actions 🕶			=*
Home 🔊		Getting Started Summary Monitor Man	age Related Objects		
Image: Second secon		Settings Alarm Definitions Tags Permiss Storage	ions Policies Scheduled Tasks		Edit VM Storage Policies
cs-ie-h02.ie.local		Name	VM Storage Policy	Compliance Status	Last Checked
cs-ie-h04 ie local		🛅 VM home	StripeWidth=2	✓ Compliant	09/04/2015 12:26
∰ ie-vcsa-09_1		🚈 Hard disk 1	StripeWidth=2	🗸 Compliant	09/04/2015 12:26
vsan6-poc-test-vm-3	>				
mp vsano-poc-test-Vm2					

Figure 9.10: VM is compliant with the policy

The next step is to select the Monitor tab > Policies and check the layout of the VM's storage objects. The first object to check is the VM home namespace. Select it, and then select the "Physical Disk Placement" tab at the lower part of the window. This continues to show that there is only one mirrored component, but no stripe width (which is displayed as a RAID 0 configuration). Why? The reason for this is that the VM home namespace object does not benefit from striping so it ignores this policy setting. Therefore this behavior is normal and to be expected.

vsan6-poc-test-vm-3 Ac	tions 👻				
etting Started Summary	Monitor Manage Related C	bjects			
)			
sues Performance Polici	es Tasks Events Utilization	ļ			
C 隆 👘					Q Filter
lame	VM Storage Poli	7/	Compliance Status	Last Checked	
VM home	🔓 StripeWidt	h=2	🗸 Compliant	09/04/2015	12:26
🔄 Hard disk 1	😭 StripeWidt	h=2	🗸 Compliant	09/04/2015	12:26
4					2 items
Compliance Failures Phy	ysical Disk Placement		-		
vsan6-poc-test-vm-3 - VM h	ome : Physical Disk Placement				
)— ·=					Q Filter
Туре	Component State	Host	Flash Disk Name	Flash Disk Uuid	HDD Disk Name
✓ RAID 1					
Component	Active	cs-ie-h02.ie.l	HP Serial Attached SCSI Dis	521963f0-33f5-eaaf-d2e1-f	7a21 🔲 HP Serial Attache
Component	Active	cs-ie-h01.ie.l	HP Serial Attached SCSI Dis	528ba019-e369-151e-01b3	3-26 🔲 HP Serial Attache
Witness	Active	cs-ie-h04.ie.l	HP Serial Attached SCSI Dis	52742cfb-d99d-cdc1-8ac4-	1527 🛄 HP Serial Attache

Figure 9.11: VM home namespace ignores stripe width policy setting

Now let's examine "Hard disk 1" and see if that layout is adhering to the policy. Here we can clearly see a difference. Each replica or mirror copy of the data now contains two components in a RAID 0 configuration. This implies that the hard disk storage objects are indeed adhering to the stripe width requirement that was placed in the VM Storage Policy.

10000000000000000000000000000000000000	tions 🕶				
tting Started Summary	Monitor Manage Related	Objects			
sues Performance Policie	es Tasks Events Utilization	n			
C 🍓 🗟				(0	Filter
ame	VM Storage Po	olicy	Compliance Status	Last Checked	
🔁 VM home	🔓 StripeWig	dth=2	🗸 Compliant	09/04/2015 12:26	
🔄 Hard disk 1	🔂 StripeWid	dth=2	 Compliant 	09/04/2015 12:26	
4					2 items 🔒
Compliance Failures Phy	sical Disk Placement				
Compliance Failures Phy	sical Disk Placement	ent			
Compliance Failures Phy /san6-poc-test-vm-3 - Hard	rsical Disk Placement	ent		Q	Filter
Compliance Failures Phy vsan6-poc-test-vm-3 - Hard	disk 1 : Physical Disk Placement	Host	Flash Disk Name	Q Flash Disk Uuid	Filter •
Compliance Failures Phy rsan6-poc-test-vm-3 - Hard 	disk 1 : Physical Disk Placement disk 1 : Physical Disk Placement Component State	Host	Flash Disk Name	Resh Disk Uuid	Filter
Compliance Failures Phy rsan6-poc-test-vm-3 - Hard 	disk 1 : Physical Disk Placement disk 1 : Physical Disk Placement Component State	Host	Flash Diak Name	Rash Disk Uuid	Filter •
Compliance Failures Phy vsan6-poc-test-vm-3 - Hard Type * RAID 1 * RAID 0 Component	disk 1 : Physical Disk Placement disk 1 : Physical Disk Placement Component State	Host	Flash Disk Name	Flash Disk Uuid 528ba019-e369-151e-01b3-26	Filter HDD Disk Name HDD Disk Name
Compliance Failures Phy vsan6-poc-test-vm-3 - Hard 	disk 1 : Physical Disk Placement disk 1 : Physical Disk Placem Component State Active Active	Host	Flash Disk Name	Flash Disk Uuid 528ba019-e369-151e-01b3-26 528ba019-e369-151e-01b3-26	Filter HDD Disk Name HDD Disk Name HP Serial Attached
Compliance Failures Phy vsan6-poc-test-vm-3 - Hard Type * RAID 1 * RAID 1 Component Component * RAID 0	disk 1 : Physical Disk Placement disk 1 : Physical Disk Placem Component State Active Active	Host Host CS-le-h01.le.l	Flash Disk Name	Flash Disk Uuid 528ba019-e369-151e-01b3-26 528ba019-e369-151e-01b3-26	Filter HDD Disk Name HDD Disk Name HP Serial Attached HP Serial Attached
Compliance Failures Phy vsan6-poc-test-vm-3 - Hard Type * RAID 1 * RAID 1 Component Component * RAID 0 Component Component	sical Disk Placement disk 1 : Physical Disk Placem Component State Active Active Active Active Active	Host Host CS-le-h01.le.L CS-le-h01.le.L CS-le-h01.le.L	Flash Disk Name HP Serial Attached SCSI Dis HP Serial Attached SCSI Dis HP Serial Attached SCSI Dis	Flash Disk Uuid 528ba019-e369-151e-01b3-26 528ba019-e369-151e-01b3-26 52a4acab-f622-6025-bee3-746	Filter HDD Disk Name HD Disk Name HP Serial Attached HP Serial Attached HP Serial Attached
Compliance Failures Phy vsan6-poc-test-vm-3 - Hard Type * RAID 1 * RAID 0 Component Component * RAID 0 Component Component Component	disk 1 : Physical Disk Placement disk 1 : Physical Disk Placem Component State Active Active Active Active Active Active	Host Host CS-Ie-h01.Ie.I CS-Ie-h01.Ie.I CS-Ie-h01.Ie.I CS-Ie-h03.Ie.I	Flash Disk Name HP Serial Attached SCSI Dis	Flash Disk Uuid 528ba019-e369-151e-01b3-26 528ba019-e369-151e-01b3-26 52a4acab-1622-6025-bee3-746 52a4acab-1622-6025-bee3-746	Filter HDD Disk Name HD Disk Name HP Serial Attached HP Serial Attached HP Serial Attached HP Serial Attached

Figure 9.12: Hard disks adhere to stripe width policy setting

Note that each striped component must be placed on its own physical disk. There are enough physical disks to meet this requirement in this POC. However, a request for a larger stripe width would not be possible in this configuration. Keep this in mind if you plan a POC with a large stripe width value in the policy.

It should also be noted that snapshots taken of this base disk continue to inherit the policy of the base disk, implying that the snapshot delta objects will also be striped.

One final item to note is the fact that this VM automatically has a *NumberOfFailuresToTolerate*=1, even though it was not explicitly requested in the policy. We can tell this from the RAID 1 configuration in the layout. Virtual SAN will always provide availability to VMs via the *NumberOfFailuresToTolerate* policy setting, even when it is not requested via the policy. The only way to deploy a VM without a replica copy is by placing *NumberOfFailuresToTolerate*=0 in the policy.

A useful rule of thumb for *NumberOfFailuresToTolerate* is that in order to tolerate n failures in a cluster, you require a minimum of 2n + 1 hosts in the cluster (to retain a >50% quorum with n host failures).

9.3 Add a New VM Storage Policy to an Existing VM

Virtual Machines may also have new VM Storage Policies added after they have been deployed to the Virtual SAN datastore. The configuration of the objects will be changed when the new policy is added. That may mean the adding of new components to existing objects, for example in the case where the *NumberOfFailuresToTolerate* is increased. It may also involve the creation of new objects that are synced to the original object, and once synchronized, the original object is discarded. This is typically only seen when the layout of the object changes, such as increasing the *NumberOfDiskStripesPerObject*.

In this case, we will add the new StripeWidth=2 policy to one of the VMs created in section 7 which still only has the default policy (*NumberOfFailuresToTolerate*=1, *NumberOfDiskStripesPerObject*=1, *ObjectSpaceReservation*=0) associated with it.

To begin, select the VM that is going to have its policy changed from the vCenter inventory, then select the Manage tab > Policies view. This VM should currently be compliant with the Virtual SAN Default Storage Policy. Now click on the Edit VM Storage Policies button as highlighted below.

	0	0 0								
đ	to vsan6-poc-test-vm-1 Actions →									
	Getting Started Summary Monitor Manage Related Objects									
(Settings Narm Definitions Tags Permissions Policies Scheduled Tasks									
	Storage			Edit VM Storage Policies						
	C 💁 📾			Q Filter 🔹						
	Name	VM Storage Policy	Compliance Status	Last Checked						
	VM home	🕞 Virtual SAN Default Storage Policy	✓ Compliant	09/04/2015 12:44						
	🚈 Hard disk 1 🔓 Virtual SAN Default Storage Policy		✓ Compliant	09/04/2015 12:44						

Figure 9.13: Edit VM Storage Policies

This takes you to the edit screen, where the policy can be changed.

🗗 vsan6-poc-test	-vm-1: Manage VM Stor	age Policies			?
Select a common s VM storage policy:	torage policy for all stora Virtual SAN Default Sto	ige objects or i rage Policy	use the table below to define different polic	ies for the different storage	objects.
Name	Disk Siz	e VM	I Storage Policy	Datastore	Datastore Type
🛅 VM home		Vi	rtual SAN Default Storage Policy	🗐 vsanDatastore	vsan
🏝 Hard disk 1	40.00	GB Vi	rtual SAN Default Storage Policy	🗐 vsanDatastore	vsan
Predicted impact or	storage consumption:				
				0	Cancel

Figure 9.14: Manage VM Storage Policies

Select the new VM Storage Policy from the drop-down list. The policy that we wish to add to this VM is the StripeWidth=2 policy.

🖞 vsan6-poc-test-v	m-1: Manage VM Storage Polic	ies		?
Select a common sto VM storage policy:	orage policy for all storage objec StripeWidth=2	ts or use the table below to define different • Apply to all	policies for the different stora	ge objects.
Name	Disk Size	VM Storage Policy	Datastore	Datastore Type
🗀 VM home		Virtual SAN Default Storage Policy	🗐 vsanDatastore	vsan
🏝 Hard disk 1	40.00 GB	Virtual SAN Default Storage Policy	🗐 vsanDatastore	vsan
Predicted impact on s	storage consumption:			
Storage policy h	as not changed. No change in s	storage consumption.		
				OK Cancel

Figure 9.15: Select a new VM Storage Policies

Once the policy is select, click on the "Apply to all" button as shown below to ensure the policy gets applied to all storage objects and not just the VM home namespace object. The VM Storage Policy should now appear updated for all objects.

🖞 vsan6-poc-test	-vm-1: Manage	VM Storage Poli	cies				?
Select a common s VM storage policy:	storage policy for StripeWidth=2	all storage obje	cts or use the table	below to define different poli	cies for the different stora	ige objects.	
Name		Disk Size	VM Storage Policy		Datastore	Datastore Type	
🛅 VM home			StripeWidth=2		🗐 vsanDatastore	vsan	
🏝 Hard disk 1		40.00 GB	StripeWidth=2		🗐 vsanDatastore	vsan	
Predicted impact or	n storage consur	mption:					
Applying the r 712.00 MB st 0.00 B reserved	new VM storage p orage space ed flash space	oolicy will affect t	he storage consum;	ption of the VM home. The n	ew storage consumption	will be:	
						OK Cancel	

Figure 9.16: Apply to all

Next, click OK and initiate the policy change. Now when you revisit the Monitor tab > Policies view, you should see the changes in the process of taking effect (Reconfiguring) or completed, as shown below.

anto-poc-test-vill-1 Actions	*				2
ting Started Summary Monit	itor Manage Related C	Dbjects			
ues Performance Policies T	Tasks Events Utilization]			
ب 😜					Filter -
me	VM Storage Poli	icy	Compliance Status	Last Checked	
] VM home	📑 StripeWidt	th=2	🗸 Compliant	09/04/2015 13:06	
Hard disk 1	🔓 StripeWidt	th=2	🗸 Compliant	09/04/2015 13:06	
			-		
Compliance Failures Physical san6-poc-test-vm-1 - Hard disk	I Disk Placement 1 : Physical Disk Placeme	nt			Filler
Compliance Failures Physical san6-poc-test-vm-1 - Hard disk '- *=	1 : Physical Disk Placemee	nt	Flash Disk Name		Filter
Compliance Failures Physical san6-poc-test-vm-1 - Hard disk	1 Disk Placement 1 : Physical Disk Placeme Component State	nt Host	Flash Disk Name	Resh Disk Uuid	Filter
Compliance Failures Physical san6-poc-test-vm-1 - Hard disk 	1 Disk Placement 1 : Physical Disk Placeme Component State	nt Host	Flash Disk Name	Q Flash Disk Uuid	Filter •) HDD Disk Name
Compliance Failures Physical san6-poc-test-vm-1 - Hard disk Type	I Disk Placement Physical Disk Placeme Component State Active	Host	Flash Disk Name	Q Flash Disk Uuid	Filter HDD Disk Name HP Serial Attached
Compliance Failures Physical san6-poc-test-vm-1 - Hard disk Type	I Disk Placement	Host	Flash Disk Name	Q Flesh Disk Uuid 52a4acab-1622-6025-bee3-746 528ba019-e369-151e-01b3-26	HDD Disk Name
Compliance Failures Physical san6-poc-test-vm-1 - Hard disk '* * Type * * RAID 1 * RAID 0 Component Component * RAID 0	I Disk Placement	Host	Flash Disk Name	Resh Disk Uuid 52a4acab-1622-6025-bee3-746 528ba019-e369-151e-01b3-26	Filter HDD Disk Name HD Disk Name HP Serial Attached HP Serial Attached
Compliance Failures Physical san6-poc-test-vm-1 - Hard disk Type RAID 1 * RAID 0 Component Component * RAID 0 Component * RAID 0 Component	I Disk Placement	Host i CS-le-h03.ie.l CS-le-h01.ie.l CS-le-h01.ie.l	Flash Disk Name HP Serial Attached SCSI Dis HP Serial Attached SCSI Dis HP Serial Attached SCSI Dis	Flesh Disk Uuid 52a4acab-1622-6025-bee3-746 528ba019-e369-151e-01b3-26 521963f0-33f5-eaaf-d2e1-f7a21	Filter HDD Disk Name HD Disk Name HP Serial Attached HP Serial Attached HP Serial Attached

Figure 9.17: Reconfiguring complete – new policy in effect

This is useful when you only need to modify the policy of one or two VMs, but what if you need to change the VM Storage Policy of a significant number of VMs.

That can be achieved by simply changing the policy used by those VMs. All VMs using those VMs can then be "brought to compliance" by reconfiguring their storage object layout to make them compliant with the policy. We shall look at this next.

9.4 Modify a VM Storage Policy

We will modify the StripeWidth=2 policy created earlier to include an *ObjectSpaceReservation*=10%. This means that each storage object will now reserve 10% of the VMDK size on the Virtual SAN datastore. Since all VMs were deployed with 40GB VMDKs, the reservation value will be 4GB.

The first step in this task is to note the amount of free space in the Virtual SAN datastore, so you can compare it later and confirm that each VMDK has 4GB of space reserved. Next, revisit the VM Storage Policy section that we visited previously. This can be accessed once again via the Home page.



Figure 9.18: VM Storage Policies: Stripewidth

Select StripeWidth=2 policy in the left hand column, and then the Manage tab. Select "Rule-set 1: Virtual SAN" and then click on "Edit" button on the far right.



Figure 9.19: Edit Policy

From the <Add rule> drop down list, select *ObjectSpaceReservation* as a new capability to be added to the policy.

StripeWidth=2: Edit VM Storage F	Policy	?₩
Name and description Rule-Set 1	Rule-Set 1 Select rules specific for a datastore type. Rules can be based on data services provided by data The VM storage policy will match datastores that satisfy all the rules in at least one of the rule-se	store or based on tags. ts.
storage compatibility	Rules based on data services VSAN Number of disk stripes per object Add rule> Flash read cache reservation (%) Number of failures to tolerate Force provisioning Object space reservation (%)	Storage Consumption Model A virtual disk with size 100 GB would consume: Storage space 200.00 GB Initially reserved storage space 0.00 B Reserved flash space 0.00 B
	Add another	rule set Remove this rule set
		OK Cancel

Figure 9.20: Add Object space reservation (%) as a rule to the policy

Set *ObjectSpaceReservation* to 10%. Note Storage Consumption calculations on right.

StripeWidth=2: Edit VM Storage Pe	olicy	(?) ₩
Name and description Rule-Set 1	Rule-Set 1 Select rules specific for a datastore type. Rules can be based on data services provided by data The VM storage policy will match datastores that satisfy all the rules in at least one of the rule-se	store or based on tags. ets.
Storage compatibility	Rules based on data services VSAN Number of disk stripes per object 2 Object space reservation (%) 10 C Add rule> Rules based on tags Add tag-based rule Add another	Storage Consumption Model A virtual disk with size 100 GB would consume: Storage space 200.00 GB Initially reserved storage space 20.00 GB Reserved flash space 0.00 B
		OK Cancel

Figure 9.21: Set Object space reservation to 10%

After clicking OK to make the change. The wizard will prompt you as to whether you want to reapply this change to the virtual machines using this policy manually later (default) or automatically now. It also tells you how many VMs in the environment are using the policy and will be affected by the change. Leave it at the default, which is "Manually later", by clicking Yes. This POC guide will show you how to do this manually shortly.

Reapply the VM storage policy to those 2 virtual machine(s) to make it in sync. Th might take significant time and system resources. Reapply to VMs: Manually later	is action
Reapply to VMs: Manually later	
Save changes?	

Figure 9.22: Manually later

Next, click on the Monitor tab next to the Manage tab. It will display the two VMs along with their storage objects, and the fact that they are no longer compliant with the policy. They are in an "Out of Date" compliance state as the policy has now been changed.

StripeWidth=2 Actions -			
Getting Started Summary Monitor Manage Rel	ated Objects		
VMs and Virtual Disks Storage Compatibility			O Filter
Name	Compliance Status	Last Checked	
👻 🔂 vsan6-poc-test-vm-1			
🥅 VM home	🙁 Out of Date	09/04/2015 13:06	
📇 Hard disk 1	🙁 Out of Date	09/04/2015 13:06	
👻 🚰 vsan6-poc-test-vm-3			
🥅 VM home	🙁 Out of Date	09/04/2015 12:26	
📇 Hard disk 1	🔇 Out of Date	09/04/2015 12:26	
A4			6 items 🔒 🖛

Figure 9.23: Out of Date

In order to bring the VM to a compliant state, we must manually reapply the VM Storage Policy to the objects. The button to do this action is highlighted in the previous screenshot. When this button is clicked, the following popup appears.



Figure 9.24: Reapply VM Storage Policy

When the reconfigure activity completes against the storage objects, and the compliance state is once again checked, everything should show as Compliant.

StripeWidth=2 Actions -		
Getting Started Summary Monitor Manage Related	Objects	
VMs and Virtual Disks Storage Compatibility C S T		Q Filter
Name	Compliance Status	Last Checked
👻 🔂 vsan6-poc-test-vm-1		
🔁 VM home	✓ Compliant	09/04/2015 13:30
📇 Hard disk 1	✓ Compliant	09/04/2015 13:30
👻 👘 vsan6-poc-test-vm-3		
🛅 VM home	✓ Compliant	09/04/2015 13:30
🚔 Hard disk 1	 Compliant 	09/04/2015 13:30
M6		6 items 🕞 🗧

Figure 9.25: Compliant once again

Since we have now included an *ObjectSpaceReservation* value in the policy, what you may notice is that the amount of free capacity on the Virtual SAN datastore will have reduced.

For example, the two VMs with the new policy change have 40GB storage objects. Therefore there is a 10% *ObjectSpaceReservation* implying 4GB is reserved per VMDK. 4GB per VMDK, 1 VMDK per VM, 2 VMs equals 8GB reserved space, right? However the VMDK is also mirrored, so there is a total of 16GB reserved on the Virtual SAN datastore.

Checking the Virtual SAN datastore, we can see this reflected in the free capacity.

VSAN6-Cluster Actions -		<i>E</i> *
Getting Started Summary Mo	nitor Manage Related Objects	
Settings Scheduled Tasks Ala	rm Definitions Tags Permissions	
	Virtual SAN is Turned ON	Edit
✓ Services	Add disks to storage Manual	
vSphere DRS vSphere HA	Resources	
- Virtual SAN	Hosts	4 hosts
General	Flash disks in use	4 of 4 eligible
Disk Management	Data disks in use	8 of 25 eligible
Fault Domains	Total capacity of Virtual SAN datastore	1.06 TB
Health	Free capacity of Virtual SAN datastore	1.03 TB
- Configuration	Network status	✓ Normal
General		
Licensing	On-disk Format Version	Upgrade
VMware EVC	Virtual SAN format version Version 2 ((latest)
VM/Host Groups	Disks with older version 🥥 0 of 12	2
VM/Host Rules	Support Assistant	United Overand Duration to Overlag Descent
VM Overrides	Support Assistant	Upload Support Bundles to Service Request
Host Options	Last upload time -	
Profiles		

Figure 9.26: ObjectSpace Reservation consuming capacity

This completes the "VM Storage Policies" section of this POC. You should now appreciate how powerful VM Storage Policies are, and how characteristics of the underlying storage can be assigned to virtual machines on a granular per VMDK basis while using a single Virtual SAN datastore.

10. Virtual SAN Monitoring

When it comes to monitoring Virtual SAN, there are a number of areas that need particular attention. In no particular order, these are considerations when it comes to monitoring Virtual SAN:

- Monitor the Virtual SAN Cluster
- Monitor Virtual Devices in the Virtual SAN Cluster
- Monitor Physical Devices in Virtual SAN Datastores
- Monitor Resynchronization & Rebalance Operations in the Virtual SAN Cluster
- Examine Default Virtual SAN Alarms
- Triggering Alarms based on Virtual SAN VMkernel Observations Alarms

10.1 Monitor the Virtual SAN Cluster

The first item to monitor is the overall health of the cluster. The Manage > General view gives you a good idea as to whether all the flash and capacity devices that you expect to be in use are in fact in use. It also shows whether the network status is normal or not. Finally, it is a good indicator as to whether or not the expected capacity of the Virtual SAN datastore is correct, and if there are any capacity concerns looming.

VSAN6-Cluster Actions -			=*
Getting Started Summary Mo	nitor Manage Related Objects		
Settings Scheduled Tasks Ala	rm Definitions Tags Permissions		
44	Virtual SAN is Turned ON		Edit
✓ Services	Add disks to storage Manual		
vSphere DRS vSphere HA	Resources		
✓ Virtual SAN	Hosts	4 hosts	
General	Flash disks in use	4 of 4 eligible	
Disk Management	Data disks in use	8 of 25 eligible	
Fault Domains	Total capacity of Virtual SAN datastore	1.06 TB	
Health	Free capacity of Virtual SAN datastore	1.03 TB	
- Configuration	Network status	✓ Normal	
General Licensing	On-disk Format Version		Upgrade
VMware EVC	Virtual SAN format version Version 2	(latest)	
VM/Host Groups	Disks with older version 🔮 0 of 1	12	
VM/Host Rules VM Overrides	Support Assistant		Upload Support Bundles to Service Request
Host Options	Last upload time -		
Profiles			

Figure 10.1: General view

Virtual SAN Health checks will display even more information regarding health and should be enabled as part of any Virtual SAN 6.1 POC.

10.2 Monitor Virtual Devices in the Virtual SAN Cluster

To monitor the virtual devices, navigate to Monitor > Virtual SAN > Virtual Disks. This will list the objects associated with each virtual machine, such as the VM home namespace and the hard disks. One can also see the policy, compliance state and health of an object. If one selects an object, physical disk placement and compliance failures are displayed in the lower half of the screen.

VSAN6-Cluster Actions -									E
Getting Started Summary Mo	nitor Manage Related C	Dbjects							
Issues Profile Compliance Per	formance Tasks Events	Resource Rese	rvation Virt	ual SAN	vSphere DRS	Utilization			
44	Virtual Disks								
Health									
Proactive Tests								Q Filter	•
Physical Disks	Name	VM Storage Poli	PY	Complia	noe Status	Last Checked		Operational State	
Virtual Disks	👻 🔂 vsan6-poc-test								
Resyncing Components	🔁 VM home	StripeWidt	h=2	V Co	mpliant	13/04/2015 10:2	:9	Healthy	
	🏝 Hard disk 1	StripeWidt	h=2	V Co	mpliant	13/04/2015 10:2	9	 Healthy 	
	▼ b vsan6-poc-test								
	VM home	Virtual SAN	Default	✓ Co	mpliant	13/04/2015 10:2	9	 Healthy 	
	Hard disk 1	E Virtual SAM	Default	✓ Co	mpliant	13/04/2015 10:2	:9	 Healthy 	
	vsano-poc-test	Christen 14/1 dt	h-0		maliant	10/04/0045 10/0	0		
	Number 1	StripeWidt	n=2	V Co	mpliant	13/04/2015 10:2	.9	Healthy	
		EL Subewidt	11-2	V C0	Inpliant	13/04/2015 10.2	.9	V Healury	
	AA							9 ite	ms 🔍 -
		_			_				
	Physical Disk Placeme	ent Compliance	Failures						
	vsan6-poc-test-vm-1 - \	/M home : Physic	al Disk Plac	ement					
)— ·=							O Filtor	-
	= Turne (Cr	In a second Cipita	last		Flash Disk Name		Elash Disk I		
	Witnocc	Activo	nost El colio bú	2 io l	Hash Disk Name	achad SCSI Dis	Flash Disk u	h f622 6025 haa2 746	
	- BAD 1	Active	CS-IE-IIU	5.IE.I	- HF Selial Au	actied SCSI DIS	52848C8	0-1022-0020-0663-740	
	Component	Active	Cs-ie-b0	2 ia l	HP Serial Att	ached SCSI Dis	521063fr	L33f5_papf_d2p1_f7p21	- HP
	Component	Active	CS-ie-h0	1 ie l	HP Serial Att	ached SCSI Dis	528ba01	9-e369-151e-01b3-26	
	Component	Active		1.10.1		actied ocor bis	5200401	3-6303-1316-0103-20	
	4								•
	M							4 item	s 📑 🕶

Figure 10.2: Virtual Disks view

All objects should be compliant and healthy. All components in the physical disk placement view should appear as "Active".

10.3 Monitor Physical Devices in the Virtual SAN Cluster

In the same monitor > Virtual SAN view, physical disks can also be displayed. Where this view is very useful is when you wish to see which objects reside on a particular physical disk. In the view below, one of the magnetic disks is selected and in the lower half of the screen, the objects that have components residing on that physical disk are displayed.

USAN6-Cluster Actions -							=*	
Getting Started Summary Monitor Manage Related Objects								
Issues Profile Compliance Per	rformance Tasks Events Resource	e Reservation Virtual SAN vSphere	e DRS Ut	tilization				
++ Physical Disks								
Health	Health 🥥 🔘 '= '≡ C				Q Filter			
Proactive Tests	Name	Name Disk Group				Capacity	Used (*	
Physical Disks	👻 📱 cs-ie-h03.ie.local							
Virtual Disks	F HP Serial Attached SCS	📄 Disk	group (020008000060	Flash	186.28 GB	0.00		
Resyncing Components	🚐 HP Serial Attached SCSI	I Disk (naa.600508b1001c2b7a3d	🗃 Disk	group (020008000060	HDD	136.70 GB	8.02	
	HP Serial Attached SCSI Disk (naa.600508b1001cb11f32			group (020008000060	HDD	136.70 GB	8.00	
	✓							
	HP Serial Attached SCSI Disk (naa.600508b1001c61cedd			group (020008000060	Flash	186.28 GB	0.00	
	HP Serial Attached SCSI	📰 Disk	group (020008000060	HDD	136.70 GB	6.75		
	HP Serial Attached SCSI Disk (naa.600508b1001c388c92			group (020008000060	HDD	136.70 GB	368.	
	F HP Serial Attached SCS	📰 Disk	group (020008000060	Flash	186.28 GB	0.00		
	HP Serial Attached SCSI	Disk (naa.600508b1001cb2234d	🚍 Disk	group (020008000060	HDD	136.70 GB	2.71 *	
	M 16 items 🔍 🗸							
	=							
	HP Serial Attached SCSI Disk (naa.6	Serial Attached SCSI Disk (naa.600508b1001c2b7a3d39534ac6beb92d): VM Objects on Disk						
	Q Filter •							
	Parent VM	VM Object Ob		Object Type	VM Storage Policy			
	vsan6-poc-test-vm-3	📇 Hard disk 1	Virtual Disk		5 StripeWidth=2			
	vsan6-poc-test-vm-1	🚈 Hard disk 1		Virtual Disk	5	StripeWidth=2		

Figure 10.3: Physical Disks view

10.4 Monitor Resynchronization and Rebalance Operations

Another very useful view in this Monitor > Virtual SAN tab is "Resyncing components". This will display any rebuilding or rebalancing operations that might be taking place on the cluster. For example, if there was a device failure, resyncing or rebuilding activity could be observed here. Similarly, if a device was removed or a host failed, and the CLOMd (Cluster Logical Object Manager daemon) timer expired (60 minutes by default), rebuilding activity would also be observed in this case.

With regards to rebalancing, Virtual SAN attempts to keep all physical disks at less than 80% capacity. If any physical disks' capacity passes this threshold, Virtual SAN will move components from this disk to other disks in the cluster in order to rebalance the physical storage.
By default, there should be no resyncing activity taking place on the Virtual SAN Cluster, as shown below. Resyncing activity usually indicates:

- (a) a failure of a device or host in the cluster
- (b) a device has been removed from the cluster
- (c) a physical disk have greater than 80% of its capacity consumed
- (d) a policy change has been implemented which necessitates a rebuilding of a VM's object layout. In this case, the new object layout is created, synchronized to the original object, and then the original object is discarded.

VSAN6-Cluster Actions -						=*		
Getting Started Summary Monito	Manage Related Ob	ojects						
Issues Profile Compliance Perform	Issues Profile Compliance Performance Tasks Events Resource Reservation Virtual SAN vSphere DRS Utilization							
Health Ri Proactive Tests Mi Physical Disks C	Health Resyncing components Health Resyncing components view displays the status of virtual machine objects that are currently being resynchronized in the Virtual SAN cluster. Monitoring object resynchronization is not available for clusters containing only hosts with version earlier than ESXI 6.0 Physical Disks							
Virtual Disks	Resyncing components							
Resyncing Components	Bytes left to resync							
E	ETA to compliance							
j								
N	lame		VM Storage Policy	Host	Bytes Left to Resync	ETA		
				This list is empty.				

Figure 10.4: Resyncing components

10.5 Default Virtual SAN Alarms

There are at least 56 Virtual SAN alarms pre-defined in vCenter server 6.0u1. Some are shown here, and the majority relate to Virtual SAN Health issues:

etting Started Summary Monitor Mar	age Related Obje	ects			
	ago				
ettings Scheduled Tasks Alarm Definitio	ns Tags Permis	sions			
+ ×				Q san	•
Name		Defined In			
😽 Expired Virtual SAN time-limited license	9	10.15	6.130.20		
🙀 Virtual SAN Health Alarm 'Fault domain	number check'	10.15	6.130.20		
😽 Virtual SAN Health Alarm 'Hosts discon	nected from VC'	10.15	i6.130.20		
😽 Virtual SAN Health Alarm 'SCSI Control'	er on Virtual SA	10.15	i6.130.20		
ᡖ Virtual SAN Health Alarm 'Limits health'		🕝 10.15	6.130.20		
ᡖ Virtual SAN Health Alarm 'Advanced Virt	ual SAN configur	🕝 10.15	6.130.20		
📷 Virtual SAN Health Alarm 'Hosts with Vi	rtual SAN disabl	🕝 10.15	6.130.20		
📷 Virtual SAN Health Alarm 'ESX Virtual S/	AN Health servic	🕝 10.15	6.130.20		
📷 Virtual SAN Health Alarm 'Hosts withou	t configured unic	🕝 10.15	6.130.20		
ᡖ Virtual SAN Health Alarm 'MTU check (p	ing with large pa	🕝 10.15	6.130.20		
📷 Virtual SAN Health Alarm Witness host	with non-existin	🝘 10.15	6.130.20		
📷 Virtual SAN Health Alarm 'Component r	netadata health'	🕝 10.15	6.130.20		
📷 Virtual SAN Health Alarm Virtual SAN H	CL DB up-to-date'	🕝 10.15	6.130.20		
📷 Virtual SAN Health Alarm 'Hosts with co	nnectivity issues'	🔁 10.15	6.130.20		
📷 Virtual SAN Health Alarm 'Host issues r	etrieving hardwa	🙋 10.15	6.130.20		
📷 Virtual SAN Health Alarm 'Virtual SAN H	ealth Service up	🕝 10.15	6.130.20		
📷 Virtual SAN Health Service Alarm for Ov	erall Health Sum	🕝 10.15	6.130.20		
📷 Virtual SAN Health Alarm 'All hosts have	e a Virtual SAN v	🕝 10.15	6.130.20		
📷 Virtual SAN Health Alarm 'Congestion'		🕝 10.15	6.130.20		
📷 Virtual SAN Health Alarm 'Stretched clu	ster health'	🕝 10.15	6.130.20		
📷 Virtual SAN Health Alarm 'Controller Re	lease Support'	🕝 10.15	6.130.20		
📷 Virtual SAN Health Alarm 'Current cluste	er situation'	🔁 10.15	6.130.20		
📷 Virtual SAN Health Alarm Virtual SAN cl	uster partition'	🕝 10.15	6.130.20		
📷 Virtual SAN Health Alarm 'Data health'		🙋 10.15	6.130.20		
📷 Virtual SAN Health Alarm 'After 1 additio	nal host failure'	🕝 10.15	6.130.20		
📷 Virtual SAN Health Alarm 'Software stat	e health'	🕝 10.15	6.130.20		
📷 Virtual SAN Health Alarm 'Network heal'	th'	🕝 10.15	6.130.20		
📷 Virtual SAN Health Alarm 'All hosts have	e matching multi	🕝 10.15	6.130.20		
📷 Virtual SAN Health Alarm 'Active multica	st connectivity ch	7 10.15	6.130.20		
🔜 Virtual SAN Haalth Alarm 'Overall dieke.	health'	<i>i</i> 10 15	6 130 20	56 of 112 ite	ams
				30 01 112 116	
/SAN6-Cluster Actions 👻					
ting Started Summary Monitor Manage Re	lated Objects				
ttings Scheduled Tasks Alarm Definitions Tags	Permissions				
			Errora occurred	the disk(a) of a listent CARL	
	Q san	•	Errors occurred on	uie uisk(s) oi a virtuai SAN hõst	E
Expired Virtual SAN time-limited license	ie-vcsa-09.ie.local		Name	Errors occurred on the dis	sk(s) of a Virtual SAN host
Registration/unregistration of a VASA vendor	ie-vcsa-09.ie.local		Defined in	ie-vcsa-09.ie.local	
Host flash capacity exceeds the licensed limit	🛃 ie-vcsa-09.ie.local		Description	Derault alarm that monito host disk(s) in the Virtual	s wnetner there are errors of SAN cluster.
Expired Virtual SAN license	ie-vcsa-09.ie.local		Monitor type	Host	
b Errors occurred on the disk(s) of a Virtual SA	🛃 ie-vcsa-09.ie.local		Enabled	Yes	

Figure 10.5: Alarm definitions

10.7 Monitor Virtual SAN with VSAN Observer

The VMware VSAN Observer is a performance monitoring and troubleshooting tool for Virtual SAN. The tool is launched from the Ruby vSphere Console (RVC) and can be utilized for monitoring performance statistics for Virtual SAN live mode or offline. When running in live mode, a web browser can be pointed at vCenter Server to see live graphs related to the performance of Virtual SAN.

The utility can be used to understand Virtual SAN performance characteristics. The utility is intended to provide deeper insights of Virtual SAN performance characteristics and analytics. VSAN Observer's user interface displays performance information of the following items:

- Host level performance statistics (client stats)
- Statistics of the physical disk layer
- Deep dive physical disks group details
- CPU Usage Statistics
- Consumption of Virtual SAN memory pools
- Physical and In-memory object distribution across Virtual SAN Clusters

The VSAN Observer UI depends on some JavaScript and CSS libraries (JQuery, d3, angular, bootstrap, font-awesome) in order to successfully display the performance statistics and other information. These library files are accessed and loaded over the Internet at runtime when the VSAN Observer page is rendered. The tool requires access to the libraries mentioned above in order to work correctly. This means that the vCenter Server requires access to the Internet. However with a little work beforehand, VSAN Observer can be configured to work in an environment that does not have Internet access.

Further discussion on VSAN Observer is outside the scope of this POC Guide. For those interested in learning more about Virtual SAN Observer, refer to the <u>VMware Virtual</u> <u>SAN Diagnostics and Troubleshooting Reference Manual</u> and <u>Monitoring VMware</u> <u>Virtual SAN with VSAN Observer</u>.

11. Performance Testing

Performance testing is an important part of evaluating any storage solution. Setting up a desirable test environment could be challenging, and customers may do it differently. Customers may also select from a variety of tools to run workloads, or choose to collect data and logs in different ways. These all add complexity to troubleshoot performance issues claimed by customers, and lengthen the evaluation process.

Virtual SAN Performance will depend on what devices are in the hosts (SSD, magnetic disks), on the policy of the virtual machine (how widely the data is spread across the devices), the size of the working set, the type of workload, and so on.

A major factor for virtual machine performance is the virtual hardware: how many virtual SCSI controllers, VMDKs, outstanding I/O and how many vCPUs can be pushing I/O. Use a number of VMs, virtual SCSI controllers and VMDKs for maximum performance.

Virtual SAN's distributed architecture dictates that reasonable performance is achieved when the pooled compute and storage resources in the cluster are well utilized. This usually means a number of VMs each running the specified workload should be distributed in the cluster and run in a consistent manner to deliver aggregated performance. Virtual SAN also depends on VSAN Observer for detailed performance monitoring and analysis, which as a separate tool is easy to become an afterthought of the testing.

11.1 Use VSAN Observer

Virtual SAN ships with a performance-monitoring tool called VSAN Observer. It is accessed via RVC – the Ruby vSphere Console. If you're planning on doing any sort of performance testing, plan on using VSAN Observer to observe what's happening.

Reference VMware Knowledgebase Article 2064240 for getting started with VSAN Observer – <u>http://kb.vmware.com/kb/2064240</u>. See detailed information in <u>Monitoring VMware Virtual SAN with VSAN Observer</u>.

11.2 Performance Considerations

There are a number of considerations you should take into account when running performance tests on Virtual SAN.

11.2.1 Single vs. Multiple Workers

Virtual SAN is designed to support good performance when many VMs are distributed and running simultaneously across the hosts in the cluster. Running a single storage test in a single VM won't reflect on the aggregate performance of a Virtual SAN- enabled cluster. Regardless of what tool you are using – IOmeter, VDbench or something else – plan on using multiple "workers" or I/O processors to multiple virtual disks to get representative results.

11.2.2 Working Set

For the best performance, a virtual machine's working set should be mostly in cache. Care will have to be taken when sizing your Virtual SAN flash to account for all of your virtual machines' working sets residing in cache. A general rule of thumb is to size cache as 10% of your consumed virtual machine storage (not including replica objects). While this is adequate for most workloads, understanding your workload's working set before sizing is a useful exercise. Consider using VMware Infrastructure Planner (VIP) tool to help with this task – <u>http://vip.vmware.com</u>.

11.2.3 Sequential Workloads versus Random Workloads

Sustained sequential write workloads (such as VM cloning operations) run on Virtual SAN will simply fill the cache and future writes will need to wait for the cache to be destaged to the spinning magnetic disk layer before more I/Os can be written to cache, so performance will be a reflection of the spinning disk(s) and not of flash. The same is true for sustained sequential read workflows. If the block is not in cache, it will have to be fetched from spinning disk. Mixed workloads will benefit more from Virtual SAN's caching design.

11.2.4 Outstanding IOs

Most testing tools have a setting for Outstanding IOs, or OIO for short. It shouldn't be set to 1, nor should it be set to match a device queue depth. Consider a setting of between 2 and 8, depending on the number of virtual machines and VMDKs that you plan to run. For a small number of VMs and VMDKs, use 8. For a large number of VMs and VMDKs, consider setting it lower.

11.2.5 Block Size

The block size that you choose is really dependent on the application/workload that you plan to run in your VM. While the block size for a Windows Guest OS varies between 512 bytes and 1MB, the most common block size is 4KB. But if you plan to run SQL Server, or MS Exchange workloads, you may want to pick block sizes appropriate to those applications (they may vary from application version to application version). Since it is unlikely that all of your workloads will use the same block size, consider a number of performance tests with differing, but commonly used, block sizes.

11.2.6 Cache Warm up Considerations

Flash as cache helps performance in two important ways. First, frequently read blocks end up in cache, dramatically improving performance. Second, all writes are committed to cache first, before being efficiently destaged to disks – again, dramatically improving performance.

However, data still has to move back and forth between disks and cache. Most realworld application workloads take a while for cache to "warm up" before achieving steady-state performance.

11.2.7 Number of Magnetic Disk Drives in Hybrid Configurations

In the getting started section, we discuss how disk groups with multiple disks perform better than disk groups with fewer, as there are more disk spindles to destage to as well as more spindles to handle read cache misses. Let's look at a more detailed example around this.

Consider a Virtual SAN environment where you wish to clone a number of VMs to the Virtual SAN datastore. This is a very sequential I/O intensive operation. We may be able to write into the SSD write buffer at approximately 200-300 MB per second. A single magnetic disk can maybe do 100MB per second. So assuming no read operations are taking place at the same time, we would need 2-3 magnetic disks to match the SSD speed for destaging purposes.

Now consider that there might also be some operations going on in parallel. Let's say that we have another Virtual SAN requirement to achieve 2000 read IOPS. Virtual SAN is designed to achieve a 90% read cache hit rate (approximately). That means 10% of all reads are going to be read cache misses; for example, that is 200 IOPS based on our requirement. A single magnetic disk can perhaps achieve somewhere in the region of 100 IOPS. Therefore an additional 2 magnetic disks will be required to meet this requirement.

If we combine the destaging requirements and the read cache misses described above, your Virtual SAN design may need 4 or 5 magnetic disks per disk group to satisfy your workload.

11.2.8 Striping Considerations

One of the VM Storage Policy settings is *NumberOfDiskStripesPerObject*. That allows you to set a stripe width on a VM's VMDK object. While setting disk striping values can sometimes increase performance, that isn't always the case.

As an example, if a given test is cache-friendly (e.g. most of the data is in cache), striping won't impact performance significantly. As another example, if a given VMDK is striped across disks that are busy doing other things, not much performance is gained, and may actually be worse.

11.2.9 Guest File Systems Considerations

Many customers have reported significant differences in performance between different guest file systems and their settings; for example, Windows NTFS and Linux. If you are not getting the performance you expect, consider investigating whether it could be a guest OS file system issue.

11.2.10 Performance during Failure and Rebuild

When Virtual SAN is rebuilding one or more components, application performance can be impacted. For this reason, always check to make sure that Virtual SAN is fully rebuilt and that there are no underlying issues prior to testing performance. Verify there are no rebuilds occurring before testing with the following RVC command, which we discussed earlier:

- vsan.check_state
- vsan.disks_stats
- vsan.resync_dashboard

11.3 Performance Testing Option 1: Virtual SAN Health Check

Virtual SAN Health Check comes with its own Storage Performance Test. This negates the need to deploy additional tools to test the performance of your Virtual SAN environment. To run the storage performance test is quite simple; navigate to the cluster's Monitor tab > Virtual SAN > Proactive Tests, select Storage Performance Test, then click on the Go arrow highlighted below.

	Proactive Tests		
Health			O Filter
Proactive Tests	Name	Last Run Result	Last Run Time
Physical Disks	VM creation test	N/A	N/A
Virtual Disks	Multicast performance test	N/A	N/A
Resyncing Components	Storage performance test 🕕	N/A	N/A
	84		3 items 📑 🗸

Figure 11.1: Storage Performance Test

A popup is then displayed, showing the duration of the test (default 10 minutes) along with the type of workload that will be run. The user can change this duration, for example, if a burn-in test for a longer period of time is desired.

There are a number of different workloads that can be chosen from the drop-down menu.

			l	Q Filter		
Proactive lests	Name	Last Run Result	Last Run Time			
Physical Disks	VM creation test	N/A	N/A			
Virtual Disks	Multicast performance test	N/A	N/A			
Resyncing Components	Storage performance test	N/A	N/A			
	M			3 items 🔒 🔻		
Run Storage p	erformance test					
Dura under a d			durber.			
Run workload t	or at least 5 minutes to get representative res	suits. Run for hours to test stability of the	cluster			
Duration:	10 Minutes					
Workload:	Basic sanity test, focus on F	lash cache laver				
			OK Cancel			
			إيرىيىت تىپىي			
		Pup the test to view it	a dotalla horo			
		Run the test to view it	o uctano nero.			

Figure 11.2: Storage Performance Test duration and workload

To learn more about the test that is being run, click on the (i) symbol next to the workload. This will describe the type of workload that the test will initiate.

When the test completed, the Storage Load Test results are displayed, including test name, workload type, IOPS, throughput, average latency and maximum latency. Keep in mind that a sequential write pattern will not benefit from caching, so the results that are shown from this test are basically a reflection of what the capacity layer (in this case, the magnetic disks) can do.

VSAN6-Cluster Actions -									E
Getting Started Summary	Ionitor Manage Related	Dbjects							
Issues Profile Compliance F	orformanco Hoalth Tacks	Events Resource Reconsticu	A Mittual CAN LICob	oro DBS USaboro HA Utiliza	100				
issues Frome compliance F	enomiance meanin rasks	Events Resource Reservation	Vintual SPAN VSpri	ere DRS Vopriere HA Ouriza	1011				
**	Proactive Tests								
Health								O Filler	
Proactive Tests	News			Last Due Danuk		Last Due T	1	(a rinter	
Physical Disks	VM creation test			N/A		List Run T	ime		
Virtual Disks	Multicast notformance for	ince test							
Resyncing Components	Muticast performance test				Monday	May 18: 2015 4:01:25 PM			
	otorage performance test			0 1 45564		monday,	May 10, 2010 4.01.201 M		
	A4							3 itor	
	101							5 1101	
	Storage performance test	t Dotaile							
	VSAN hosts Storage Perfor	rmance Test Result							
	Used.	Wedland Tree	Martin Disk Museling	Duration (and	1070	Theory is a 10 la	Automa I also automa	Mandanian Latence (and	
	Host	Workload Type	VMDK Disk Number	Duration (sec)	IOPS	Throughput MB/s	Average Latency (ms)	Maximum Latency (ms)	
	Cs-le-h02.le.local	Basic sanity test, focus o	1	60	922	3.60	1.61	44.10	
	Co-le-h02.le.local	Basic sanity test, focus o	2	60	919	3.59	1.62	42.20	
	cs-ie-h02.ie.local	Basic sanity test, focus o	2	60	1059	4.12	1.04	57.96	
	Cs-ie-h02.ie.local	Basic sanity test, focus o	4	60	449	4.15	3.05	85.16	
	cs-ie-b02 ie local	Basic sanity test, focus o	5	60	451	1.76	3.02	73.32	
	cs-ie-b02 ie local	Basic sanity test focus o	6	60	998	3.90	1.46	51.68	
	cs-ie-h02.ie.local	Basic sanity test, focus o	7	60	999	3.90	1.46	54.04	
	cs-ie-h02.ie.local	Basic sanity test, focus o	8	60	983	3.84	1.49	56.33	
	cs-ie-h02.ie.local	Basic sanity test, focus o	9	60	999	3.90	1.46	61.53	
	cs-ie-h03.ie.local	Basic sanity test, focus o	0	60	896	3.50	1.70	44.07	_
	cs-ie-h03.ie.local	Basic sanity test, focus o	1	60	881	3.44	1.74	44.13	
	cs-ie-h03.ie.local	Basic sanity test, focus o	2	60	913	3.57	1.63	52.12	
	cs-le-h03.le.local	Basic sanity test, focus o	3	60	887	3.46	1.72	35.82	
	cs-ie-h03.ie.local	Basic sanity test, focus o	4	60	873	3.41	1.75	63.67	
	cs-ie-h03.ie.local	Basic sanity test, focus o	5	60	993	3.88	1.47	42.28	
	cs-ie-h03.ie.local	Basic sanity test, focus o	6	60	887	3.47	1.71	44.02	
	cs-ie-h03.ie.local	Basic sanity test, focus o	7	60	447	1.75	3.04	68.41	
	cs-ie-h03.ie.local	Basic sanity test, focus o	8	60	946	3.70	1.58	41.15	
	cs-le-h03.le.local	Basic sanity test, focus o	9	60	903	3.53	1.68	62.78	

Figure 11.3: Virtual SAN Cluster Storage Load Test results

The proactive test could then be repeated with different workloads

As before, when the test completes, the results are once again displayed. You will notice a major difference in results when the workload can leverage the caching layer versus when it cannot.

11.4 Performance Testing Option 2: HClbench

In a hyper-converged architecture, each server is intended to support both many application VMs, as well as contribute to the pool of storage available to applications. This is best modeled by invoking many dozens of test VMs, each accessing multiple stored VMDKs. The goal is to simulate a very busy cluster. Unfortunately, popular storage performance testing tools do not directly support this model. As a result performance testing a hyper-converged architecture such as Virtual SAN presents a different set of challenges. To accurately simulate workloads of a production cluster it is best to deploy multiple VMs dispersed across hosts with each VM having multiple disks. In addition, the workload test needs to be run against each VM and disk simultaneously.

To address the challenges of correctly running performance testing in hyperconverged environments, VMware has created a storage performance testing automation tool called HCIbench that automates the use of the popular Vdbench testing tool. Users simply specify the parameters of the test they would like to run, and HCIbench instructs Vdbench what to do on each and every node in the cluster.

HCIbench aims to simplify and accelerate customer Proof of Concept (POC) performance testing in a consistent and controlled way. The tool fully automates the end-to-end process of deploying test VMs, coordinating workload runs, aggregating test results, and collecting necessary data for troubleshooting purposes. Evaluators choose the profiles they are interested in; HCIbench does the rest quickly and easily.

This section provides an overview and recommendations for successfully using HCIbench. For complete documentation and use procedures, refer to the HCIbench Installation and User guide which is accessible from the download directory.

11.4.1 Where to Get HClbench

HCIbench and complete documentation can be downloaded from the following location: <u>HCIbench Automated Testing Tool.</u>

This tool is provided free of charge and with no restrictions. Support will be provided solely on a best-effort basis as time and resources allow, by the <u>VMware</u> <u>Virtual SAN Community Forum</u>.

11.4.2 Deploying HClbench

Step 1 – Deploy the OVA

To get started, you deploy a single HCIbench appliance called *HCIbench.ova*. The process for deploying the HCIbench OVA is no different from deploying any other OVA.

Step 2 – HCIbench Configuration

After deployment, navigate to <u>http://Controller_VM_IP:8080/</u> to start configuration and kick off the test.

There are three main sections in this configuration file:

• vSphere Environment Information

In this section, all the parameters are required except for the **Network Name** field. You must provide the vSphere environment information where the Virtual SAN Cluster is configured, including vCenter IP address, vCenter credential, name of the datacenter, name of the Virtual SAN Cluster, and name of the Datastore.

- The **Network Name** parameter defines which network the Vdbench Guest VMs should use. The default value is VM Network.
- If DHCP serviceses are not available, the Enable DHCP Service on the Network parameter allows user to enable DHCP service on the network which "HCIBench Internal Network" mapped on.
- The Datastore Name parameter specifies the datastores to be tested. All VM data will be deployed on this datastore. For the purposes of this guide the Virtual SAN datastore should be specified. Testing multiple datastores in parallel is also supported. You can enter the datastore names, one per line. In this case, virtual machines are distributed evenly across the datastores. For example, if you enter two datastores and 100 virtual machines, 50 virtual machines will be deployed on each datastore.

Performance Automation Tool Configuration Page

vSphere Environment Informati	on
vCenter Hostname/IP	
10.156.169.96	*
vCenter Username	
administrator@vsphere.local	*
vCenter Password	
	*
Datacenter Name	
	1
Lab	*
Cluster Name	
VSAN	*
)
Network Name	
VM Network-1284)
Enable DHCP Service on the N	letwork
Datastore Name	
vsanDatastore	
nfsDatastore	
	*
	1

DHCP Service could be enabled if the specified Network doesn't have DHCP Server (OPTIONAL), if checked, HCIBench Internal Network needs to be mapped on the same Network

Figure 11.4: Performance Automation Tool Configuration

Step 3 - Virtual SAN Cluster Hosts Information

Configuring the Cluster Hosts information is optional. If this parameter is lift unchecked HCIbench will create a VDbench Guest VM, then clone it to all hosts in the

Virtual SAN Cluster in a round-robin fashion. The naming convention of Vdbench Guest VMs deployed in this mode is "vdbench-vc-<*DATASTORE_NAME*>-<#>".

If this option is checked, each hosts you wish to deploy HClbench guest VMs on must be manually added to the Hosts section. As a best practice it is recommended to leave the Cluster host information parameter unchecked and let HClbench evenly distribute virutal machines on each host.

Virtual SAN Clu	ster Hosts I	Informatio	n	
Directly Deploy	on Hosts			
🖉 Deploy on F	losts			
Hosts				
10.156.28.21				
10.156.28.22				
10.156.28.23				
10.156.28.24				
Host Username	ļ			
root				
Host Password	I			

Figure 11.5: Virtual SAN Cluster Hosts Information

Step 4 - VDbench Guest VM Specification

In this section, the only required parameter is **Number of VMs** that specifies the total number of Vdbench Guest VMs to be deployed for testing. If you enter multiple datastores, these VMs are deployed evenly on the datastores. The **Number of Data Disk** and **Size of Data Disk** parameters are optional:

- The **Number of Data Disk** parameter specifies how many VMDKs to be tested are added to each Vdbench Guest VM.
- The Size of Data Disk parameter specifies the size (GB) of each VMDK to be tested. The total number of simulated workload instances is Number of VM * (times) Number of Data Disk.

The default value of both parameters is 10.

NOTE: Prior to setting the number and size of each data disk careful consideration should be given to ensure that there is sufficeent compute and storage resrouces to supporte the target workload. In addition, the cumulative size of all test VMs should not exceed the size of cache available on the cluster as a whole. You should take a carefule sizing exercise to make sure there is sufficient compute and storage resoruces to support the target level of workload instances.

		_
10		8
Numbor	e of Data Dick	
Number	r ur Data Disk	
2	r ur Data Disk	
2	r ui Data Disk	
2		

Figure 11.5: Vdbench Guest VM Specification

Step 5 - Download and add vdbench zip file, and add parameter file

Once this is done, users need to provide access to the *vdbench* tool. Due to licensing issues, we are not allowed to distribute the vdbench benchmarking tool, so it needs to be downloaded from Oracle if you do not have it already. There is a link provided to the Oracle website to down the vdbench zip file, but you will need to have an account on Oracle's site to access it. Once the vdbench zip file has been downloaded locally, you must then uploaded to the appliance. The next part of the setup is to generate a vdbench parameter file, which has information such as I/O size, R/W ratio and whether the I/O should be random or sequential in nature. You should also state how long you want the test to run (3600 seconds = 1 hour below), as well as whether you want to *dd* the storage first (initialize it). Finally, decide if you want the benchmark VMs cleaned up once the test completes. Save the configuration. To make sure that everything is OK, run the validate test. This will verify that all the configuration parameters are correct, and will state whether it is OK to start the test.

Vdbench Testing Configuration	
Test Name	
DemoTest	
Select a Vdbench parameter file	
vdb-2vmdk-100ws-16k-60rdp	
Refresh Delete Upload and use a Vdbench parameter OPERATION WILL OVERWRITE YOUR S Browse No file selected.	file for testing. (THIS ELECTION ABOVE)
Generate Vdbench Parameter File by Y	ourself
Generate	

Figure 11.6: Vdbench Testing Configuration

11.4.3 Considerations for Defining Test Workloads

Working set

Working set is one of the most important factors for correctly running performance test and obtaining accurate results. For the best performance, a virtual machine's working set should be mostly in cache. Care will have to be taken when sizing your Virtual SAN flash to account for all of your virtual machines' working sets residing in cache. A general rule of thumb is to size cache as 10% of your consumed virtual machine storage (not including replica objects). While this is adequate for most workloads, understanding your workload's working set before sizing is a useful exercise. Consider using VMware Infrastructure Planner (VIP) tool to help with this task – http://vip.vmware.com.

The following process is an example of sizing an appropriate working set for performance testing with HClbench. Consider a four node cluster with one 400GB SSD per node. This gives the cluster a total cache size of 1.6TB. The total cache available in Virtual SAN is split 70% for read cache and 30% for write cache. This gives the cluster in our example 1120GB of available read cache and 480GB of available write cache. In order to correctly fit the HClbench within the available the total capacity of all VMDKs should not exceed 1,120GB.

Designing a test scenario with 4 VMs per host, each VM having 5 X 10GB VMDKs, resulting in a total size of 800GB. This will allow the test working set to fit within cache. The default setting for both the number and size of data disks is 10. This value should. If the total of the **Size of Data Disk** parameter should exceed the total cache size of the cluster. The total size of data disk is the **Number of VM** * (times) **Number of Data Disk**.

Sequential workloads versus random workloads

Before doing performance tests it is important to understand the performance characteristics of the production workload to be tested. Different applications have different performance characteristics. Understanding these characteristics is crucial to successful performance testing. When it is not possible to test with the actual application or application specific testing tool it is important to design a test which matches the production workload as closely as possible. Different workload types will perform differently on Virtual SAN.

Sustained sequential write workloads (such as VM cloning operations) run on Virtual SAN will simply fill the cache and future writes will need to wait for the cache to be destaged to the spinning magnetic disk layer before more I/Os can be written to cache, so performance will be a reflection of the spinning disk(s) and not of flash. The same is true for sustained sequential read workflows. If the block is not in cache, it will have to be fetched from spinning disk. Mixed workloads will benefit more from Virtual SAN's caching design.

HCIbench allows you to change the percentage read and the percentage random parameters. As a starting point it is recommended to set the percentage read parameter to 70 and the percentage random parameter to 30%.

Initializing Storage

During configuration of the workload the recommendation is to select the option to initialize storage. This option will zero the disks for each VM being used in the test, helping to alleviate a first write penalty during the performance testing phase.

Test Run Considerations

As frequently read blocks end up in cache, read performance will improve. In a production environment active blocks will already be in cache. When running any kind of performance testing it is important to keep this in mind. As a best practice performance tests should include at least a 15 minute warm up period. Also keep in mind that the longer testing runs the more accurate the results will be. In addition to the cache warming period HCIbench tests should be configured to for at least an hour.

Results

After the Vdbench testing is completed, the test results are collected from all Vdbench instances in the test VMs. And you can view the results at http://Controller_VM_IP/results in a web browser. You can find all of the original result files produced by Vdbench instances inside the subdirectory corresponding to a test run. In addition to the text files, there is another subdirectory named iotest-vdbench-<*VM#*>vm inside, which is the statistics directory generated by Virtual SAN Observer. Virtual SAN performance data can be viewed by opening the stats.html file within the test directory.

12. Testing Hardware Failures

12.1 Understanding Expected Behavior

When doing failure testing with Virtual SAN, it is important to understand the expected behavior for different failure scenarios. You should compare the results of your test to what is expected. The previous section should be read to understand expected failure behaviors.

12.2 Important: Test one Thing at a Time

By default, virtual machines are deployed on Virtual SAN with the ability to tolerate one failure. If you do not wait for the first failure to be resolved, and then try to test another failure, you will have introduced two failures to the cluster. Virtual Machines will not be able to tolerate the second failure and will become inaccessible.

12.3 VM Behavior when Multiple Failures Encountered

Previously we discussed VM operational states and availability. To recap, a VM remains accessible when a full mirror copy of the objects are available, as well as greater than 50% of the components that make up the VM; the witnesses are there to assist with the latter requirement.

Let's talk a little about VM behavior when there are more failures in the cluster than the *NumberOfFailuresToTolerate* setting in the policy associated with the VM.

12.3.1 VM Powered on and VM Home Namespace Object Goes Inaccessible

If a running VM has its VM Home Namespace object go inaccessible due to failures in the cluster, a number of different things may happen. Once the VM is powered off, it will be marked "inaccessible" in the vSphere web client UI. There can also be other side effects, such as the VM getting renamed in the UI to its ".vmx" path rather than VM name, or the VM being marked "orphaned".

12.3.2 VM Powered on and Disk Object Goes Inaccessible

If a running VM has one of its disk objects go inaccessible, the VM will keep running, but its VMDK's I/O is stalled. Typically, the Guest OS will eventually time out I/O. Some operating systems may crash when this occurs. Other operating systems, for example some Linux distributions, may downgrade the filesystems on the impacted VMDK to read-only. The Guest OS behavior, and even the VM behavior is not Virtual SAN specific. It can also be seen on VMs running on traditional storage when the ESXi host suffers an *APD* (All Paths Down).

Once the VM becomes accessible again, the status should resolve, and things go back to normal. Of course, data remains intact during these scenarios.

12.4 What Happens when a Server Fails or is Rebooted?

A host failure can occur in a number of ways. It could be a crash, or it could be a network issue (which is discussed in more detail in the next section). However, it could also be something as simple as a reboot, and that the host will be back online when the reboot process completes. Once again, Virtual SAN needs to be able to handle all of these events.

If there are active components of an object residing on the host that is detected to be failed (due to any of the stated reasons) then those components are marked as ABSENT. I/O flow to the object is restored within 5-7 seconds by removing the ABSENT component from the active set of components in the object.

The ABSENT state is chosen rather than the DEGRADED state because in many cases a host failure is a temporary condition. A host might be configured to auto-reboot after a crash, or the host's power cable was inadvertently removed, but plugged back in immediately. Virtual SAN is designed to allow enough time for a host to reboot before starting rebuilds on other hosts so as not to waste resources. Because Virtual SAN cannot tell if this is a host failure, a network disconnect or a host reboot, the 60minute timer is once again started. If the timer expires, and the host has not rejoined the cluster, a rebuild of components on the remaining hosts in the cluster commences.

If a host fails, or is rebooted, this event will trigger a "Host connection and power state" alarm, and if vSphere HA is enabled on the cluster, it will also cause a" vSphere HA host status" alarm and a "Host cannot communicate with all other nodes in the Virtual SAN Enabled Cluster" message.

If *NumberOfFailuresToTolerate=1* or higher in the VM Storage Policy, and an ESXi host goes down, VMs not running on the failed host continue to run as normal. If any VMs with that policy were running on the failed host, they will get restarted on one of the other ESXi hosts in the cluster by vSphere HA, as long as it is configured on the cluster.

Caution: If VMs are configured in such a way as to not tolerate failures, (*NumberOfFailuresToTolerate=0*), a VM that has components on the failing host will become inaccessible through the vSphere web client UI.

12.5 Simulate Host Failure without vSphere HA

Without vSphere HA, any virtual machines running on the host that fails will not be automatically started elsewhere in the cluster, even though the storage backing the virtual machine in question is unaffected.

Let's take an example where a	VM is running on a host	(cs-ie-h02.ie.local)
-------------------------------	-------------------------	----------------------

Solutions → Actions →	
Getting Started Summary Monitor M	lanage Related Objects
Next hot Ins. Int. [200] Spright (2) 200-2011 Owner, Sin.	6-poc-test-vm2
Guest	t OS: Microsoft Window's Server 2008 R2 (64-bit)
Peretty Sprine and Second	atibility: ESXi 6.0 and later (∨M version 11)
	are Tools: Not running, not installed
DNS 1	Name:
P Powered Op	dresses:
Host:	cs-ie-h02.ie.local
Launch Remote Console 🛛 🛛 🍠	
Download Remote Console 🚺 😽	
▲ VMware Tools is not installed on this	virtual machine.

Figure 12.1: host failure without vSphere HA

It would also be a good test if this VM also had components located on the local storage of this host. However it does not matter if it does not as the test will still highlight the benefits of vSphere HA.

Next, the host is rebooted:



Figure 12.2: Reboot the host

As expected, the host is not responding in vCenter, and the VM becomes disconnected. The VM will remain in a disconnected state until the ESXi host has fully rebooted, as there is no vSphere HA enabled on the cluster, so the VM cannot be restarted on another host in the cluster.



Figure 12.3: ESXi host not responding, VM disconnected

If you now examine the policies of the VM, you will see that it is non-compliant. You can also see the reason why in the lower part of the screen. This VM should be able to tolerate one failure, but due to the failure currently in the cluster (for example: one ESXi host is currently rebooting), this VM cannot tolerate another failure, thus it is non-compliant with its policy.

What can be deduced from this is that not only was the VM's compute running on the host which was rebooted, but that it also had some components residing on the storage of the host that was rebooted. We can confirm this when the host fully reboots.

Stand-poc-test-vm2 Actions -			E.		
Getting Started Summary Monitor Man	age Related Objects				
Issues Performance Policies Tasks Ev	ents Utilization				
C 隆 👘			Q Filter		
Name	VM Storage Policy	Compliance Status	Last Checked		
🫅 VM home	📅 Virtual SAN Default Storage Policy	😣 Noncompliant	13/04/2015 12:04		
🚈 Hard disk 1	🔓 Virtual SAN Default Storage Policy	🙁 Noncompliant	13/04/2015 12:04		
M			2 items 📑 👻		
Compliance Failures Physical Disk Place	ement				
vsan6-poc-test-vm2 - Hard disk 1: VM Storage Policy Compliance Failures					
Property Name	Expected Value	Current \	/alue		
VSAN - Number of failures to tolerate	1	0			

Figure 12.4: VM is non-compliant

Once the ESXi host has rebooted, we see that the VM is no longer disconnected. However it is left in a powered off state.



Figure 12.5: ESXi host rebooted, VM powered off

And as mentioned previously, if the physical disk placement is examined, we can clearly see that the storage on the host that was rebooted, cs-ie-h02.ie.local, was used to store components belonging to the VM.

	1						
b vsan6-poc-test-vm2	Actions 👻						=
Getting Started Summ	ary Monitor Man	age Related Object	s				
Issues Performance	Policies Tasks Ev	ents I Itilization					
133063 Tenomance		ounzadon					
C 🍖 👘						Q Filter	•
Name		VM Storage Policy		Compliance	Status	Last Checked	
🗀 VM home		🔓 Virtual SAN Defa	ult Storage Policy	🗸 Compli	ant	13/04/2015 12:14	
🏝 Hard disk 1		📅 Virtual SAN Defa	ult Storage Policy	🗸 Compli	ant	13/04/2015 12:14	
M						2 item:	: 📑 -
Osmalianas Failuras							
Compliance Failures	Physical Disk Plac	cement					
vsan6-poc-test-vm2 -	VM home : Physical	Disk Placement					
≡` -(Q Filter	•
Туре	Component State	Host	Flash Disk Name		Flash Disk Uuid	HDD Disk Name	HDD
Witness	Active	cs-ie-h04.ie.l	🚐 HP Serial Attache	d SCSI Dis	52742cfb-d99d-cdc1-8ac4-1527	. 🔄 HP Serial Attached SCSI Dis	524
→ RAID 1							
Component	Active	cs-ie-h02.ie.l	🚐 HP Serial Attache	d SCSI Dis	521963f0-33f5-eaaf-d2e1-f7a21	🚐 HP Serial Attached SCSI Dis	52f
Component	Active	cs-ie-h01.ie.l	🛄 HP Serial Attache	d SCSI Dis	528ba019-e369-151e-01b3-26	🚐 HP Serial Attached SCSI Dis	525

Figure 12.6: Components on host that was rebooted

12.6 Simulate Host Failure with vSphere HA

Let's now repeat the same scenario, but with vSphere HA enabled on the cluster. First, power on the VM from the last test.

Next, select the cluster object, and navigate to the Manage tab, then Settings > Services > vSphere HA. vSphere HA is turned off currently.



Figure 12.7: vSphere HA is turned off

Click on the "Edit" button to enable vSphere HA. When the wizard pops up, click on the "Turn on vSphere HA" checkbox as shown below, then click OK.

🚯 VSAN6-Cluster - Edit Cluster Setti	ngs		(?))
vSphere DRS	Turn on vSphere HA		
vSphere HA	Host Monitoring		
	ESX/ESXi hosts in this cluster exchang might cause isolation responses.	e network heartbeats. Disable this feature when performing network maintenance that	
	✓ Host Monitoring		
	Host Hardware Monitoring - VM Comp	onent Protection	
	ESX/ESXi hosts have the capability to d could deem them unusable (for examp	etect various failures that do not necessarily cause virtual machines to go down, but le, losing network/disk communication)	
	Protect against Storage Connectivit	Loss	
	Virtual Machine Monitoring		
	VM Monitoring restarts individual VMs if restarts individual VMs if their in-quest	their VMware Tools heartbeats are not received within a set time. Application Monitoring application heartbeats are not received within a set time.	
	Disabled v		
	► Failure conditions and VM response	Expand for details	
	 Admission Control 	Expand for details	
	 Datastore for Heartbeating 	Expand for details	
	 Advanced Options 	Expand for advanced options	
		OK Can	cel

Figure 12.8: Turn on vSphere HA

This will launch a number of tasks on each ode in the cluster. These can be monitored via the Monitor > Tasks view. When the configuring of vSphere HA tasks complete, select the cluster object, then the Summary tab, then the vSphere HA window and ensure it is configured and monitoring. The cluster should now have Virtual SAN, DRS and vSphere HA enabled.

VSAN6-Cluster Actions -								
Getting Started Summary Monitor Manage Related Objects								
VSAN6-Cluster Total Processors: Total vMotion Migrations	48 : 16							
▼ vSphere HA		▼ vSphere DRS						
CPU Memory 0% 50%	100%	Migration automation level:	ed	Fully Automated Apply priority 1				
CPU reserved for failover: 25 % Memory reserved for failover: 25 % Host Monitoring: Enabled		migration theshold.		priority 2, and priority 3 recommendations				
Virtual SAN		Power management automatio DRS recommendations:	n level:	Off 0				
Add disks to storage	Manual	DRS faults:		0				
Hosts	4 hosts	Cluster Resources						
Flash disks in use	4 of 4 eligible	Hosts	4 Hos	ts				
Data disks in use	8 of 25 eligible	Total virtual flash resources	0.00 B	3				
Total capacity of Virtual SAN datastore	1.06 TB	EVC mode Disabled		led				
Free capacity of Virtual SAN datastore	1.03 TB	L Tago						
Network status	✓ Normal							
C	isk management	Kelated Objects Cluster Consumers						

Figure 12.9: Virtual SAN, DRS and vSphere HA enabled

Verify that the test VM is still residing on host cs-ie-h02.ie.local. Now repeat the same test as before by rebooting host cs-ie-h02.ie.local and examine the differences with vSphere HA enabled.



Figure 12.10: Reboot the host, this time with vSphere HA enabled

On this occasion, a number of HA related events should be displayed on the Summary tab of the host being rebooted (you may need to refresh the web client to see these):



Figure 12.11: vSphere HA messages

However, rather than the VM becoming disconnected for the duration of the host reboot like was seen in the last test, the VM in instead restarted on another host, in this case cs-ie-h03.ie.local.

san6-poc-test-vm2 Ac	tions 🔻	
Getting Started Summary	Monitor Manage	Related Objects
Actions April Topy (this 2000) 	vsan6-poc-tes	t-vm2
Dependent für 1997-2000 - Benar Derpenstram 11.007 -002 AMB- 00 VA HE HE HE KEITE - 12064003-0442-1123-2023-10122200003 1007	Guest OS:	Microsoft Windows Server 2008 R2 (64-bit)
	Compatibility:	ESXi 6.0 and later (VM version 11)
	∨Mw are Tools:	Not running, not installed
	DNS Name:	
Powered On	IP Addresses: Host:	cs-ie-h03.ie.local
Launch Remote Console	🔊 🔊	
Download Remote Console	0 💙 🛸	
⚠ VMware Tools is not ins	alled on this virtual ma	achine.

Figure 12.12: VM restarted on a different host

If you remember earlier we stated that there were some components belonging to the objects of this VM residing on the local storage of the host that was rebooted. These components now show up as "Absent" in the VM > Monitor > policies > Physical Disk Placement view as shown below.

vsan6-poc-test-vm2 Actions -									
Getting Started Summary Monitor	Manage Related Obje	cts							
Issues Performance Policies Tasks Events Utilization									
C 隆 🗟				Q, F	-ilter -				
Name	VM Storage Policy		Compliance Status	Last Checked					
C VM home	France Virtual SAN Def	ault Storage Policy	😣 Noncompliant	13/04/2015 12:33					
📇 Hard disk 1	📅 Virtual SAN Def	ault Storage Policy	😣 Noncompliant	13/04/2015 12:33					
A6					2 items 🔒				
Compliance Failures Physical Dis	sk Placement								
vsan6-poc-test-vm2 - VM home : Ph	ysical Disk Placement								
				Q Fil	ter 🔹				
Туре	Component State	Host	Flash Disk Name	Flash Disk Uuid	HDD Disk Name				
→ RAID 1									
Component	Active	cs-ie-h01.ie.l	HP Serial Attached SCSI Dis	528ba019-e369-151e-01b3-26	🚐 HP Serial Attach				
Component	Absent	Object not found	Object not found	Object not found	🚐 Object not found				
Witness	Active	cs-ie-h04.ie.l	HP Serial Attached SCSI Dis	52742cfb-d99d-cdc1-8ac4-1527	. 🔄 HP Serial Attach				

Figure 12.13: Absent components

However, once the ESXi host completes rebooting, assuming it is within 60 minutes, these components will be rediscovered, resynchronized and placed back in an Active state.

Should the host be disconnected for longer than 60 minutes (the CLOMD timeout delay default value), the "Absent" components will be rebuilt elsewhere in the cluster.

12.7 Disk is Pulled Unexpectedly from ESXi Host

When a magnetic disk is pulled from an ESXi hosts that is using it to contribute storage to Virtual SAN without first decommissioning the disk, all the Virtual SAN components residing on the disk go ABSENT and are inaccessible.

The ABSENT state is chosen over DEGRADED because Virtual SAN knows the disk is not lost, but rather just removed. If the disk gets put back into the server before the 60-minute timeout, no harm is done and Virtual SAN syncs it back up. In this scenario, Virtual SAN is back up with full redundancy without wasting resources on an expensive rebuild.

12.7.1 Expected Behaviors

- If the VM has a policy that includes *NumberOfFailuresToTolerate=1* or greater, the VM's objects will still be accessible from another ESXi host in the Virtual SAN Cluster.
- The disk state is marked as ABSENT and can be verified via vSphere web client UI.
- At this point, all in-flight I/O is halted while Virtual SAN reevaluates the availability of the object (e.g. VM Home Namespace or VMDK) without the failed component as part of the active set of components.
- If Virtual SAN concludes that the object is still available (based on a full mirror copy and greater than 50% of the components being available), all in-flight I/O is restarted.
- The typical time from physical removal of the disk, Virtual SAN processing this event, marking the component ABSENT halting and restoring I/O flow is approximately 5-7 seconds.
- If the same disk is placed back on the same host within 60 minutes, no new components will be re-built.
- If 60 minutes passes, and the original disk has not been reinserted in the host, components on the removed disk will be built elsewhere in the cluster, if capacity is available, including any newly inserted disks claimed by Virtual SAN.
- If the VM Storage Policy has *NumberOfFailuresToTolerate=0*, the VMDK will be inaccessible if one of the VMDK components (think one component of a stripe or a full mirror) resides on the removed disk. To restore the VMDK, the same disk has to be placed back in the ESXi host. There is no other option for recovering the VMDK.

12.8 SSD is Pulled Unexpectedly from ESXi Host

When a solid-state disk drive is pulled without decommissioning it, all the Virtual SAN components residing in that disk group go ABSENT and are inaccessible. In other words, if an SSD is removed, it will appear as a removal of the SSD as well as all associated magnetic disks backing the SSD from a Virtual SAN perspective.

12.8.1 Expected Behaviors

- If the VM has a policy that includes *NumberOfFailuresToTolerate=1* or greater, the VM's objects will still be accessible.
- Disk group and the disks under the disk group states will be marked as ABSENT and can be verified via the vSphere web client UI.
- At this point, all in-flight I/O is halted while Virtual SAN reevaluates the availability of the objects without the failed component(s) as part of the active set of components.
- If Virtual SAN concludes that the object is still available (based on a full mirror copy and greater than 50% of components being available), all in-flight I/O is restarted.
- The typical time from physical removal of the disk, Virtual SAN processing this event, marking the components ABSENT halting and restoring I/O flow is approximately 5-7 seconds.
- When the same SSD is placed back on the same host within 60 minutes, no new objects will be re-built.
- When the timeout expires (default 60 minutes), components on the impacted disk group will be rebuilt elsewhere in the cluster, if capacity is available.
- If the VM Storage Policy has *NumberOfFailuresToTolerate=0*, the VMDK will be inaccessible if one of the VMDK components (think one component of a stripe or a full mirror) exists on disk group whom the pulled SSD belongs to. To restore the VMDK, the same SSD has to be placed back in the ESXi host. There is no option to recover the VMDK.

12.9 What Happens When a Disk Fails?

If a disk drive has an unrecoverable error, Virtual SAN marks the disk as DEGRADED as the failure is permanent.

12.9.1 Expected Behaviors

- If the VM has a policy that includes *NumberOfFailuresToTolerate=1* or greater, the VM's objects will still be accessible.
- The disk state is marked as DEGRADED and can be verified via vSphere web client UI.
- At this point, all in-flight I/O is halted while Virtual SAN reevaluates the availability of the object without the failed component as part of the active set of components.
- If Virtual SAN concludes that the object is still available (based on a full mirror copy and greater than 50% of components being available), all in-flight I/O is restarted.
- The typical time from physical removal of the drive, Virtual SAN processing this event, marking the component DEGRADED halting and restoring I/O flow is approximately 5-7 seconds.
- Virtual SAN now looks for any hosts and disks that can satisfy the object requirements. This includes adequate free disk space and placement rules (e.g. 2 mirrors may not share the same host). If such resources are found, Virtual SAN will create new components on there and start the recovery process immediately.
- If the VM Storage Policy has *NumberOfFailuresToTolerate=0*, the VMDK will be inaccessible if one of the VMDK components (think one component of a stripe) exists on the pulled disk. This will require a restore of the VM from a known good backup.

12.10 What Happens When an SSD Fails?

An SSD failure follows a similar sequence of events to that of a disk failure with one major difference; Virtual SAN will mark the entire disk group as DEGRADED. Virtual SAN marks the SSD and all disks in the disk group as DEGRADED as the failure is permanent (disk is offline, no longer visible, and others).

12.10.1 Expected Behaviors

- If the VM has a policy that includes *NumberOfFailuresToTolerate=1* or greater, the VM's objects will still be accessible from another ESXi host in the Virtual SAN Cluster.
- Disk group and the disks under the disk group states will be marked as DEGRADED and can be verified via the vSphere web client UI.
- At this point, all in-flight I/O is halted while Virtual SAN reevaluates the availability of the objects without the failed component(s) as part of the active set of components.
- If Virtual SAN concludes that the object is still available (based on available full mirror copy and witness), all in-flight I/O is restarted.
- The typical time from physical removal of the drive, Virtual SAN processing this event, marking the component DEGRADED halting and restoring I/O flow is approximately 5-7 seconds.
- Virtual SAN now looks for any hosts and disks that can satisfy the object requirements. This includes adequate free SSD and disk space and placement rules (e.g. 2 mirrors may not share the same hosts). If such resources are found, Virtual SAN will create new components on there and start the recovery process immediately.
- If the VM Storage Policy has *NumberOfFailuresToTolerate=0*, the VMDK will be inaccessible if one of the VMDK components (think one component of a stripe) exists on disk group whom the pulled SSD belongs to. There is no option to recover the VMDK. This may require a restore of the VM from a known good backup.

Warning: Test one thing at a time during the following POC steps. Failure to resolve the previous error before introducing the next error will introduce multiple failures into Virtual SAN which it may not be equipped to deal with, based on the NumberOfFailuresToTolerate setting, which is set to 1 by default.

12.11 Virtual SAN Disk Fault Injection Script for POC Failure Testing

When the Virtual SAN Health Check VIB is installed (installed by default in vSphere 6.0U1), a python script to help with POC disk failure testing is available on all ESXi hosts. The script is called vsanDiskFaultInjection.pyc and can be found on the ESXi hosts in the directory /usr/lib/vmware/vsan/bin. To display the usage, run the following command:

Warning: This command should only be used in pre-production environments during a POC. It should not be used in production environments. Using this command to mark disks as failed can have a catastrophic effect on a Virtual SAN Cluster.

Readers should also note that this tool provides the ability to do "hot unplug" of drives, which is similar to the testing that was done with the *hpssacli* command previously. This is an alternative way of creating a similar type of condition. However, in this POC guide, this script is only being used to inject permanent errors.

12.12 Pull Magnetic Disk/Capacity Tier SSD and Replace before Timeout Expires

In this first example, we shall remove a disk from the host using the *vsanDiskFaultInjection.pyc* python script rather than physically removing it from the host.

It should be noted that the same tests can be run by simply removing the disk from the host. If physical access to the host is convenient, literally pulling a disk would test exact physical conditions as opposed to emulating it within software.

Also note that not all I/O controllers support hot unplugging drives. Check the Virtual SAN Compatibility Guide to see if your controller model supports the hot unplug feature.

We will then examine the effect this operation has on Virtual SAN, and virtual machines running on Virtual SAN. We shall then replace the component before the CLOMD timeout delay expires (default 60 minutes), which will mean that no rebuilding activity will occur during this test.

Pick a host with a running VM.

vsan6-poc-tes	vsan6-poc-test-vm2									
Guest OS:	Microsoft Window's Server 2008 R2 (64-bit)									
Compatibility:	ESXi 6.0 and later (VM version 11)									
VMw are Tools:	Not running, not installed									
DNS Name:										
IP Addresses:										
Host:	cs-ie-h02.ie.local									
灯 🏹										

Figure 12.14: Select host with running VM

Next, navigate to the VM's Monitor tab > Policies, select a Hard Disk and then select Physical Disk Placement tab in the lower half of the screen. Identify a Component object. The column that we are most interested in is HDD Disk Name, as it contains the NAA SCSI identifier of the disk. The objective is to remove one of these disks from the host (other columns may be hidden by right clicking on them).

vsan6-poc-test-vm2	Actions 👻						· ·
etting Started Summa	ary Monitor Mar	age Related Object	S				
sues Performance	Policies Tasks E	vents Utilization					
୯ 🐁 🗟						Q Filter	
lame		VM Storage Policy		Compliance Status	Last Checked		
🔁 VM home		📅 Virtual SAN Defa	ult Storage Policy	🗸 Compliant	13/04/2015 13:0	6	
🔤 Hard disk 1		📑 Virtual SAN Defai	ult Storage Policy	 Compliant 	13/04/2015 13:0	6	
HA.						:	2 items 🔒
Compliance Failures	Physical Disk Pla	cement					
vsan6-poc-test-vm2 - I	" Hard disk 1 : Physic	al Disk Placement					
)- '=						\ Filter	-
Туре	Component State	Host	HDD Disk Name				
→ RAID 1							
Component	Active	cs-ie-h04.ie.l	🚐 HP Serial Attac	hed SCSI Disk (naa.600508b10)	D1c846c000c3d9114ed71b3)		
Component	Active	cs-ie-h01.ie.l	🚐 HP Serial Attac	hed SCSI Disk (naa.600508b10)	01c388c92e817e43fcd5237)		
		-					

Figure 12.15: Display disk identifiers

From figure 12.15, let us say that we wish to remove the disk containing the component residing on host cs-ie-h01.ie.local. That component resides on physical disk with an NAA ID string of naa.600508b1001c388c92e817e43fcd5237. Make a note of your NAA ID string. Next, SSH into the host with the disk to pull. Inject a hot unplug event using the *vsanDiskFaultInjection.pyc* python script:

```
[root@cs-ie-h01:~] python /usr/lib/vmware/vsan/bin/vsanDiskFaultInjection.pyc -u
-d naa.600508b1001c388c92e817e43fcd5237
Injecting hot unplug on device vmhba1:C0:T4:L0
vsish -e set /reliability/vmkstress/ScsiPathInjectError 0x1
vsish -e set /storage/scsifw/paths/vmhba1:C0:T4:L0/injectError 0x004C040000002
```

Let's now check out the VM's objects and components and as expected, the component that resided on that disk in host cs-ie-h01 quickly shows up as absent.

By vsan6-poc-test-vm2 Actions -									
Getting Started Summary Monitor Manage Related Objects									
Issues Performance Policies Tasks Events Utilization									
C 🍖 👘			Q Fi	ter 👻					
Name	VM Storage Policy	Compliance Status	Last Checked						
🥅 VM home	📅 Virtual SAN Default Storage Policy	🗸 Compliant	13/04/2015 13:36						
📇 Hard disk 1	F Virtual SAN Default Storage Policy	🙁 Noncompliant	13/04/2015 13:36						
86				2 items 📑 🗸					
Compliance Failures Physical Disk Place	cement								
vsan6-poc-test-vm2 - Hard disk 1 : Physical	al Disk Placement								
)- `E			Q Filte	r •					
Type Compor	nent State Host	Flash Disk Name	Flash Disk Uuid	HDD Disk Name					
→ RAID 1									
Component 🗧 Ac	ctive cs-ie-h04.ie.l	HP Serial Attached SCSI Dis	52742cfb-d99d-cdc1-8ac4-1527	🛄 HP Serial Attach					
Component 🔲 Ab	osent Object not found	Object not found	Object not found	🚐 Absent VSAN Di:					
Witness Ac	ctive cs-ie-h02.ie.l	HP Serial Attached SCSI Dis	521963f0-33f5-eaaf-d2e1-f7a21	HP Serial Attach					

Figure 12.16: Disk Removed, Component Absent

To put the disk drive back in the host, one simply rescans the host for new disks. Navigate to the host > Manage > Storage > Storage Devices and click the rescan button.

Getting Started	Summary Mo	nitor Manage Related Obje	ects					
Settings Storag	ge Networking	Alarm Definitions Tags Per	missions					
44		Storage Devices						
Storage Ada	pters	🛃 🚊 🛛 🖬 📖 🌾) 💿 🖬	🙆 All Actic	ons 🔁 🛨	(Q Filter	
Storage Dev	rices	Name 🗄	Туре	Capacity	Operational S	Hardware Acceler	Drive Type	Transport
Host Cache	Configuration	SEAGATE Rescans all stor	rage adapters	on the ho	st to	Unknown	HDD	Block Ada
Protocol End	dpoints	SEAGATE VMFS volumes.	auueu storagi	e devices a	and/or	Unknown	HDD	Block Ada
		Local USB Direct-Access (mp	ox disk	0.00 B	Attached	Not supported	HDD	Block Ada
		Local ATA Disk (naa.55cd2e4	40 disk	186.3	Attached	Not supported	Flash	Block Ada
		Local USB CD-ROM (mpx.vm	nhb cdrom		Attached	Not supported	HDD	Block Ada
		SEAGATE Serial Attached So	CSI disk	558.9	Attached	Unknown	HDD	Block Ada
		Device Details			=			
		General						
		Name	SEAGATE Se	rial Attached	d SCSI Disk (n	aa.5000c50071a9	11c3)	
		Identifier	naa.5000c500	71a911c3				
		Туре	disk					
		Location	/vmfs/devices/	disks/naa.5	000c50071a9 ²	11c3		,

Figure 12.17: Rescan storage adapters

Look at the list of storage devices for the NAA ID that was removed. If for some reason, the disk doesn't return after refreshing the screen, try rescanning the host

again. If it still doesn't appear, reboot the ESXi host. Once the NAA ID is back, clear any hot unplug flags set previously with the -c option:

```
[root@cs-ie-h01:~] python /usr/lib/vmware/vsan/bin/vsanDiskFaultInjection.pyc -c
-d naa.600508b1001c388c92e817e43fcd5237
Clearing errors on device vmhba1:C0:T4:L0
vsish -e set /storage/scsifw/paths/vmhba1:C0:T4:L0/injectError 0x00000
vsish -e set /reliability/vmkstress/ScsiPathInjectError 0x00000
```

12.13 Pull Magnetic Disk/Capacity Tier SSD and Do not Replace before Timeout Expires

In this example, we shall remove the magnetic disk from the host, once again using the *vsanDiskFaultInjection.pyc* script. However, this time we shall wait longer than 60 minutes before scanning the HBA for new disks. After 60 minutes, Virtual SAN will rebuild the components on the missing disk elsewhere in cluster.

The same process as before can now be repeated. However this time we shall leave the disk drive removed for more than 60 minutes and see the rebuild activity take place. Begin by identifying the disk on which the component resides.

b vsan6-poc-test-vm2	Actions 👻						=*
Getting Started Summ	ary Monitor	Manage Related Objects					
Issues Performance	Policies Tasks	Events Utilization					
C 🌯 🖏						Q Filter	•
Name		VM Storage Policy		Compliance Status	Last Checked		
🛅 VM home		📅 Virtual SAN Default Stora	ige Policy	🗸 Compliant	13/04/2015 13	:42	
And disk 1		😽 Virtual SAN Default Stora	ige Policy	🗸 Compliant	13/04/2015 13	:42	
86						2 iten	ns 📑 🗝
Compliance Failures	Physical Disk F	Placement					
vsan6-poc-test-vm2 -	Hard disk 1 : Phy	sical Disk Placement					
≡ - ((Q Filter	•
Туре	Component St	Host	HDD Disk Nam	e			
Witness	Active	cs-ie-h02.ie.local	🚐 HP Seria	al Attached SCSI Disk (naa.600	508b1001c0cc0ba2a3866cf8e28	be)	
Component	Active	cs-ie-h01.ie.local	🚐 HP Seria	al Attached SCSI Disk (naa.600	508b1001cd292e69df3a3291db0)c1)	
Component	Active	cs-ie-h04.ie.local	HP Seria	al Attached SCSI Disk (naa.600	508b1001c846c000c3d9114ed7	1b3)	

Figure 12.18: Identify NAA id

```
[root@cs-ie-h01:~] date
Mon Dec 14 13:36:02 UTC 2015
[root@cs-ie-h01:~] python /usr/lib/vmware/vsan/bin/vsanDiskFaultInjection.pyc
-u -d naa.600508b1001c388c92e817e43fcd5237
Injecting hot unplug on device vmhba1:C0:T4:L0
vsish -e set /reliability/vmkstress/ScsiPathInjectError 0x1
vsish -e set /storage/scsifw/paths/vmhba1:C0:T4:L0/injectError 0x004C040000002
```

At this point, we can once again see that the component has gone *absent*. After 60 minutes have elapsed, the component should now be rebuilt.

vsan6-poc-test-vm2	Actions -					
tting Started Summ	ary Monitor	Manage Related Objects				
ues Performance	Policies Tasks	Events Utilization				
2 🍋 👘					Q Filter	
ame		VM Storage Policy		Compliance Status	Last Checked	
VM home		📑 Virtual SAN Default	t Storage Policy	🗸 Compliant	13/04/2015 14:37	
Hard disk 1		F Virtual SAN Default	t Storage Policy	8 Noncompliant	13/04/2015 14:37	
A						2 items 🔒
Compliance Failures	Physical Disk I	Placement		-		
san6-poc-test-vm2 -	Hard disk 1 : Phy	sical Disk Placement				
<u>≡ (</u>					Q Filter	•
Туре	Component St	Host	HDD Disk Nar	ne		
→ RAID 1						
Component	Active	cs-ie-h04.ie.local	🚐 HP Ser	ial Attached SCSI Disk (naa.60050)	8b1001c846c000c3d9114ed71b3)	
Component	Absent	Object not found	🚐 Absent	VSAN Disk (VSAN UUID:52fc637f-e	ecf9-2b53-ff31-9e8d75d2b43f)	
Witness	Active	cs-ie-h02.ie.local	A HP Ser	ial Attached SCSI Disk (naa.60050)	8b1001c0cc0ba2a3866cf8e28be)	

Figure 12.19: Component is absent

After the 60 minutes has elapsed, the component should be rebuilt on a different disk in the cluster. That is what is observed. Note the component resides on a new disk (NAA id is different).

stan6-poc-test-vm2 Actions 🕶					=*					
Getting Started Summary Monitor Manage Related Objects										
Issues Performance Policies Tasks Events Utilization										
C 🍄 🖏				Q Filter	•					
Name	VM Storage Policy		Compliance Status	Last Checked						
🛅 VM home	Virtual SAN Default Stora	age Policy	✓ Compliant	13/04/2015 14:51						
📇 Hard disk 1	Virtual SAN Default Stora	age Policy	✓ Compliant	13/04/2015 14:51						
M					2 items 🛛 🔒 🕶					
Compliance Failures Physical D	Disk Placement		-							
vsan6-poc-test-vm2 - Hard disk 1	: Physical Disk Placement									
= -(Q Filter	-					
Type Component	St Host	HDD Disk Name								
→ RAID 1										
Component 📕 Active	cs-ie-h04.ie.local	🚐 HP Serial /	Attached SCSI Disk (naa.600508b1001c846	c000c3d9114ed71b3)						
Component Active	cs-ie-h02.ie.local	🚐 HP Serial /	Attached SCSI Disk (naa.600508b1001cb22	34d6ff4f7b1144f59)						
Witness Active	cs-ie-h03.ie.local	🚐 HP Serial /	Attached SCSI Disk (naa.600508b1001cb11	(3292fe743a0fd2e7)						

Figure 12.20: Component is rebuilt

The removed disk can now be re-added by scanning the HBA:

Navigate to the host > Manage > Storage > Storage Devices and click the rescan button. See Figure 12.18 above for a screenshot.

Look at the list of storage devices for the NAA ID that was removed. If for some reason, the disk doesn't return after refreshing the screen, try rescanning the host again. If it still doesn't appear, reboot the ESXi host. Once the NAA ID is back, clear any hot unplug flags set previously with the -c option:

```
[root@cs-ie-h01:~] python /usr/lib/vmware/vsan/bin/vsanDiskFaultInjection.pyc -c
-d naa.600508b1001c388c92e817e43fcd5237
Clearing errors on device vmhbal:C0:T4:L0
vsish -e set /storage/scsifw/paths/vmhbal:C0:T4:L0/injectError 0x00000
vsish -e set /reliability/vmkstress/ScsiPathInjectError 0x00000
```

That completes this part of the POC.

12.14 Pull Cache Tier SSD and Do Not Reinsert/Replace

For the purposes of this test, we shall remove an SSD from one of the disk groups in the cluster. Navigate to the cluster > Manage > Settings > Virtual SAN > Disk Management. Select a disk group from the top window and identify its SSD in the bottom window. If All-Flash, make sure it's the Flash device in the "Cache" Disk Role. Make a note of the SSD's NAA ID string.

C1 Actions -					=-
Getting Started Summary Mor	nitor Manage Related Objects				
Settings Scheduled Tasks Ala	rm Definitions Tags Permissions				
	Disk Groups				
- Services	🥶 🚉 👛 🐂 📜 Č		Q Filter	Filter -	
vSphere DRS	Disk Group	Disks in Use	State		Vir
vSphere HA	w2-stsds-139.eng.vmware.com	3 of 7	Connect	ed	н.:
	Bisk group (02000000055cd2e404b66fcc5494e54454c20)	3	Mounted		H_
General	✓	3 of 7	Connect	ed	н
Disk Management	Disk group (02000000055cd2e404b5b7fa0494e54454c20)	3	Mounted		ΗŢ
Fault Domains					
Health	M			8 items	
- Configuration	Execution ====================================				
General					
Licensing	🚢 🙀 🥥 💿 🛫 🎡 All Actions 👻	Show:	In use (3)		•
VMware EVC	Name	Drive Type	Disk Role	Capacity	
VM/Host Groups	E Local ATA Disk (naa.55cd2e404b66fcc5)	Flash	Cache	186.31	GB
VM/Host Rules	SEAGATE Serial Attached SCSI Disk (naa.5000c50071a912f3)	HDD	Capacity	558.91	GB
VM Overrides	SEAGATE Serial Attached SCSI Disk (naa.5000c50071a9151b)	HDD	Capacity	558.91	GB
Host Options	•				•
Profiles	A4			3 items	

Figure 12.21: Locate a caching-tier SSD

In the above screenshot, we have located an SSD on host w2-stsds-139 with an NAA ID string of naa.55cd2e404b66fcc5. Next, SSH into the host with the SSD to pull. Inject a hot unplug event using the *vsanDiskFaultInjection.pyc* python script:

```
[root@w2-stsds-139:~] python /usr/lib/vmware/vsan/bin/vsanDiskFaultInjection.pyc
-u -d naa.55cd2e404b66fcc5
```

```
Injecting hot unplug on device vmhba1:C0:T4:L0
vsish -e set /reliability/vmkstress/ScsiPathInjectError 0x1
vsish -e set /storage/scsifw/paths/vmhba1:C0:T4:L0/injectError 0x004C040000002
```

Now we observe the impact that losing an SSD (flash device) has on the whole disk group.
C1 Actions -				=					
Getting Started Summary Mo	nitor Manage Related Objects								
Settings Scheduled Tasks Ala	rm Definitions Tags Permissions								
44	Disk Groups								
✓ Services									
vSphere DRS Disks in Use State Virtual SAN H F									
vSphere HA	w2-stsds-139.eng.vmware.com	3 of 7 Co	onnected	Jnhealthy ::					
✓ Virtual SAN	式 Disk group (527f728d-7af8-870f-607b-9a7d3ff23a86)	3 Me	ounted	Jnhealthy					
General		3 of 7 Co	onnected	Healthy					
Disk Management	Bisk group (02000000055cd2e404b5b7fa0494e54454c20)	3 M	ounted I	-lealthy -					
Fault Domains				•					
Health	1996 -			8 items 📑 🍷					
- Configuration	E ====================================								
General									
Licensing	🏭 🙀 🥥 🔍 🚬 🎡 All Actions 👻	Show:	In use (3)	•					
VMware EVC	Name	Drive Type	Disk Role	Capacity					
VM/Host Groups	Absent VSAN Disk (VSAN UUID:527f728d-7af8-870f-607b-9a7d3ff	Flash	Cache	0.00 B					
VM/Host Rules	SEAGATE Serial Attached SCSI Disk (naa.5000c50071a912f3)	HDD	Capacity	558.91 GB					
VM Overrides	SEAGATE Serial Attached SCSI Disk (naa.5000c50071a9151b)	HDD	Capacity	558.91 GB					
Host Options									
Profiles	M			3 items 斗 🕶					

Figure 12.22: Absent cache tier SSD = Unhealthy Disk Group

And finally, let's look at the components belonging to the virtual machine. This time, any components that were residing on that disk group are absent.

san6-poc-test-vm2	Actions +							=	
Getting Started Summa	ry Monitor	Manage Rel	ated Objects						
Issues Performance F	Policies Tasks	Events Utili:	zation						
C 💁 🗞							Q Filter	•	
Name			VM Storage Policy		Compliance Status	Last Checked			
🛅 VM home			📴 Virtual SAN De	fault Storage Policy	😒 Noncompliant	13/04/2015 15:04			
📇 Hard disk 1			😚 Virtual SAN De	fault Storage Policy	🗸 Compliant	13/04/2015 15:04			
86								2 items 🔒 🗸	
					-				
Compliance Failures	Physical Disk	Placement							
vsan6-poc-test-vm2 - V	/M home : Physi	cal Disk Place	ment						
≣*(Q Filter	•	
Туре	Component St	Host		HDD Disk Name					
Witness	Active	cs-ie-h04	.ie.local	🚐 HP Serial Attached SCSI [HP Serial Attached SCSI Disk (naa.600508b1001c258181f0a088f6e40dab)				
→ RAID 1									
Component	Component 📄 Active 🔋 cs-ie-h02.ie.local 🔄 HP Serial Attached SCSI Disk (naa.600508b1001c0cc0ba2a3866df8e28be)								
Component	Absent	Object not fo	und	HP Serial Attached SCSI [Disk (naa.600508b1001cf23cc9b38261ddc75f09)				

Figure 12.23: SSD removed – all components absent

To show that this impacts all VMs, here is another VM that had a component on local storage on host cs-ie-h01.ie.local.

6	vsan6-poc-test-vm-3	Actions -							≡*	
G	etting Started Summ	ary Monitor	Manage Rel:	ated Objects						
ŀ	ssues Performance	Policies Tasks	Events Utilia	zation						
	C 🌚 🕼							Q Filter	•	
	Name			VM Storage Policy		Compliance Status	Last Checked			
	🛅 VM home			StripeWidth=2	2	Noncompliant	13/04/2015 15:04			
	🚑 Hard disk 1			StripeWidth=2	2	O Noncompliant	13/04/2015 15:04			
	86								2 items 🛛 🖨 🕶	
	Compliance Failures	Physical Disk	Placement							
	vsan6-poc-test-vm-3 -	- Hard disk 1 : Ph	ysical Disk Pla	cement						
)— "≣							Q Filter	-	
	Туре	Component St	Host		HDD Disk Name					
	→ RAID 1									
	+ RAID 0									
	Component	Active	cs-ie-h03	.ie.local	IP Serial Attached St	CSI Disk (naa.600508b1001c2b7a3d39534	ac6beb92d)			
	Component	Active	cs-ie-h02	.ie.local	I HP Serial Attached So	CSI Disk (naa.600508b1001c0cc0ba2a3866	(cf8e28be)			
	→ RAID 0		_							
	Component 🗖 Active 🗊 cs-ie-h04 ie local 🖉 HP Serial Attached SCSI Disk (naa 600508b1001c258181f0a088f6e40dab)									
	Component	Absent	Object not fo	und	HP Serial Attached St	CSI Disk (naa.600508b1001cf23cc9b38261	ddc75f09)			
			,				,			

Figure 12.24: SSD removed – all components absent

If you search all your VMs, you will see that each VM that had a component on the disk group on cs-ie-h07 now has absent components. This is expected since an SSD failure impacts the whole of the disk group.

After 60 minutes have elapsed, new components should be rebuilt in place of the absent components. If you manage to refresh at the correct moment, you should be able to observe the additional components synchronizing with the existing data.

C1 Actions -								=*
Getting Started Summary Mor	nitor Manage Relate	d Objects						
Issues Performance Profile Co	ompliance Tasks Ever	nts Resour	ce Reservation	vSphere DRS	Utilization	Virtual SAN]	
44	Resyncing Components	6						
Physical Disks Virtual Disks Resyncing Components	Resyncing components Virtual SAN cluster. Mon earlier than ESXi 6.0	view displays itoring object	the status of vin tresynchronizati	rtual machine o on is not availa	bjects that are ble for cluster	e currently be rs containing	ing resynchronized only hosts with ver	l in the sion
Health	G							
Proactive Tests	Resyncing components	5						
	Bytes left to resync	8.11 GE	3					
	ETA to compliance	0 secon	d					
	, ≡		C Filter					•
	Name		VM Storage Policy	н	lost		Bytes Left to Resync	
	👻 🚰 NFS04						1.62 GB	
	👻 🏝 Hard disk 1		📑 Virtual SAI	N Default			1.62 GB	
	Component				w2-stsds-1	138.eng.v	1.62 GB	
	🔫 🔂 NFS01						1.62 GB	
	👻 🏝 Hard disk 1		📑 Virtual SAI	N Default			1.62 GB	
	Component			1	w2-stsds-1	137.eng.v	1.62 GB	T
	4		::					•
	<i>8</i> 6						15 items	→ •

Figure 12.25: New components resynchronizing after clomd timeout expires

To complete this POC, re-add the SSD logical device back to the host by rescanning the HBA:

Navigate to the host > Manage > Storage > Storage Devices and click the rescan button. See Figure 12.18 above for a screenshot.

Look at the list of storage devices for the NAA ID of the SSD that was removed. If for some reason, the SSD doesn't return after refreshing the screen, try rescanning the host again. If it still doesn't appear, reboot the ESXi host. Once the NAA ID is back, clear any hot unplug flags set previously with the -c option:

```
[root@cs-ie-h01:~] python /usr/lib/vmware/vsan/bin/vsanDiskFaultInjection.pyc -c
-d naa.55cd2e404b66fcc5
```

```
Clearing errors on device vmhbal:C0:T4:L0
vsish -e set /storage/scsifw/paths/vmhbal:C0:T4:L0/injectError 0x00000
vsish -e set /reliability/vmkstress/ScsiPathInjectError 0x00000
```

USAN6-Cluster Actions -								≡*
Getting Started Summary Mo	nitor Manage Related Objects							
Settings Scheduled Tasks Ala	rm Definitions Tags Permissions							
	Disk Groups							
✓ Services							Q. Filter	•
vSphere DRS	Disk Group	Disks in Use	State	Virtual SAN	Fault Domain	Network Parti	Disk Format Version	
vSphere HA		3 of 7	Connected	Healthy		Group 1		
✓ Virtual SAN	Disk group (0200080000600508b1001c9c8b5f6f0d7a2be	3		Healthy			2	
General	✓ □ cs-ie-h01.ie.local	3 of 7	Connected	Healthy		Group 1		
Disk Management	Disk group (0200070000600508b1001cb683ff0e29252f9e	3		Healthy			2	
Fault Domains		3 of 7	Connected	Healthy		Group 1		
Health	Disk group (0200080000600508b1001c64b76c8ceb56e8	3		Healthy			2	
- Configuration		3 of 7	Connected	Healthy		Group 1		
General	Disk group (0200080000600508b1001c29d8145d6cc192	3		Healthy			2	
Licensing								
VMware EVC								
VM/Host Groups								
VM/Host Rules	A6						8 items	-€
VM Overrides		O 42 441 Disks	-					
Host Options	DISK group (02000/0000600508b1001Cb685ff0e29252f9e6dcc4c4f4/4	(94341): DISKS						
Profiles							Show: In use (3)	•
	Name	Drive Type	apacity Virtual S	AN Health Status	Operational	Transport Type		
	HP Serial Attached SCSI Disk (naa.600508b1001cb683ff0e2925	HDD	186.28 GB Health	iy.	Mounted	Block Adapte	r	
	HP Serial Attached SCSI Disk (naa.600508b1001c6dcca2f50488	HDD	136.70 GB Health	iy	Mounted	Block Adapte	r	
	HP Serial Attached SCSI Disk (naa.600508b1001cf23cc9b38261	HDD	136.70 GB Health	iy	Mounted	Block Adapte	r	

Figure 12.26: Verify that the disk group is back in a health state

Warning: If you delete an SSD drive that was marked as an SSD, and a logical RAID 0 device was rebuilt as part of this test, you may have to mark the drive as an SSD once more.

12.15 Checking Rebuild/Resync Status

Virtual SAN 6.0 displays details on resyncing components. Navigate to Monitor tab > Virtual SAN > Resyncing Components.

VSAN6-Cluster Actions 🕶] VSAN6-Cluster Actions →									
Getting Started Summary M	onitor Manage Related O	bjects								
Issues Profile Compliance Pe	erformance Tasks Events	Resource Reservation Virtual SAN	VSphere DRS VSphere	HA Utilization						
	Resyncing Components									
Health	Resyncing components view	w displays the status of virtual mach	ine objects that are currently	being resynchronized in the Virtual S	AN cluster. Monitoring object resynchronization					
Proactive Tests	is not available for clusters containing only hosts with version earlier than ESXI 6.0									
Physical Disks	G									
Virtual Disks	Resyncing components									
Resyncing Components	Bytes left to resync									
	ETA to compliance									
)				Q Filter -					
	Name	VM Storage Policy	Host	Bytes Left to Resync	ETA					
			This list is en	hpty.						

Figure 12.27: Resyncing Components

To check the status of component resync/rebuild on a Virtual SAN Cluster using RVC commands, the following command will display useful information:

vsan.resync_dashboard

When resynchronization is complete, this command will report "0 bytes to sync".

12.16 Injecting a Disk Error

The first step is to select a host, and the select a disk that is part of a disk group on that host. The –d DEVICENAME argument requires the SCSI identifier of the disk, typically the NAA id. You might also wish to verify that this disk does indeed contain VM components. This can be done by selecting a VM, then selecting the Monitor > Policies > Physical Disk Placement tab.



cs-ie-03, and has an NAA id of 600508b1001c1a7f310269ccd51a4e83:

Figure 12.28: Healthy Disk Group

The error can only be injected from the command line of the ESXi host. To display the NAA ids of the disks on the ESXi host, you will need to SSH to the ESXi host, login as the *root* user, and run the following command:

```
[root@cs-ie-h03:/usr/lib/vmware/vsan/bin] esxcli storage core device list| grep
^naa
naa.600508b1001ceefc4213ceb9b51c4be4
naa.600508b1001cd259ab7ef213c87eaad7
naa.600508b1001c9c8b5f6f0d7a2be44433
naa.600508b1001c2b7a3d39534ac6beb92d
naa.600508b1001cb11f3292fe743a0fd2e7
naa.600508b1001cla7f310269ccd51a4e83
naa.600508b1001c1a7f310269ccd51a4e83
naa.600508b1001c9b93053e6dc3ea9bf3ef
naa.600508b1001c626dcb42716218d73319
```

Once a disk has been identified, and has been verified to be part of a disk group, and that the disk contains some virtual machine components, we can go ahead and inject the error as follows:

```
[root@cs-ie-h03:/usr/lib/vmware/vsan/bin] python vsanDiskFaultInjection.pyc -p
-d naa.600508b1001c1a7f310269ccd51a4e83
Injecting permanent error on device vmhbal:C0:T0:L4
vsish -e set /reliability/vmkstress/ScsiPathInjectError 0x1
vsish -e set /storage/scsifw/paths/vmhbal:C0:T0:L4/injectError
0x0311030000002
[root@cs-ie-h03:/usr/lib/vmware/vsan/bin]
```

Before too long, the disk should display an error and the disk group should enter an unhealthy state.

VSAN6-Cluster Actions -						≡*
Getting Started Summary Mor	Nitor Manage Related Objects					
Settings Scheduled Tasks Alar	m Definitions Tags Permissions					
**	Disk Groups					
▼ Services	📾 🛤 🐂 T 🖬 🙋				Q Filte	• 1
vSphere DRS	Disk Group	Disks in Use	State	Virtual SAN	Network Parti	Disk Format Version
vSphere HA	cs-ie-h03.ie.local	4 of 7	Connecte	d Unhealthy	Group 1	
✓ Virtual SAN	B Disk group (0200080000600508b1001c9c8b5f6f0d7a2be	4		Unhealthy		2
General	cs-ie-h01.ie.local	3 of 7	Connecte	d Healthy	Group 1	
Disk Management	Disk group (0200070000600508b1001cb683ff0e29252f9e	3		Healthy		2
Fault Domains		3 of 7	Connecte	d Healthy	Group 1	
Health	Disk group (0200080000600508b1001c62313d3c49ad8e	3		Healthy		2
- Configuration	✓ ☐ cs-ie-h04.ie.local	3 of 7	Connecte	d Healthy	Group 1	
General	Disk group (0200080000600508b1001c29d8145d6cc192	3		Healthy		2
Licensing						
VMware EVC	4 III					e itoma
VM/Host Groups	in the second se	_				o iteriis 📑
VM/Host Rules	Disk group (0200080000600508b1001c9c8b5f6f0d7a2be444334c4f47	494341): Disks				
VM Overrides	🖴 🗔 🥝 💿 🖬 晃				Show: In use	e (4)
Host Options	Name	Drive Type	Capacity	Virtual SAN Health Status	Operations	Transport Type
Profiles	HP Serial Attached SCSI Disk (naa 600508b1001c9c8b5f6f0d7a	Flash	186.28 GB	Healthy	Mounted	Block Adapter
	HP Serial Attached SCSI Disk (naa.600508b1001ceefc4213ceb9	HDD	136.70 GB	Healthy	Mounted	Block Adapter
	HP Serial Attached SCSI Disk (naa.600508b1001c1a7f310269cc	HDD	136.70 GB	Permanent disk loss	Mounted	Block Adapter
	HP Serial Attached SCSI Disk (naa.600508b1001c9b93053e6dc	HDD	136.70 GB	Healthy	Mounted	Block Adapter

Figure 12.29: Unhealthy Disk Group

Notice that the disk group is in an Unhealthy state and the status of the disk is "Permanent disk loss". This should place any components on the disk into a degraded state (which can be observed via the VM's Physical Disk Placement tab, and initiate an immediate rebuild of components. Navigating to Cluster > Monitor > Virtual SAN > Resyncing Components should reveal the components resyncing.

VSAN6-Cluster Actions	-					=*			
Getting Started Summary	Monitor Manage Relate	d Objects							
(1 -			1				
Issues Profile Compliance	Performance Health Tas	ks Events	Resource Reservation	Virtual SAN VSphere DRS	S VSphere HA Utilization				
44	Resyncing Component	5							
Health	Resyncing components	s view displa	avs the status of virtual ma	chine objects that are curre	ntly being resynchronized in th	e Virtual SAN cluster. Monitoring object			
Proactive Tests	resynchronization is no	t available fo	or clusters containing only	hosts with version earlier t	nan ESXi 6.0				
Physical Disks	C								
Virtual Disks	Resyncing componen	ts 4							
Resyncing Components Bytes left to resync 29.77 GB									
	ETA to compliance	11 mi	nutes						
				-					
	·- •≡		Q. Filter						
	Name		VM Storage Policy	Host	Bytes Left to Resync	ETA			
	👻 🚰 🛛 vRealize Operat	ions Man	-		17.79 GB	11 minutes			
	👻 📤 Hard disk 2		😚 Virtual SAN Default		17.79 GB	11 minutes			
	Component		-	cs-ie-h04.ie.local	17.79 GB	11 minutes			
	👻 🎒 w2k12-iometer-	1	-	-	979.00 MB	0 second			
	👻 🚈 Hard disk 3		📅 Virtual SAN Default		979.00 MB	0 second			
	Component		-	cs-ie-h02.ie.local	979.00 MB	0 second			
	👻 🎒 Nexenta IO Eng	ne Image	570	-	3.36 GB	0 second			
	👻 🎰 Hard disk 1		📅 Virtual SAN Default		3.36 GB	0 second			
	Component		-	cs-ie-h04.ie.local	3.36 GB	0 second			
	👻 💭 win2k12-vsan				7.66 GB	5 minutes			
	👻 🚈 Hard disk 1		G Virtual SAN Default		7.66 GB	5 minutes			
	Component		-	cs-ie-h04.ie.local	7.66 GB	5 minutes			

Figure 12.30: Resyncing components after disk failure

12.16.2 Clear a Permanent Error

At this point, we can clear the error. We use the same script that was used to inject the error, but this time we provide a -c (clear) option:

```
[root@cs-ie-h03:/usr/lib/vmware/vsan/bin] python vsanDiskFaultInjection.pyc -c
-d naa.600508b1001c1a7f310269ccd51a4e83
Clearing errors on device vmhba1:C0:T0:L4
vsish -e set /storage/scsifw/paths/vmhba1:C0:T0:L4/injectError 0x00000
vsish -e set /reliability/vmkstress/ScsiPathInjectError 0x00000
[root@cs-ie-h03:/usr/lib/vmware/vsan/bin]
```

Note however that since the disk failed, it will have to be removed, and re-added from the disk group. This is very simple to do. Simply select the disk in the disk group, and remove it by clicking on the icon highlighted below.

Disk group (0200080000600508b1001c9c8b5f6f0d7a2be444334c4f47494341): Disks											
쓰 🗔 🥝 🗉 토	In use (4)										
Name	Drive Type	Capacity	Virtual SAN Health Status	Operational	Transport Type						
F HP Serial Attached SCSI Disk (naa.600508b1001c9c8b5f6f0d7a	Flash	186.28 GB	Healthy	Mounted	Block Adapter						
IP Serial Attached SCSI Disk (naa.600508b1001ceefc4213ceb9	HDD	136.70 GB	Healthy	Mounted	Block Adapter						
HP Serial Attached SCSI Disk (naa.600508b1001c1a7f310269cc	HDD	136.70 GB	Permanent disk fa	Mounted	Block Adapter						
HP Serial Attached SCSI Disk (naa.600508b1001c9b93053e6dc	HDD	136.70 GB	Healthy	Mounted	Block Adapter						
A items 🕒 🗸											

Figure 12.31: Remove disk from disk group

This will display a pop-up window regarding which action to take regarding the components on the disk. You can choose to migrate the components or not. By default it is shown as "Evacuate Data", shown here.



Figure 12.32: Data is evacuated by default, but can be unchecked in this test

For the purposes of this POC, you can <u>uncheck</u> this option as you are adding the disk back in the next step. When the disk has been removed and re-added, the disk group will return to a healthy state. That completes the disk failure test.

12.17 When Might a Rebuild of Components Not Occur?

There are a couple of reasons why a rebuild of components might not occur.

12.17.1 Lack of Resources

Verify that there are enough resources to rebuild components before testing with the following RVC command:

• vsan.whatif_host_failures

Of course, if you are testing with a 3-node cluster, and you introduce a host failure, there will be no rebuilding of objects. Once again, if you have the resources to create a 4-node cluster, then this is a more desirable configuration for evaluation Virtual SAN.

12.17.2 Underlying Failures

Another cause of a rebuild not occurring is due to an underlying failure already present in the cluster. Verify there are none before testing with the following RVC command:

- vsan.hosts_info
- vsan.check_state
- vsan.disks_stats

If these commands reveal underlying issues (ABSENT or DEGRADED components for example), rectify these first or you risk inducing multiple failures in the cluster, resulting in inaccessible virtual machines.

13. Virtual SAN Management

In this section, we shall look at a number of management tasks, such as the behavior when placing a host into maintenance mode, and the evacuation of a disk and a disk group from a host. We will also look at how to turn on and off the identifying LEDs on a disk drive.

13.1 Put a Host into Maintenance Mode

There are a number of options available when placing a host into maintenance mode. The first step is to identify a host that has a running VM, as well as components belonging to virtual machine objects.

Select the Summary tab of the virtual machine to verify which host it is running on.



Figure 13.1: VM Summary tab

Then select the Monitor tab > Policies > Physical Disk Placement and verify that there are components also residing on the same host.

vsan6-poc-test-vm2 Actions -								7
etting Started Summary Monitor Man	age Related Obj	ects						
ssues Performance Policies Tasks Ew	ents Utilization							
୯ କ୍ଷ 🗟							Q Filter	
Name		Compliance Status	L	ast Checked				
🛅 VM home	🔂 Virtual	SAN Default Storage Po	olicy	✓ Compliant		14/04/2015 02:29		
🗠 Hard disk 1	SAN Default Storage Po	blicy	 Compliant 		14/04/2015 02:29			
M							21	tems 🔒 🕇
Compliance Failures Physical Disk Place	ement		=					
vsan6-poc-test-vm2 - Hard disk 1 : Physica	I Disk Placement							
= =						(Q Filter	•
Type Compo	nent State	Host	Flash Disk Name		Flash Disk Uuid	HDD Disk Name		HDD Disk U
→ RAID 1								
Component Ac	tive	cs-ie-h04.ie.l	🚐 HP Serial At	tached SCSI Dis	52742cfb-d99d-cdc1-8ac4-1527	- HP Serial Atta	ched SCSI Dis	52c448d
Component 🗧 Ac	tive	cs-ie-h02.ie.l) 🚐 HP Serial At	tached SCSI Dis	521963f0-33f5-eaaf-d2e1-f7a21	IP Serial Atta	iched SCSI Dis	52edaed
Witness	tivo	El cs-ie-b03 ie l	I UR Social At	lacked SCSI Die	52949c9h_f622_6025_hee3_746	HP Serial Atta	shad SCSI Dia	507aada

Figure 13.2: Physical Disk Placement

From the screenshots shown here, we can see that the VM selected is running on host cs-ie-h02 and also has components residing on that host. This is the host that we shall place into maintenance mode.

Right click on the host, select Maintenance Mode from the dropdown menu, then select the option "Enter Maintenance Mode" as shown here.



Figure 13.3: Enter Maintenance Mode

There are three options displays when maintenance mode is selected; (i) Ensure accessibility, (ii) Full data migration and (iii) No data migration.



Figure 13.4: Maintenance Mode options

In this first part of the maintenance mode testing, we shall select the option "Ensure accessibility". This means that although components may go missing, the VMs shall remain accessible.

When this option is chosen, a popup is displayed regarding migrating running virtual machines. Since this is a fully automated DRS cluster, the virtual machines should be automatically migrated.



Figure 13.5: Migration warning

After the host has entered maintenance mode, we can now examine the state of the components that were on the local storage of this host. What you should observe is that these components are now in an "Absent" state. However the VM remains accessible as we chose the option "Ensure Accessibility" when entering Maintenance Mode.

vsan6-poc-test-vm2 Acti	ons 🔻					-
Setting Started Summary	Monitor Manage Rela	ated Objects				
ssues Performance Policie	s Tasks Events Utiliz	zation				
୯ 💊 🗟					Q Filter	-
Name	VM	Storage Policy	Compliance Status	L	ast Checked	
🛅 VM home	5	Virtual SAN Default Storage Po	licy 🗸 Compliant		14/04/2015 08:41	
🚈 Hard disk 1	Hard disk 1 🔂 Virtual SAN Default Storage Policy 🗸 Compliant 14/04/2015 08:41					
MA						2 items 🔒 🕇
Compliance Failures Phys	sical Disk Placement					
vsan6-poc-test-vm2 - Hard d	lisk 1 : Physical Disk Plac	ement				
)— <u> </u>					Q Filter	•
Туре	Component State	Host	Flash Disk Name	Flash Disk Uuid	HDD Disk Name	HDD Disk U
Witness	Active	cs-ie-h03.ie.l	IP Serial Attached SCSI Dis	52a4acab-f622-6025-bee3-746	🚐 HP Serial Attached SCSI Dis	527aade
→ RAID 1						
Component	Absent	cs-ie-h02.ie.l	IP Serial Attached SCSI Dis	521963f0-33f5-eaaf-d2e1-f7a21	🚐 HP Serial Attached SCSI Dis	52edaed
		E		E0740alb 4004 adot 0ao4 4507		

Figure 13.6: Components are Absent during Maintenance Mode

The host can now be taken out of maintenance mode. Simply right click on the host as before, select Maintenance Mode and then Exit Maintenance Mode.



Figure 13.7: Exit Maintenance Mode

After exiting Maintenance Mode, the "Absent" component becomes Active once more. This is assuming that the host exited maintenance mode before the vsan.ClomdRepairDelay expires (default 60 minutes).

Actions -	Gy vsan6-poc-test-vm2 Actions → ===										
Getting Started Summary Monitor Mana	age Related Objects										
Issues Performance Policies Tasks Events Utilization											
C S a											
Name	VM Storage Policy		Compliance Status	1	Last Checked						
VM home	📑 Virtual SAN Default Storag	ge Policy	🗸 Compliant		14/04/2015 08:41						
🖻 Hard disk 1 📑 Virtual SAN Default Storage Policy 🗸 Compliant 14/04/2015 08:41											
At 2 items 🔒 🗸											
Compliance Failures Physical Disk Place	ement										
vsan6-poc-test-vm2 - Hard disk 1 : Physical	I Disk Placement										
)— —					Q Filter	•					
Type Compon	nent State Host	Flash Disk Name		Flash Disk Uuid	HDD Disk Name	HDD Disk U					
✓ RAID 1											
Component Act	tive Cs-ie-h04.ie	I 🚐 HP Serial A	ttached SCSI Dis	52742cfb-d99d-cdc1-8ac4-1527.	🚐 HP Serial Attached SC	SI Dis 52c448d					
Component 🗧 Act	tive 📋 cs-ie-h02.ie	.I 🛄 HP Serial A	ttached SCSI Dis	521963f0-33f5-eaaf-d2e1-f7a21.	🚐 HP Serial Attached SC	SI Dis 52edaed					
Witness Act	tive Cs-ie-h03.ie	I 🔲 HP Serial A	ttached SCSI Dis	52a4acab-f622-6025-bee3-746	. IP Serial Attached SC	SI Dis 527aade					

Figure 13.8: Component is Active once more

We shall now place the host into maintenance mode once more, but this time instead of "Ensure Accessibility", we shall choose "Full data migration". This means that although components on the host in maintenance mode will no longer be available, those components will be rebuilt elsewhere in the cluster, implying that there is full availability of the virtual machine objects.

Note: This is only possible when NumberOfFailuresToTolerate = 1 and there are 4 or more hosts in the cluster. It is not possible with 3 hosts and NumberOfFailuresToTolerate = 1, as another host needs to be available to rebuild the components. This is true for higher values of NumberOfFailuresToTolerate also.

Confirm Maintenance Mode
A host in maintenance mode does not perform any activities on virtual machines, including virtual machine provisioning. The host configuration is still enabled. The Enter Maintenance Mode task does not complete until the above state is completed. You might need to either power off or migrate the virtual machines from the host manually. You can cancel the Enter Maintenance Mode task at any time.
A There are hosts in a Virtual SAN cluster. Once the hosts are removed from the cluster, they will not have access to the Virtual SAN datastore and the state of any virtual machines on that datastore.
Move powered-off and suspended virtual machines to other hosts in the cluster
Virtual SAN data might reside on the hosts in a Virtual SAN cluster. Select an option to set the migration mechanism for the Virtual SAN data that will be enforced before the hosts enter maintenance mode.
Virtual SAN data migration:
Put the selected hosts in maintenance mode?
OK Cancel

Figure 13.9: Full data migration

Now if the components on host cs-ie-h02.ie.local are monitored, you will see that no components are placed in an "Absent" state, but rather they are rebuilt on the other hosts in the cluster. When the host enters maintenance mode, you will notice that all components of the virtual machines are active, but none reside on the host placed into maintenance mode.

Actions -							E.
Getting Started Summary Monitor Manage	Related Objects						
Issues Performance Policies Tasks Events I	Utilization						
C 🦉 🖏						Q Filter	•
Name	VM Storage Policy		Compliance Status		Last Checked		
VM home	📅 Virtual SAN Default Storage P	olicy	🗸 Compliant		14/04/2015 08:41		
🚈 Hard disk 1	F Virtual SAN Default Storage P	olicy	🗸 Compliant		14/04/2015 08:41		
86						21	tems 📑 🕶
Compliance Failures Physical Disk Placement	1	=					
vsan6-poc-test-vm2 - Hard disk 1 : Physical Disk	Placement						
)- `=						Q Filter	•
Type Component Sta	ate Host	Flash Disk Name		Flash Disk Uuid	HDD Disk Name		HDD Disk U
→ RAID 1							
Component Active	cs-ie-h04.ie.l	🚐 HP Serial At	tached SCSI Dis	52742cfb-d99d-cdc1-8ac4-1527	🚐 HP Serial Att	tached SCSI Dis	52c448ď
Component Active	cs-ie-h03.ie.l	I HP Serial At	tached SCSI Dis	52a4acab-f622-6025-bee3-746.	. 🔲 HP Serial At	tached SCSI Dis	527aade
Witness Active	cs-ie-h01.ie.l	IP Serial At	tached SCSI Dis	528ba019-e369-151e-01b3-26	IP Serial At	tached SCSI Dis	5255fd2t

Figure 13.10: All components are Active when host is in mode (full data migration)

Exit maintenance mode. This completes this part of the POC.

13.2 Remove and Evacuate a Disk

In this example, we show a feature introduced in version 6.0. This is the ability to evacuate a disk prior to removing it from a disk group.

Note: The cluster must be left in manual mode. The operations are not available when a cluster is in automatic mode.

Navigate to the cluster > Manage tab > Virtual SAN > Disk Management, and select a disk group in one of the hosts as shown below. Then select one of the capacity disks from the disk group, also shown below. Note that the disk icon with the red x becomes visible. This is not visible if the cluster is in automatic mode.

USAN6-Cluster Actions -							=
Getting Started Summary M	onitor Manage Related Objects						
Settings Scheduled Tasks A	arm Definitions Tags Permissions						
44	Disk Groups						
✓ Services						O Filter	-
vSphere DRS	Disk Group	Disks in Use	State	Virtual SAN	Fault Domain	Network Parti	Disk Format Version
vSphere HA	✓ □ cs-ie-h03.ie.local	3 of 7	Connecte	d Healthy		Group 1	
✓ Virtual SAN	Disk group (0200080000600508b1001c9c8b5f6f0d7a2be	3		Healthy			2
General	Cs-ie-h01.ie.local	3 of 7	Connecte	d Healthy		Group 1	
Disk Management	Disk group (0200070000600508b1001cb683ff0e29252f9e	3		Healthy			2
Fault Domains	✓ ☐ cs-ie-h02.ie.local	3 of 7	Connecte	d Healthy		Group 1	
Health	Disk group (0200080000600508b1001c64b76c8ceb56e8	3		Healthy			2
- Configuration	✓	3 of 7	Connecte	d Healthy		Group 1	
General	Disk group (0200080000600508b1001c29d8145d6cc192	3		Healthy			2
Licensing							
VMware EVC	4						Þ
VM/Host Groups	M6						8 items 🛛 🔒 🔻
VM/Host Rules	Disk group (020008000060050854001c20d8145d6cc1025e0fb04c4f47		e				
VM Overrides		4343417. Diak	3				
VM Overhues						Show: In use	• (3) •
Host Options	Name	Drive Type	Capacity	Virtual SAN Health Status	Operational	Transport Type	
Promes	F HP Serial Attached SCSI Disk (naa.600508b1001c29d8145d6cc	Flash	186.28 GB	Healthy	Mounted	Block Adapte	er
	HP Serial Attached SCSI Disk (naa.600508b1001c846c000c3d9	HDD	136.70 GB	Healthy	Mounted	Block Adapte	er
	HP Serial Attached SCSI Disk (naa.600508b1001c258181f0a088	HDD	136.70 GB	Healthy	Mounted	Block Adapte	er
	AA						3 items
							o nemo

Figure 13.11: Remove a disk

Make a note of the devices in the disk group, as you will need these later to rebuild the disk group. There are a number of new icons on this view of disk groups in Virtual SAN 6.0. It is worth spending some time understanding that they mean. The following table should help to explain that.

2	Add a disk to the selected disk group
_	Remove (and optionally evacuate data) from a disk in a disk group
0	Turn on the locator LED on the selected disk
0	Turn off the locator LED on the selected disk
F	Tag a device as a flash device (useful when RAID 0, non-passthru in use)
	Tag a device as a local device (useful when SAS controllers in use)

Table 13.1: Disk group icons

To continue with the option of removing a disk from a disk group and evacuating the data, click on the icon to remove a disk highlighted earlier. This pops up the following window, which gives you the option to evacuate data (selected automatically). Click "Yes" to continue:

Remov	e Disk
1	Disk "HP Serial Attached SCSI Disk (naa.600508b1001c258181f0a088f6e40dab)" is about to be removed from the disk group "0200080000600508b1001c29d8145d6cc1925e9fb94c4f474943 41". Unless the data on the disk(s) is evacuated first, removing the disk(s) might disrupt working VMs. Evacuate data (2.70 GB) Remove disk?
	Yes No

Figure 13.12: Evacuate data

When the operation completes, there should be one less disk in the disk group, but if you examine the components of your VMs, there should be none found to be in an "Absent" state. All components should be "Active", and any that were originally on the disk that was evacuated should now be rebuilt elsewhere in the cluster.

13.3 Evacuate a Disk Group

Let's repeat the previous task for the rest of the disk group. Instead of removing the original disk, let's now remove the whole of the disk group. Make a note of the devices in the disk group, as you will need these later to rebuild the disk group.

p v SANG-Cluster Actions							
Getting Started Summary	Monitor Manage Related Objects						
Settings Scheduled Tasks	Alarm Definitions Tags Permissions						
44	Disk Groups						
✓ Services						O Filter	r
vSphere DRS	Disk Group	Disks in Use	State	Virtual SAN	Fault Domain	Network Parti.	. Disk Format Versi
vSphere HA		3 of 7	Connect	ted Healthy		Group 1	
✓ Virtual SAN	Disk group (0200080000600508b1001c9c8b5f6f0d7a2be	3		Healthy			2
General	✓ ☐ cs-ie-h01.ie.local	3 of 7	Connect	ted Healthy		Group 1	
Disk Management	Disk group (0200070000600508b1001cb683ff0e29252f9e	3		Healthy			2
Fault Domains	👻 📋 cs-le-h02.ie.local	3 of 7	Connect	ted Healthy		Group 1	
Health	Disk group (0200080000600508b1001c64b76c8ceb56e8	3		Healthy			2
 Configuration 	👻 🥫 cs-ie-h04.ie.local	2 of 7	Connect	ted Healthy		Group 1	
General	Disk group (0200080000600508b1001c29d8145d6cc192	2		Healthy			2
Licensing							
VMwaro EVC	4					N. I.	
VM/Host Groups	M						8 items
VM/Host Bules	Disk group (0200080000600508b1001c29d8145d6cc1925e9fb94c4f42	(494341): Dist					
VM Overrides		1010111.010				_	
Via Overnues						Show: In us	e (2)
nost Options	Name	Drive Type	Capacity	Virtual SAN Health Status	Operational	Transport Type	e
Profiles	F HP Serial Attached SCSI Disk (naa.600508b1001c29d8145d6cc	Flash	186.28 GB	Healthy	Mounted	Block Adapt	er
	AP Serial Attached SCSI Disk (naa.600508b1001c846c000c3d9	HDD	136.70 GB	Healthy	Mounted	Block Adapt	er

Figure 13.13: Delete disk group

As before, you are prompted as to whether or not you wish to evacuate the data from the disk group. The amount of data is also displayed, and the option is selected by default. Click "Yes" to continue.

Remov	e Disk Group
	Data on the disks from the disk group "0200080000600508b1001c29d8145d6cc1925e9fb94c4f474943 41" will be deleted.
1	Unless the data on the disks is evacuated first, removing the disks might disrupt working VMs.
	✓ Evacuate data (6.41 GB) 🚯
	Remove disk group?
	Yes No

Figure 13.14: Evacuate data

Once the evacuation process has completed, the disk group should no longer be visible in the Disk Groups view.

VSANG-Cluster Actions	*						
etting Started Summary	Monitor Manage Related Objects						
Settings Scheduled Tasks	Aarm Definitions Tags Permissions						
•	Disk Groups						
 Services 	a 🕼 🔒 🦕 🐘 👌 🐂 🖉					Q Filter	0
vSphere DRS	Disk Group	Disks in Use	State	Virtual SAN	Fault Domain	Network Parti.	Disk Format Versio
vSphere HA	- cs-ie-h03.ie.local	3 of 7	Connected	Healthy		Group 1	
 Virtual SAN 	Disk group (0200080000600508b1001c9c8b5f6f0d7a2be	3		Healthy			2
General	✓ ☐ cs-ie-h01.ie.local	3 of 7	Connected	Healthy		Group 1	
Disk Management	Disk group (0200070000600508b1001cb683ff0e29252f9e	3		Healthy			2
Fault Domains		3 of 7	Connected	Healthy		Group 1	
Health	Disk group (0200080000600508b1001c64b76c8ceb56e8	3		Healthy			2
 Configuration 	cs-ie-h04.ie.local	0 of 7	Connected	Healthy		Group 1	
General							
Licensing							
Literating	4	0				1	
VMWare EVC	A6						7 items 🔒
VM/Host Groups		-					
VM/Host Rules	cs-ie-h04.ie.local: Disks						
VM Overrides						Show: In use	e (0)
Host Options	Name	Drive Type	Capacity \	/irtual SAN Health Statu	s Operational	Transport Type	
Profiles							

Figure 13.15: Disk group now removed and evacuated

Once again, if you examine the components of your VMs, there should be none found to be in an "Absent" state. All components should be "Active", and any that were originally on the disk that was evacuated should now be rebuilt elsewhere in the cluster.

13.4 Add Disk Groups Back Again

At this point, we can recreate the deleted disk group. This was already covered in section 6.1 of this POC guide. Simply select the host that the disk group was removed from, and click on the icon to create a new disk group. Once more, select a flash device and the two magnetic disk devices that you previously noted were members of the disk group. Click OK to recreate the disk group.

rst, select a single flash disk to serve as a write cache ar	nd read buffer					
				Q.F	ïlter	-
Name	Drive Type	Capacity	Transport Type			
F HP Serial Attached SCSI Disk (naa.600508b10	Flash	186.28 GB	Block Adapter			
4					1 items	
nen, select one or more HDD disks to serve as data disk	S.					
				OF	ilter	
				(-	11521	
Name			Drive Type	Capacity	Transport Type	
Name	c6a664d5d57	76299cec941)	Drive Type	Capacity 136.70 GB	Transport Type Block Adapter	
Name Image: Provide the second sec	c6a664d5d5 c4b820b4d8(76299cec941) 0f9f8acfa95)	Drive Type HDD HDD	Capacity 136.70 GB 136.70 GB	Transport Type Block Adapter Block Adapter	
Name Image: The Serial Attached SCSI Disk (naa.600508b1001	c6a664d5d57 c4b820b4d80 c846c000c3c	76299cec941))f9f8acfa95) 19114ed71b3)	Drive Type HDD HDD HDD	Capacity 136.70 GB 136.70 GB 136.70 GB	Transport Type Block Adapter Block Adapter Block Adapter	
Name Image: The Serial Attached SCSI Disk (naa.600508b1001	c6a664d5d5 c4b820b4d8 c846c000c3c c258181f0a0	76299cec941))1918ac(a95) 19114ed71b3) 88f6e40dab)	Drive Type HDD HDD HDD HDD	Capacity 136.70 GB 136.70 GB 136.70 GB 136.70 GB	Transport Type Block Adapter Block Adapter Block Adapter Block Adapter	
Name Image: All Attached SCSI Disk (naa.600508b1001	c6a664d5d57 c4b820b4d8(c846c000c3c c258181f0a0 c51f3a696fe(76299cec941))f9f8acfa95) 19114ed71b3) 88f6e40dab) 1bbbcb5096)	Drive Type HDD HDD HDD HDD HDD HDD	Capacity 136.70 GB 136.70 GB 136.70 GB 136.70 GB 136.70 GB	Transport Type Block Adapter Block Adapter Block Adapter Block Adapter Block Adapter Block Adapter	
Name Image: Provide the state of the	c6a664d5d5 c4b820b4d8 c846c000c3c c258181f0a0 c51f3a696fe0 cadff5d80ba7	76299cec941) 0f9f8acfa95) 19114ed71b3) 88f6e40dab) 0bbbcb5096) 7665b8f09a)	Drive Type HDD HDD HDD HDD HDD HDD HDD	Capacity 136.70 GB 136.70 GB 136.70 GB 136.70 GB 136.70 GB 136.70 GB	Transport Type Block Adapter Block Adapter Block Adapter Block Adapter Block Adapter Block Adapter Block Adapter	
Name Image: Provide the state of the	c6a664d5d57 c4b820b4d80 c846c000c3c c258181f0a0 c51f3a696fe0 cadff5d80ba7	76299cec941) 1918acfa95) 19114ed71b3) 88f6e40dab) 1bbbcb5096) 7665b8f09a)	Drive Type HDD HDD HDD HDD HDD HDD HDD	Capacity 136.70 GB 136.70 GB 136.70 GB 136.70 GB 136.70 GB 136.70 GB	Transport Type Block Adapter Block Adapter Block Adapter Block Adapter Block Adapter Block Adapter	
Name Image: Perial Attached SCSI Disk (naa.600508b1001	c6a664d5d57 c4b820b4d80 c846c000c3c c258181f0a0 c51f3a696fe0 cadff5d80ba7	76299cec941))1918acfa95))9114ed71b3) 88f6e40dab))bbbcb5096) 7665b8f09a)	Drive Type HDD HDD HDD HDD HDD HDD	Capacity 136.70 GB 136.70 GB 136.70 GB 136.70 GB 136.70 GB 136.70 GB	Transport Type Block Adapter Block Adapter Block Adapter Block Adapter Block Adapter Block Adapter	
Name Image: Provide the state of the	c6a664d5d57 c4b820b4d80 c846c000c3c c258181f0a0 c51f3a696fe0 cadff5d80ba7	76299cac941))1918acfa95) 19114ed71b3) 88f6e40dab) 1bbbcb5096) 7665b8f09a)	Drive Type HDD HDD HDD HDD HDD HDD	Capacity 136.70 GB 136.70 GB 136.70 GB 136.70 GB 136.70 GB 136.70 GB	Transport Type Block Adapter Block Adapter Block Adapter Block Adapter Block Adapter Block Adapter	

Figure 13.16: Recreate disk group

13.5 Turning on and off Disk LEDs

Our final maintenance task is to turn on and off the locator LEDs on the disk drives. This is a new feature of Virtual SAN 6.0. In chapter 12, we spoke about the importance of the *hpssacli* utility for removing and adding logical devices. This was a "nice to have". However for turning on and off the disk locator LEDs, the utility is a necessity when using HP controllers. Refer to section 12.10 for information on how to locate and install this utility.

Note: This is not an issue for LSI controllers, and all necessary components are shipped with ESXi for these controllers.

The icons for turning on and off the disk locator LEDs are shown in table 13.1. To turn on a LED, select a disk in the disk group and then click on the icon highlighted below.

VSAN6-Cluster Actions	*							E
Getting Started Summary	Monitor Manage Related Objects							
Settings Scheduled Tasks	Aarm Definitions Tags Permissions							
44	Disk Groups							
✓ Services	2 🗮 🚵 🤭 📜 😋					Q Filter		•
vSphere DRS	Disk Group	Disks in Use	State	Virtual SAN	Fault Domain	Network Parti	Disk Format	Version
vSphere HA	✓	3 of 7	Connect	ed Healthy		Group 1		
✓ Virtual SAN	Disk group (0200080000600508b1001c9c8b5f6f0d7a2be	3		Healthy			2	
General	✓ ☐ cs-ie-h01.ie.local	3 of 7	Connect	ed Healthy		Group 1		
Disk Management	Disk group (0200070000600508b1001cb683ff0e29252f9e	3		Healthy			2	
Fault Domains	✓	3 of 7	Connect	ed Healthy		Group 1		
Health	Disk group (0200080000600508b1001c64b76c8ceb56e8	3		Healthy			2	
- Configuration	👻 📳 cs-ie-h04.ie.local	3 of 7	Connect	ed Healthy		Group 1		
General	Bisk group (0200080000600508b1001c29d8145d6cc192	3		Healthy			2	
Licensing								
VMware EVC								
VM/Host Groups	86						8 item:	; <u>i</u> ∎∙
VM/Host Rules	Disk group (0200080000600508b1001c29d8145d6cc1925e9fb94c4f4	7494341): Disl	(S					
VM Overrides						Ohana (Jerusa	(2)	
Host Options						snow. In use	3 (3)	
Profiles	Name	Drive Type	Capacity	Virtual SAN Health Status	Operational	Transport Type		
1101100	P Senal Attached SCSI Disk (naa.600508b1001c29d8145d6cc	Flash	186.28 GB	Healthy	Mounted	Block Adapte	er	
	HP Serial Attached SCSI Disk (naa.600508b1001c846c000c3d9	HDD	136.70 GB	Healthy	Mounted	Block Adapte	er	
	HP Serial Attached SCSI Disk (naa.600508b1001c258181f0a088	HDD	136.70 GB	Healthy	Mounted	Block Adapte	er	

Figure 13.17: Turn on disk locator LED

This will launch a task to "Turn on disk locator LEDs". To see if the task was successful, go to the Monitor tab and check the Events. If there is no error, the task was successful. At this point you can also take a look in the data center and visually check if the LED of the disk in question is lit.

Once completed, the locator LED can be turned off by clicking on the "Turn off disk locator LEDs" as highlighted in the screen shot below. Once again, this can be visually checked in the data center if you wish.

VSAN6-Cluster Actions -								-
Getting Started Summary M	Ionitor Manage Related Objects							
Settings Scheduled Tasks A	larm Definitions Tags Permissions							
44	Disk Groups							
✓ Services	2 S S S S S S S S S S S S S S S S S S S					Q Filter		•
vSphere DRS	Disk Group	Disks in Use	State	Virtual SAN	Fault Domain	Network Parti	Disk Format V	'ersion
vSphere HA		3 of 7	Connect	ed Healthy		Group 1		
➡ Virtual SAN	Disk group (0200080000600508b1001c9c8b5f6f0d7a2be	3		Healthy			2	
General	✓ ☐ cs-ie-h01.ie.local	3 of 7	Connecte	ed Healthy		Group 1		
Disk Management	Disk group (0200070000600508b1001cb683ff0e29252f9e	3		Healthy			2	
Fault Domains	✓ ☐ cs-ie-h02.ie.local	3 of 7	Connecte	ed Healthy		Group 1		
Health	Disk group (0200080000600508b1001c64b76c8ceb56e8	3		Healthy			2	
- Configuration	✓ ☐ cs-ie-h04.ie.local	3 of 7	Connecte	ed Healthy		Group 1		
General	Bisk group (0200080000600508b1001c29d8145d6cc192	3		Healthy			2	
Licensing								
VMware EVC	4					1		
VM/Host Groups	86						8 items	-
VM/Host Rules	Disk group (020008000060050851001c29d8145d6cc1925e9fb94c4f42	/494341): Dis						
VM Overridee		4545417.015						
Vin Overlines						Show: In use	e (3)	
Host Options	Name Turns off the locator LED of the selected	ve Type	Capacity	Virtual SAN Health Status	Operational	Transport Type	÷	
Profiles	HP Serial Attacher disk(s).	ash	186.28 GB	Healthy	Mounted	Block Adapt	er	
	HP Serial Attached SCSI Disk (naa.600508b1001c846c000c3d9	HDD	136.70 GB	Healthy	Mounted	Block Adapt	er	
	HP Serial Attached SCSI Disk (naa.600508b1001c258181f0a088	HDD	136.70 GB	Healthy	Mounted	Block Adapt	er	

Figure 13.18: Turn off disk locator LED

This completes this section of the Virtual SAN 6.0 Proof-Of-Concept (POC) guide. Before handing over the environment to the customer, do one final check on the health and ensure all checks pass.

sues Profile Compliance Pe	erformance Tasks E	vents Resource Reservation Virtual SAN vSphere DRS vSphere HA Utilization	
	Virtual SAN Health (I	ast checked: 14 April 2015 10:15:24)	Retes
Health	Test Result	Test Name	
Proactive Tests	Passed	Advanced Vidual SAN configuration in sync	
Physical Disks	Passed	ESX VSAN Health service installation	
Virtual Disks	Passed	VSAN CLOMD liveness	
Resyncing Components	Passed	VSAN Health Service up-to-date	
	Passed	▼ Data health	
	Passed	Virtual SAN object health	
	Passed	✓ Limits health	
	Passed	After 1 additional host failure	
	Passed	Current cluster situation	
	Passed	 Network health 	
	Passed	All hosts have a VSAN vmknic configured	
	Passed	All hosts have matching multicast settings	
	Passed	All hosts have matching subnets	
	Passed	Basic (unicast) connectivity check (normal ping)	
	Passed	Hosts disconnected from VC	
	Passed	Hosts with connectivity issues	
	Passed	Hosts with VSAN disabled	
	Passed	MTU check (ping with large packet size)	
	Passed	Multicast assessment based on other checks	
	Passed	Unexpected VSAN cluster members	
	Passed	VSAN cluster partition	
	Passed	 Physical disk health 	
	Passed	Component metadata health	
	Passed	Congestion	
	Passed	Disk capacity	
	86		32 items

Figure 13.19: Final health check

14. Virtual SAN 6.1 Stretched Cluster Configuration

As per of the vSphere 6.0U1 release in September 2015, a number of new Virtual SAN features were included. The features included a Stretched Cluster solution, which is the purpose of this report. Note that the Virtual SAN version in vSphere 6.0U1 is Virtual SAN 6.1.

A good working knowledge of how Virtual SAN Stretched Cluster is designed and architected is assumed. Readers unfamiliar with the basics of Virtual SAN Stretched Cluster are urged to review the relevant documentation before proceeding with this part of the proof-of-concept. Details on how to configure a Virtual SAN Stretched Cluster are found in the <u>Virtual SAN 6.1 Stretched Cluster Guide</u>.

14.1 Virtual SAN 6.1 Stretched Cluster Network Topology

As per the *Virtual SAN 6.1 Stretched Cluster Guide*, a number of different network topologies are supported for Virtual SAN Stretched Cluster. The network topology deployed in this lab environment is a full layer 3 stretched Virtual SAN network. L3 multicast is implemented for the Virtual SAN network between data sites, and L3 unicast is implemented for the Virtual SAN network between data sites and the witness site. While VMware also supports stretched L2 between the data sites, L3 is the only supported network topology for the Virtual SAN network between the data sites.

14.2 Virtual SAN 6.1 Stretched Cluster Hosts

There are four ESXi hosts in this cluster, two ESXi hosts on data site A (the "preferred" site) and two hosts on data site B (the "secondary" site). There is one disk-group per host (all flash). The witness host/appliance is deployed on a 3rd, remote data center. The configuration is referred to as 2+2+1.

Disk Group	Disks in Use	State	Virtual SAN	Fault Domain	Network Parti.
esxi-a-pref.rainpole.com	2 of 2	Connected	Healthy	Preferred	Group 1
esxi-b-pref.rainpole.com	2 of 2	Connected	Healthy	Preferred	Group 1
esxi-c-scnd.rainpole.com	2 of 2	Connected	Healthy	Secondary	Group 1
esxi-d-scnd.rainpole.com	2 of 2	Connected	Healthy	Secondary	Group 1
esxi-e-witn.rainpole.com	2 of 2	Connected	Healthy		Group 1



VMs are deployed on both the "Preferred" and "Secondary" sites of the Virtual SAN Stretched Cluster. VMs are running/active on both sites.

14.3 Virtual SAN 6.1 Stretched Cluster Diagram

Below is a diagram detailing the POC environment used for the Stretched Cluster testing.



Figure 14.2: Virtual SAN Stretch Cluster network diagram

- This configuration uses L3 (route) for the Virtual SAN network between all sites.
- Static routes are required to enable communication between sites.
- The Virtual SAN network VLAN for the ESXi hosts on the preferred site is VLAN id 4. The gateway is 172.4.0.1.
- The Virtual SAN network VLAN for the ESXi hosts on the secondary site is VLAN id 3. The gateway is 172.3.0.1.
- The Virtual SAN network VLAN for the witness host on the witness site is VLAN id 80.
- The VM network is stretched L2 between the data sites. This is VLAN id 30. Since no VMs are run on the witness, there is no need to extend this network to the third site.

14.4 Preferred Site Details

In Virtual SAN Stretched Clusters, "preferred" site simply means the site that the witness will 'bind' to in the event of an inter-site link failure between the data sites. Thus, this will be the site with the majority of VM components, so this will also be the site where <u>all</u> VMs will run when there is an inter-site link failure between data sites.

In this example, Virtual SAN traffic is enabled on vmk1 on the hosts on the preferred site, which is sitting on routable VLAN 4.

VMkernel adapters							
2 6	/ 🗙 🗈-			Q Filter	•		
Device	Network Label	Switch	IP Address	TCP/IP Stack			
👥 vmk0	🧕 Management Networ	k 🗊 vSwitch0	172.40.0.11	Default			
📖 vmk1	👳 VMkernel-VSAN	T vSwitch0	172.4.0.11	Default			
👥 vmk2	Apreferred-vmotion-vla	n10 astretched-vswitch	h 10.10.0.11	vMotion			
4	::				•		
VMkernel ne	etwork adapter: vmk1						
	perties IP Settings Pol	icies					
]							
Port prope	erties				F		
Network	label	VMkernel-VSAN					
VLAN ID		4					
ICP/IP s	stack	Default	_				
Enabled	services	Virtual SAN traffic					

Figure 14.3: Virtual SAN preferred site networking details

Static routes need to be manually configured on these hosts. This is because the default gateway is on the management network, and if the preferred site hosts tried to communicate to the secondary site hosts, the traffic would be routed via the default gateway and thus via the management network. Since the management network and the Virtual SAN network are entirely isolated, there would be no route.

Since this is L3 everywhere, including between the data sites, the Virtual SAN interface on the preferred site, vmk1, has to route to "Secondary site (VLAN 3)" and "Witness Site (VLAN 80)".

TCP/IP Stacks								
/					Q	Filter		-
TCP/IP Stack		VMkernel Adapters		IPv4 Gatew	ay Address		IPv6 Gat	eway *
System stacks								::
Default			2	172.40.0.	1			-
4		::		·				Þ
86							3 it	tems
TCP/IP Stack: Defaul	t							
DNS Routing	IPv4 Routing Ta	ble IPv6 Routing	Table Advance	ed				
	static	route to	"Secon	dary	Site "	via vn	nk1	-
Network Addre	Prefix	Length	Gateway		Devi	ice		
172.3.0.0	24		172.4.0.1		vm	k1		
147.80.0.0	24		172.4.0.1		vm	k1		
172.4.0.0	24		0.0.0.0		vm	k1		
172.40.0.0	24		0.0.0.0		vm	k0		
0.0.0.0	0		172.40.0.	1	vm	k0		
	statio	c route to	o "Witne	ess S	ite" vi	a vmk	6	

Figure 14.4: Primary site routing table with static routes to remote sites

14.4.1 Commands to Add Static Routes

The following command is used to add static routes is as follows:

esxcli network ip route ipv4 add -n REMOTE-NETWORK -g LOCAL-GATEWAY

To add a static route from a preferred host to hosts on the secondary site in this POC:

esxcli network ip route ipv4 add -n 172.3.0.0/24 -g 172.4.0.1

To add a static route from a preferred host to the witness host in this POC:

esxcli network ip route ipv4 add -n 147.80.0.0/24 -g 172.4.0.1

Note: L3 Multicast routing must be enabled between VLAN 3 and 4. This is configured on the physical switch or router.

14.5 Secondary Site Details

The secondary site is the site that contains ESXi hosts whose objects do not bind with the witness components in the event of an inter-site link failure. However that is the only significant difference. Under normal conditions, the secondary site behaves exactly like the preferred site, and virtual machines may also be deployed there. In this POC, Virtual SAN traffic is enabled on vmk1, which is sitting on routable VLAN 3.

Mkernel ad	lapters				
ا جو 😒	/ 🗙 🗈 🖿			Q Filter	
Device	Network Label	Switch	IP Address	TCP/IP Stack	
👥 vmk0	🧕 Management Network	T vSwitch0	192.60.0.13	Default	
👥 vmk1	👳 VMkernel-VSAN	T vSwitch0	172.3.0.13	Default	
💷 vmk2	aconary-vmotion-vlan2	0 astretched-vswitch	10.20.0.13	vMotion	
4	::				
/Mkernel ne	twork adapter: vmk1				
	oerties IP Settings Polici	es			
Port prope	erues	Ularray VOAN			
Network label VMkernel-VSAN					
	stack D	ofoult			
TOP/IP S		idual OAN haffia			
Enabled	services V				

Figure 14.5: Virtual SAN secondary site networking details

Once again, static routes need to be manually configured on the Virtual SAN network interface, vmk1, to route to "Preferred site (VLAN 4)" and "Witness Site (VLAN 80)".

CP/IP Stacks								
/					(Q Filter		-
TCP/IP Stack	V	/Mkernel Adapter	5	IPv4 Gat	teway Address		IPv6	Gateway
System stacks								
Default			2	192.60	0.0.1			
4				·				P.
86								3 items
CP/IP Stack: Defa	ult							
CP/IP Stack: Defa	ult							
CP/IP Stack: Defa	ult	IPv6 Rout	ting Table Advance	ced				
TCP/IP Stack: Defa	ult IPv4 Routing Tab stati	IPv6 Rout	ting Table Advance	ced 1ess	Site"	via v	/mk1	-
CP/IP Stack: Defa	IPv4 Routing Tab	ength	to "With Gateway	ed 1ess	Site"	viaev	/mk1	-
DNS Routing Network Address 172.3.0.0	IPv4 Routing Tab	le IPv6 Rout Croute	to "Witt Gateway 0.0.0	ced 1ess	Site"	viaev wice mk1	/mk1	-
Network Address 172.3.0.0 147.80.0.0	IPv4 Routing Tab Statio Prefix L 24 24	IPv6 Rout C route	Advante Advante Advante Contraction Contra	ced IESS 1	Site"	via v vice nk1 nk1	/mk1	
CP/IP Stack: Defa DNS Routing Network Address 172.3.0.0 147.80.0.0 172.4.0.0	IPv4 Routing Tab Statio Prefix L 24 24 24 24	le IPv6 Rout C route	ting Table Advance Control Control Co	ced 1885 1	Site"	viaev nk1 nk1 nk1	/mk1	-
CCP/IP Stack: Defa DNS Routing Network Address 172:3.0.0 147:80.0.0 192:60.0.0	IPv4 Routing Tab Prefix L 24 24 24 24 24 24	IPv6 Rout	ting Table Advance Control Control Co	ced 1855 1	Site" De VI VI VI VI VI VI	vice nk1 nk1 nk1 nk1 nk1	/mk1	

Figure 14.6: Secondary site routing table with static routes to remote sites

14.5.1 Commands to Add Static Routes

The following command is used to add static routes is as follows:

esxcli network ip route ipv4 add -n REMOTE-NETWORK -g LOCAL-GATEWAY

To add a static route from a secondary host to hosts on the preferred site in this POC:

esxcli network ip route ipv4 add -n 172.4.0.0/24 -g 172.3.0.1

To add a static route from a secondary host to the witness host in this POC:

esxcli network ip route ipv4 add -n 147.80.0.0/24 -g 172.3.0.1

Note: L3 Multicast routing must be enabled between VLAN 3 and 4. This is configured on the physical switch or router.

14.6 A note on IGMP v3

IGMP Version 2, specified in [RFC-2236], added support for "low leave latency". That is, a reduction in the time it takes for a multicast router to learn that there are no longer any members of a particular group present on an attached network.

IGMP Version 3 adds support for "source filtering". That is, the ability for a system to report interest in receiving packets *only* from specific source addresses, or from *all but* specific source addresses, sent to a particular multicast address.

It should be noted that in our POC testing with the DELL network switch, the Stretched Cluster would not configure properly after failures until the network switch was forced to talk IGMP v3 between VLANs.

Recommendation: Use IGMP v3 for multicast configurations.

14.7 Witness Site Details

The witness site only contains a single host for the Stretched Cluster, and the only VM objects stored on this host are "witness" objects. No data components are stored on the witness host. In this POC, we are using the witness appliance, which is an "ESXi host running in a VM". If you wish to use the witness appliance, it should be downloaded from VMware. This is because it is preconfigured with various settings, and also comes with a pre-installed license. Note that this download requires a login to <u>My VMware</u>.

Alternatively, customers can use a physical ESXi host for the appliance.

Virtual SAN traffic must be enabled on the Virtual SAN interface of the witness appliance, in this case vmk1, which is sitting on routable VLAN 80 (tagged on the underlying physical ESXi).

VMkernel ad	lapters				
2 6	🖊 🗙 🗈 🖿			Q Filter	•
Device	Network Label	Switch	IP Address	TCP/IP Stack	
👥 vmk0	🧕 Management Netw	ork 🗊 vSwitch0	147.70.0.15	Default	
ị vmk1	👳 witnessPg	打 witnessSwitch	147.80.0.15	Default	
4					Þ
VMkernel ne	etwork adapter: vmk1 perties IP Settings F	Policies			
Port prope	erties				•
Network	label	witnessPg			
VLAN ID		None (0)			
TCP/IP s	stack	Default			
Enabled	l services	Virtual SAN traffic			

Figure 14.7: Virtual SAN witness host networking details

Once again, static routes should be manually configured on Virtual SAN vmk1 to route to "Preferred site (VLAN 4)" and "Secondary Site (VLAN 3)".

/			Q Filter	r -
TCP/IP Stack	VMkernel Adap	ters IPv4 Gate	eway Address	IPv6 Gateway
System stacks				::
Default		2 147.70.	0.1	
4				•
84				3 items
CP/IP Stack: Default				
DNS Routing IPv	4 Routing Table IPv6 Ro	outing Table Advanced		
			Q. Filter	
Network Address	atic route to	o <u>Secondary</u>	Site via	VMKI
147.80.0.0 🦯	24	0.0.0.0	vmk1	
147.70.0.0	24	0.0.00	vmk0	
172.3.0.0	24	147.80.0.1	vmk1	
172.4.0.0	24	147.80.0.1	vmk1	
0.0.0.0	0	147.70.0.1	vmk0	

Figure 14.8: Witness host routing table with static routes to remote sites

14.7.1 Commands to Add Static Routes

The following command is used to add static routes is as follows:

esxcli network ip route ipv4 add -n REMOTE-NETWORK -g LOCAL-GATEWAY

To add a static route from the witness host to hosts on the preferred site in this POC:

esxcli network ip route ipv4 add -n 172.4.0.0/24 -g 172.80.0.1

To add a static route from the witness host to hosts on the secondary site in this POC:

esxcli network ip route ipv4 add -n 147.3.0.0/24 -g 172.80.0.1

Note: L3 Multicast is not required for Witness Virtual SAN Traffic. Also VLAN tagging is enabled on ESXi host hosting witness appliance.

14.8 vSphere HA Settings

vSphere HA plays a critical part in Stretched Cluster. HA is required to restart virtual machines on other hosts and even the other site depending on the different failures that may occur in the cluster. The following section covers the recommended settings for vSphere HA when configuring it in a Stretched Cluster environment.

14.8.1 Response to Host Isolation

The recommendation is to "Power off and restart VMs" on isolation, as shown below. In cases where the virtual machine can no longer access the majority of its object components, it may not be possible to shut down the guest OS running in the virtual machine. Therefore the "Power off and restart VMs" option is recommended.

Failure	Respo	nse	Details
Host failure	Resta	rt VMs	Restart VMs using VM restart priority ordering.
Host Isolation	Power	off and restart VMs	VMs on isolated hosts will be powered off and restarted on available hosts.
Datastore with Permanent Device Loss	Disab	led	Datastore protection for All Paths Down and Permanent Device Loss is disabled.
Datastore with All Paths Down	Disab	led	Datastore protection for All Paths Down and Permanent Device Loss is disabled.
Guest not heartbeating	Disab	led	VM and application monitoring disabled.
Response for Host Isolati	on	When Disabled is selected, virtual failure. In addition, they remain Pro	machines are not restarted in the event of a host tected when Turn on vSphere HA is enabled.
Response for Datastore w Permanent Device Loss (I	rith PDL)	Disabled	•
Response for Datastore w Paths Down (APD)	ith All	Disabled	•
Delay for VM failover for AF	D	3 minutes	
Response for APD recove after APD timeout	ry	Disabled	•
VM monitoring sensitivity		Preset Low High Custom	
			OK Cance

Figure 14.9: vSphere HA Host Isolation recommended setting

14.8.2 Admission Control

If a full site fails, the desire is to have all virtual machines run on the remaining site. To allow a single data site to run all virtual machines if the other data site fails, the recommendation is to set Admission Control to 50% for CPU and Memory as shown below.

····	
Define failover capacity by reserving	a percentage of the cluster resources.
Reserved failover CPU capacity:	50 🔷 % CPU
Reserved failover Memory capacity	50 🔷 % Memory

Figure 14.10: vSphere HA Admission Control setting recommendation

14.8.3 Advanced Settings

The default isolation address uses the default gateway of the management network. This will not be useful in a Virtual SAN Stretched Cluster, when the Virtual SAN network is broken. Therefore the default isolation response address should be turned off. This is done via the advanced setting *das.usedefaultisolationaddress* to false.

To deal with failures occurring on the Virtual SAN network, VMware recommends setting two isolation addresses, each of which is local to one of the data sites. In this POC, one address is on VLAN 4, which is reachable from the hosts on the preferred sites. The other address is on VLAN 3, which is reachable from the hosts on the secondary site. Use advance settings *das.isolationaddress0* and *das.isolationaddress1* to set these isolation addresses respectively.

Configuration Parameters	Add Delete	- Prilmary
	Option	Value
	das.isolationaddress0	172.4.0.1
	das.isolationaddress1	172.3.0.1
	das.respectVmHostSoftAffinityRules	true
	das.respectVmVmAntiAffinityRules	false
	das.usedefaultisolationaddress	false Secondary

Figure 14.11: vSphere HA advanced options isolation address recommendations

These advanced settings are added in the Advanced Options > Configuration Parameter section of the vSphere HA UI. The other advanced settings get filled in automatically based on additional configuration steps. There is no need to add them manually.

14.9 VM Host Affinity Groups

The next step is to configure VM/Host affinity groups. This allows administrators to automatically place a virtual machine on a particular site when it is powered on. In the event of a failure, the virtual machine will remain on the same site, but placed on a different host. The virtual machine will be restarted on the remote site only when there is a catastrophic failure or a significant resource shortage.

To configure VM/Host affinity groups, the first step is to add hosts to the host groups. In this example, the Host Groups are named Preferred and Secondary, as shown below.

Preferred	Host Group	
VM/Host Group Members	1 Secondary	Host Group
Add Remove	VM/Host Group Members	
Preferred Group Members	Add Remove	
sxi-b-pref.rainpole.com	Secondary Group Members	
esxi-a-pref.rainpole.com	esxi-d-scnd.rainpole.com	
	esxi-c-scnd.rainpole.com	

Figure 14.12: Host affinity groups

The next step is to add the virtual machines to the host groups. Note that these virtual machines must be created in advance.

Preferred-Site-V	7M	VM Group					
VM/Host Group Mem	VM/Host Group Members						
Add Rem	love						
Preferred-Site-VM Grou	🖽 Secondary-Site-VMs		VM Group				
preferrred-vm8	VM/Host Group Members						
	Add Remove						
	Secondary-Site-VMs Group Members						
	secondary-vm2						
	secondary-vm3						

Figure 14.13: Host affinity groups with VMs

Note that these VM/Host affinity rules are "should" rules and not "must" rules. "Should" rules means that every attempt will be made to adhere to the affinity rules. However, if this is not possible (due lack of resources), the other site will be used for hosting the virtual machine.

Also note that the vSphere HA rule settings is set to "should". This means that if there is a catastrophic failure on the site to which the VM has affinity, HA will restart the virtual machine on the other site. If this was a "must" rule, HA would not start the VM on the other site.

VM/Host Rules					
Add Edit Delete					
Name	Туре		Enabled	Conflicts	Defined By
🧯 Preferred-Rule	Run VMs on Hosts		Yes	0	User 🗸 🕶
VM/Host Rule Details	=	=			
Virtual Machines that are members of the VM Group should run on hosts that are members of the Host Group.					
Add Remove		Add Ren	nove		
Preferred-Site-VM Group Memb	bers	Preferred Group Memi	bers		
preferrred-vm8		esxi-b-pref.rain	pole.com		
preferred-vm1		esxi-a-pref.rain	pole.com		
preferred-vm/					
preferred-vm4					
preferred ym6					
preferred-vm3					
preferred-vm10					
preferred-vm5					
preferred-vm2					
vSphere HA Rule Settings					Edit
vSphere HA can enforce VM/Host rules when restarting virtual machines.					
VM anti-affinity rules Ignore rules					
VM to Host affinity rules vSphere HA should espect rules during failover					

Figure 14.14: Set vSphere HA VM to Host affinity rules to "should", not "must"

The same settings are necessary on both the primary VM/Host group and the secondary VM/Host group.

VM/Host Rules						
Add Edit D	elete					
Name	Туре		Enabled	Conflicts	Defined By	
📁 Secondary-Rule	Run VMs on Hosts		Yes	0	User ::	
VM/Host Rule Details Virtual Machines that are members of the VM Group should run on hosts that are members of the Host Group.						
Add Remove		Add Remo	ove			
Secondary-Site-VMs Group Member	k	Secondary Group Memb	bers			
secondary-vm2		esxi-d-scnd.rain	pole.com			
isecondary-vm3		esxi-c-scnd.rain	pole.com			
secondary-vm1						
secondary-vm4						
Secondary-vm5						
vSphere HA Rule Settings					Edit	
vSphere HA can enforce VM/Host rules when restarting virtual machines.						
VM anti-affinity rules Ignore rules						
VM to Host affinity rules VSp	here HA should espect rules	during failover				

Figure 14.15: Set vSphere HA VM to Host affinity rules to "should" on Secondary too

14.10 DRS Settings

In this POC, partially automated mode has been chosen. However, this could be set to Fully Automated if customers wish, but note that it should be changed back to partially automated when a full site failure occurs. This is to avoid failback of VMs occurring whilst rebuild activity is still taking place. More on this later.

vSphere DRS is Turned ON	Schedule DRS 🚯 Edit.
 DRS Automation 	Partially Automated
 Power Management 	Off
 Advanced Options 	None

Figure 14.16: Virtual SAN stretch cluster DRS settings

15. Virtual SAN Stretched Cluster Network Failover Scenarios

In this section, we will look at how to inject various network failures in a Virtual SAN Stretched Cluster configuration. We will see how the failure manifests itself in the cluster, focusing on the Virtual SAN health check and the alarms/events as reported in the vSphere web client.

15.1 Network Failure between Secondary Site and Witness



Figure 15.1: Path failure between secondary site and witness site

15.1.1 Trigger the Event

To make the secondary site lose access to the witness site, one can simply remove the static route on the witness host that provides a path to the secondary site.

On witness host issue:

```
esxcli network ip route ipv4 remove -g 147.80.0.1 -n 172.3.0.0/24
```

On secondary host(s) issue:

esxcli network ip route ipv4 remove -g 172.3.0.1 -n 147.80.0.0/24

15.1.2 Cluster Behavior on Failure

To begin with, the **Cluster Summary** view shows one configuration issue related to 0 witness hosts.

ting Started Summary Monitor Manage Related Objects		
stretched-cluster	CPU	FREE: 152.54 GHz
Total Processors: 64	USED: 1.00 GHz	CAPACITY: 153.54 GHz
Total vMotion Migrations: 54	MEMORY	FREE: 321.02 GB
	USED: 62.60 GB	CAPACITY: 383.63 GB
	STORAGE	FREE: 8.38 TB
	USED: 307.09 GB	CAPACITY: 8 68 TB

Figure 15.2: Cluster summary view – 0 witness hosts

This same event is visible in the Cluster > Monitor > Issues > All Issues view.



Figure 15.3: Cluster Issue – missing witness

Note that this event may take some time to trigger. Next, looking at the **health check** alarms, a number of them get triggered (Triggering alarms from health check test failures is a new feature in Virtual SAN 6.1).

• Virtual SAN Health Alarm "Virtual SAN cluster partition"	Acknowledge	Reset To Green
Virtual SAN Health Alarm 'Stretched cluster health'	Acknowledge	Reset To Green
Virtual SAN Health Alarm 'MTU check (ping with large packet size)'	Acknowledge	Reset To Green
♦ Virtual SAN Health Alarm 'Data health'	Acknowledge	Reset To Green
(10 issues total - show all)		

Figure 15.4: Virtual SAN Health Alarms triggered

In the **Cluster summary** view, an error is also shown. This directs the administrator to go to "Monitor Virtual SAN health".

Virtual SAN is Turned ON			Edit
Add disks to storage Manua	1		
Resources			
Hosts		4 hosts	
Flash disks in use		8 of 8 eligible	
Data disks in use		0 of 0 eligible	
Total capacity of Virtual SAN dat	astore	5.04 TB	
Free capacity of Virtual SAN dat	astore	4.76 TB	
Network status		Misconfiguration detected	Monitor VSAN health
On-disk Format Version			Upgrade
Disk format version	2.0 (lates	st)	
Disks with outdated version	🛛 0 of	8	

Figure 15.5: Virtual SAN cluster summary view

On navigating to the Virtual SAN Health > Monitor view, there are a lot of checks showing errors. One should also note that there is a set of new Stretched Cluster health checks in 6.1. These are also failing.

Virtual SAN Health	n (Last checked: Today at 16:44)	Retest
Test Result	Test Name	
😣 Failed	✓ Data health	<u>ـ</u>
😣 Failed	Virtual SAN object health	
😆 Failed	 Network health 	
😣 Failed	Basic (unicast) connectivity check (normal ping)	
😣 Failed	MTU check (ping with large packet size)	
😣 Failed	Virtual SAN cluster partition	
🔥 Warning	All hosts have matching subnets	
Passed	All hosts have a Virtual SAN vmknic configured	
Passed	All hosts have matching multicast settings	
Passed	Hosts disconnected from VC	
Passed	Hosts with connectivity issues	
Passed	Hosts with Virtual SAN disabled	
Passed	Multicast assessment based on other checks	
Passed	Unexpected Virtual SAN cluster members	
😣 Failed	✓ Stretched cluster health	
😣 Failed	Stretched cluster without a witness host	

Figure 15.6: Virtual SAN Health Check detects the problems

One final place to examine is the virtual machines. Navigate to a VM on the secondary site, then Monitor > Policies > Physical Disk Placement. It should show the witness absent from secondary site perspective. However the virtual machines should still be running and fully accessible.
secondary-wn1 Actions -						
etting Started Summary Monitor Manage Related Objects						
Issues Performance Policies Tasks Events Utilization						
					Q Filter	
Name	VM Storage Poli	ay	Compliance Status	Last Checked		
🔁 VM home	🚰 Virtual SAI	N Default Storage Policy	Oncompliant	27/08/2015 18:42		
🚈 Hard disk 1	📑 Virtual SAI	N Default Storage Policy	8 Noncompliant	27/08/2015 18:42		
M	A 2 items					
Compliance Failures Physical Dis	sk Placement					
secondary-vm1 - VM home : Physica	I Disk Placement					
)— "≣				(A Filter -	
Туре	Component State	Host	Flash Disk Name	Flash Disk Uuid	HDD Disk Name	
Witness	Absent	esxi-e-witn.rainpole.com	Local VMware Disk (mpx.vm	52c2ea15-c820-e930-de9c-e3b.	🚐 Local VMware Disk (r	
→ RAID 1						
Component	Active	esxi-d-scnd.rainpole.com	Local FUSIONIO Disk (eui.4	52473e4d-0c88-01f2-75fc-0b60	Local ATA Disk (t10.4	
Component	Active	esxi-a-pref.rainpole.com	Local FUSIONIO Disk (eui.c	521b0339-7379-1833-0310-57	Local ATA Disk (t10.4	
		_				

Figure 15.7: VM shows the witness component is absent

Returning to the health check client, selecting "Basic (unicast) connectivity check (normal ping), you can see that the Secondary Site can't talk to witness or vice versa.

-					
3 Failed	Basic (unicast) connectivity check (normal p	ing)			
S Failed	MTU check (ping with large packet size)				
S Failed	Virtual SAN cluster partition				
🔥 Warning	All hosts have matching subnets				
Passed	All hosts have a Virtual SAN vmknic configu	red			
4			29 items 🔒		
asic (unicast) connecti	vity check (normal ping)		Ask VMwar		
erforms a small packet	size ping test from all bosts to all other bosts				
Ping results Only fa	ailed pings				
Ping results Only fa	To Host	Ping result	1 A To Device		
Ping results Only fa rom Host	To Host	Ping result S Failed	1 ▲ To Device vmk1		
Ping results Only fa rom Host esxi-d-scnd.rainpole esxi-d-scnd.rainpole	To Host 2.com Sexie-with.rainpole.com 2.com sexi-c-scnd.rainpole.com	Ping result Failed Passed	1 ▲ To Device vmk1 vmk1		
Ping results Only fa rom Host sci-d-scnd.rainpole esxi-d-scnd.rainpole esxi-d-scnd.rainpole esxi-d-scnd.rainpole sci-d-scnd.rainpole	To Host Com Sexie-with.rainpole.com e.com sexi-c-scnd.rainpole.com e.com sexi-c-scnd.rainpole.com	Ping result S Failed Passed Passed	1 To Device vmk1 vmk1 vmk1		
Ping results Only farmer rom Host esxi-d-scnd.rainpole esxi-d-scnd.rainpole esxi-d-scnd.rainpole esxi-d-scnd.rainpole esxi-d-scnd.rainpole esxi-d-scnd.rainpole esxi-d-scnd.rainpole	To Host Com Contraction Contraction Com Contraction Com Contraction Com Contraction Com Contraction Com Contraction	Ping result S Failed Passed Passed S Failed	1 To Device vmk1 vmk1 vmk1 vmk1		
Ping results Only fa rom Host esxi-d-scnd rainpole esxi-d-scnd rainpole esxi-d-scnd rainpole esxi-e-with.rainpole esxi-e-with.rainpole<	To Host Com Contraction Contr	Ping result Failed Passed Passed Failed Passed	1 ▲ To Device vmk1 vmk1 vmk1 vmk1		
Ping results Only fa rom Host esxi-d-scnd rainpole esxi-d-scnd rainpole esxi-d-scnd rainpole esxi-e-witn.rainpole esxi-e-witn.rainpole<	To Host com com com com com com com com	Ping result Failed Passed Passed Failed Passed Passed Passed	1 ▲ To Device vmk1 vmk1 vmk1 vmk1 vmk1 vmk1 vmk1		
Ping results Only fa rom Host esxi-d-scnd.rainpole esxi-d-scnd.rainpole esxi-d-scnd.rainpole esxi-e-witn.rainpole esxi-e-witn.rainpole esxi-e-witn.rainpole esxi-e-witn.rainpole esxi-e-witn.rainpole esxi-e-witn.rainpole	To Host Com	Ping result	1 To Device vmk1 vmk1 vmk1 vmk1 vmk1 vmk1 vmk1 vmk1 vmk1		
Ping results Only fa rom Host esxi-d-scnd.rainpole esxi-d-scnd.rainpole esxi-d-scnd.rainpole esxi-d-scnd.rainpole esxi-e-witn.rainpole esxi-e-witn.rainpole esxi-e-witn.rainpole esxi-e-witn.rainpole esxi-a-pref.rainpole	To Host Com Com Comparison Comparison Com	Ping result	1 To Device vmk1 vmk1		
Ping results Only fa rom Host esxi-d-scnd.rainpole esxi-d-scnd.rainpole esxi-d-scnd.rainpole esxi-d-scnd.rainpole esxi-e-with.rainpole esxi-e-with.rainpole esxi-e-with.rainpole esxi-e-with.rainpole esxi-a-pref.rainpole esxi-a-pref.rainpole esxi-a-pref.rainpole	To Host a.com is esxi-e-scnd.rainpole.com a.com is esxi-b-pref.rainpole.com a.com is esxi-b-pref.rainpole.com a.com is esxi-b-pref.rainpole.com com is esxi-b-pref.rainpole.com com is esxi-b-pref.rainpole.com com is esxi-a-pref.rainpole.com com is esxi-a-pref.rainpole.com com is esxi-c-scnd.rainpole.com com is esxi-c-scnd.rainpole.com com is esxi-a-scnd.rainpole.com com is esxi-a-scnd.rainpole.com com is esxi-a-scnd.rainpole.com	Ping result	1 To Device		
Ping results Only fa rom Host esxi-d-scnd.rainpole esxi-d-scnd.rainpole esxi-d-scnd.rainpole esxi-d-scnd.rainpole esxi-e-with.rainpole esxi-e-with.rainpole esxi-e-with.rainpole esxi-a-pref.rainpole esxi-a-pref.rainpole esxi-a-pref.rainpole esxi-d-scnd.rainpole	To Host a.com is esxi-e-with.rainpole.com a.com is esxi-o-scnd.rainpole.com a.com is esxi-o-pref.rainpole.com a.com is esxi-d-scnd.rainpole.com com is esxi-d-scnd.rainpole.com	Ping result	1 To Device		

Figure 15.8: Virtual SAN Health Check ping test results

15.1.3 Conclusion

Loss of the witness does not impact the running virtual machines on the secondary site. There is still a quorum of components available per object, available from the data sites. Since there is only a single witness host/site, and only three fault domains, there is no rebuilding/resyncing of objects.

15.1.4 Repair the Failure

Add back the static routes that were removed earlier, and rerun the health check tests. Verify that all tests are passing before proceeding. **Remember to test one thing at a time**.

15.2 Network Failure between Preferred Site and Witness



Figure 15.9: Path failure between preferred site and witness site

15.2.1 Trigger the Event

To make the preferred site lose access to the witness site, one can simply remove the static route on the witness host that provides a path to the preferred site.

On witness host issue:

esxcli network ip route ipv4 remove -g 147.80.0.1 -n 172.4.0.0/24

On preferred host(s) issue:

esxcli network ip route ipv4 remove -g 172.4.0.1 -n 147.80.0.0/24

15.2.2 Cluster Behavior on Failure

As per the previous test, it may take some time for alarms to trigger when this event occurs. However, the events are similar to those seen previously.

ing Started Summary Monitor Manage Related Objects		
stretched-cluster Total Processors: 64 Total vMotion Migrations: 54	CPU USED: 1.00 GHz MEMORY USED: 62.60 GB STORAGE	FREE: 152.54 GHz CAPACITY: 153.54 GHz FREE: 321.02 GB CAPACITY: 383.63 GB FREE: 8.38 TB
	USED: 207.09 GB	CAPACITY: 8 69 TR

Figure 15.10: Cluster summary view – 0 witness hosts

anage Related Objects e Tasks Events Resource Reservation	vSphere DRS	vSphere HA	Utilization	Virtual SAN	
e Tasks Events Resource Reservation	vSphere DRS	vSphere HA	Utilization	Virtual SAN	
			Q	Filter	
	Туре	Trigg	ger Time	St	atus
d 0 witness hosts on stretched cluster. The	Configuratio	on Iss		4	Warning
	10 witness hosts on stretched cluster. The	10 witness hosts on stretched cluster. The Configuration	1 0 witness hosts on stretched cluster. The Configuration Iss	1 0 witness hosts on stretched cluster. The Configuration Iss	Type Trigger Time St 1 0 witness hosts on stretched cluster. The Configuration Iss

Figure 15.11: Cluster issue – missing witness

One can also see various health checks fail, and their associated alarms being raised.

Navigator I	Stretched-cluster Actions	-		=*
Home	Getting Started Summary Mo	nitor Manage R	lated Objects	
👻 🕑 mgmt-vc01.rainpole.com 🔽	Issues Performance Profile C	ompliance Tasks	vents Resource Reservation vSphere DRS vSphere HA Utilization Virtual SAN	
Stretched-Datacenter				
	**	Virtual SAN Health	Last checked: Today at 20:20)	Retest
esxi-a-pref.rainpole.com	Physical Disks	Test Result	Test Name	
esxi-b-pref.rainpole.com	Virtual Disks	A Failed	Data health	
esxi-c-scnd.rainpole.com	Resyncing Components	C Epiled	 Network health 	
esxi-d-scnd.rainpole.com		V Falleu	P Network freatur	
preferred-vm1	Health	8 Failed	 Stretched cluster health 	
preferred-vm10	Proactive Tests	🔥 Warning	 Virtual SAN HCL health 	
preferred-vm2		Passed	 Cluster health 	
preferred-vm3		Passed	Limits health	
preferred-vm4		Passed	Physical disk health	
preferred-vm5		•		

Figure 15.12: Virtual SAN Health Check detects the problems

Just like the previous test, the witness component goes absent.

preferred-vm1 Actions	*						=
Getting Started Summary	Monitor Manage Related	Objects					
Issues Performance Polic	ies Tasks Events Utilization	n					
C 🍓 🚓						Q Filter	•)
Name	VM Storage Policy		Compliand	e Status	Last Ch	ecked	
🗀 VM home	Virtual SAN Defau	ult Storage Po	O Nonc	ompliant	27/08	2015 22:29	
🚈 Hard disk 1	Virtual SAN Defau	ult Storage Po	8 Nonc	ompliant	27/08	2015 22:29	
86							2 items 📑 👻
Compliance Failures Ph	ysical Disk Placement						
preferred-vm1 - VM home :	Physical Disk Placement						
)— ·=						Q Filter	•
Туре	Compone	Host		Flash Disk Name		Flash Disk Uuid	I.
Witness	Absent	esxi-e-	witn.rai	Local VMware D	sk (mpx.vm	52c2ea15-c	820-e930-de9c-
* RAID 1							
Component	Active	esxi-b-	pref.rai	Local FUSIONIO	Disk (eui.a	5273dae8-ft	b30-4772-c1f3-€
Component	Active	esxi-c-	scnd.r	Local FUSIONIO	Disk (eui.d	5241c4d3-6	804-20d3-66ab

Figure 15.13: VM's storage policy is out of compliance

We did not look at the "Data health" health check during the previous test. If this health check **"Virtual SAN object health"** is selected, it displays X number of objects with "reduced-availability-with-no-rebuild-delay-timer". In this POC, there are 52 objects impacted by the failure.

Virtual SAN Health (Las	t checked: Today at 20:39)			Retest
Test Result	Test Name			
😣 Failed				
🙁 Failed	Virtual SAN object health			
😣 Failed	 Network health 			
😣 Failed	 Stretched cluster health 			
🔥 Warning	 Virtual SAN HCL health 			
Passed	 Cluster health 			
Passed	 Limits health 			
Passed	 Physical disk health 			
A4			Ę	3 items 🛛 🔒 🕶
Virtual SAN object heal	th		Repair Objects Immediately	Ask VMware
Checks the health statu	is of all Virtual SAN objects.			0
Object health overview				
Health/Objects		Number	1 V Objects UUID	
reduced-availability-wi	th-no-rebuild-delay-timer	52	328bdd55-26f3	3-580e-2 🔺
inaccessible		0		::
data-move		0		
healthy		0		
reduced-availability-wi	th-no-rebuild	0		*

Figure 15.14: Virtual SAN Health Check for object health

This health check behavior appears whenever components go 'absent' and Virtual SAN is waiting for the 60-minute *clomd* timer to expire before starting any rebuilds. If an administrator clicks on "Repair Objects Immediately", the objects switch state and now the objects are no longer waiting on the timer, and will start to rebuild immediately under general circumstances. However in this POC, with only three fault domains and no place to rebuild witness components, there is no syncing/rebuilding.

15.2.3 Conclusion

Just like the previous test, a witness failure has no impact on the running virtual machines on the preferred site. There is still a quorum of components available per object, as the data sites can still communicate. Since there is only a single witness host/site, and only three fault domains, there is no rebuilding/resyncing of objects.

15.2.4 Repair the Failure

Add back the static routes that were removed earlier, and rerun the health check tests. Verify that all tests are passing before proceeding. **Remember to test one thing at a time.**

15.3 Network Failure between Witness and Both Data Sites



Figure 15.15: Complete witness site outage

15.3.1 Trigger the Event

To introduce a network failure between the preferred and secondary data sites and the witness site, one can simply remove the static route on the witness host that provides a path to both the preferred and secondary sites, and remove the static routes to the witness on the preferred and secondary hosts.

On Witness host issue:

esxcli network ip route ipv4 remove -g 147.80.0.1 -n 172.3.0.0/24 esxcli network ip route ipv4 remove -g 147.80.0.1 -n 172.4.0.0/24 On Preferred host(s) issue: esxcli network ip route ipv4 remove -g 172.4.0.1 -n 147.80.0.0/24 On Secondary host(s) issue: esxcli network ip route ipv4 remove -g 172.3.0.1 -n 147.80.0.0/24

15.3.2 Cluster Behavior on Failure

The events observed are for the most part identical to those observed in failure scenario #1 and #2.

15.3.3 Conclusion

When the Virtual SAN network fails between the witness site and both the data sites (as in the witness site fully losing its WAN access), it does not impact the running virtual machines. There is still a quorum of components available per object, available from the data sites. However, as explained previously, since there is only a single witness host/site, and only three fault domains, there is no rebuilding/resyncing of objects.

15.3.4 Repair the Failure

Add back the static routes that were removed earlier, and rerun the health check tests. Verify that all tests are passing. **Remember to test one thing at a time.**

16. Further Information

16.1 VMware Virtual SAN Community

16.2 Links to Existing Documentation

- <u>VMware Virtual SAN Resources</u>
- <u>Administering VMware Virtual SAN</u>
- <u>VMware Compatibility Guide</u>
- <u>VMware Virtual SAN Diagnostics and Troubleshooting Reference Manual</u>
- <u>VMware Virtual SAN 6.0 Design and Sizing Guide</u>
- <u>Virtual SAN Hosted Evaluation</u>
- VMware Virtual SAN Health Check Plugin Guide

16.3 VMware Support

- <u>My VMware</u>
- <u>How to file a Support Request in My VMware</u>
- Location of log files for VMware Products
- Location of ESXi 5.1 and 5.5 log files
- <u>Collecting Virtual SAN support logs and uploading to VMware</u>

Appendix A—Fault Domains

In this four-node environment, we now look at the benefits of failure domains, a new feature introduced in vSphere 6.0. In this scenario, we will assume that the 4 nodes are in two racks, something as follows.



Figure A.1 Fault Domains

The objective now is to match the rack with fault domains. This implies that if there is a rack failure, the virtual machine components will have been distributed in a fashion such that they remain available even when a complete rack fails.

A1. Setting up Fault Domains

As shown above, we will create three fault domain, two of which only contain a single host, but one which contains two hosts. Navigate to the Manage tab, and under Virtual SAN select Fault Domains as shown below. Initially, there are no hosts in any fault domains.

VSAN6-Cluster Actions -	U VSAN6-Cluster Actions -					
Getting Started Summary Mo	Initor Manage Related Objects					
Settings Scheduled Tasks Alarm Definitions Tags Permissions						
44	Virtual SAN Cluster Fault Domains					
▼ Services	You can group Virtual SAN hosts that could potentially fail together by creating a fault domain and assigning one or more hosts to it. Failure of all hosts within a single fault domain					
vSphere DRS	one failure. If fault domains are specified, Virtual SAN will never put more than one replica of the same object in the same fault domain.					
vSphere HA	+ '- '≡ C'	-				
	Fault Domain/Host					
General	✓ Hosts not in fault domain (4 hosts)					
Disk Management	Cs-le-h03,le,local					
Fault Domains	Fault Domains ü cs-le-h01.le.local Health ü cs-le-h02.le.local ~ Configuration ü cs-le-h04.ie.local					
Health						
- Configuration						

Figure A.2 No hosts in Fault Domains

Click on the green "+" symbol to create a fault domain. Initially, we will add host csie-h01.ie.local to the first fault domain. Let's call the domain FD1.

New Fault Domain		?
Name: FD1 Select the hosts to move into th	nis fault domain.	
	Show: Hosts not in fault domains	 ▼ Q Filter ▼
Host		Fault Domain
🔲 📱 cs-ie-h03.ie.local		
S-ie-h01.ie.local		
cs-ie-h02.ie.local		
cs-ie-h04.ie.local		
A4.		4 items 🕒 🗸
		OK Cancel

Figure A.3 Add single host to Fault Domain FD1

Repeat this operation for the second fault domain, but this time add host cs-ie-h02.ie.local to this domain FD2. For the third fault domain, add the remaining two hosts, cs-ie-h03 and cs-ie-h04 as shown here.

New Fault Domain		?
Name: FD3		
Select the hosts to move into the	nis fault domain.	
	Show: Hosts not in fault domains	 ▼ Q Filter ▼
Host		Fault Domain
🗹 📋 cs-ie-h03.ie.local		
🗹 🔋 cs-ie-h04.ie.local		
A4		2 items 📑 🗸
		OK Cancel

Figure A.4 Add two hosts to third Fault Domain FD3

At this point, three fault domains have been created.



Figure A.5 Fault Domain Overview

A2. Create a Policy to Leverage Fault Domains

The next step is to create a VM storage policy that highlights the behavior of fault domains. In the event of a failure of any single rack, there should still be enough components available belonging to the VM to continue running. In essence, there should still be a full copy of the data even when a rack fails. Let's create a policy so that we can observe how a VM's components. We have chosen a policy that has *NumberOfFailuresToTolerate* = 1 and *NumberOfDiskStripesPerObject* = 3.

We have already created policies back in chapter 9. Here are the steps once again.

١	/mware® vSphere Web Client			
	Navigator I	To VM Storage Policies		
	(Home 🕑	Objects		
	📅 VM Storage Policies 🗾			
	R Stripe/Width=2	🔁 🥖 🌯 🚰 🔯 Actions	*	
		Name	Description	VC
	E Virtual SAN Default Storage Policy	F Virtual SAN Default Storage	Storage policy used as default f	😰 ie-vcsa-09.ie.local
	E VVol No Requirements Policy	📅 VVol No Requirements Policy	Allow the datastore to determine	😰 ie-vcsa-09.ie.local
		StripeWidth=2	NumberOfDiskStripesPerObject	😰 ie-vcsa-09.ie.local

Figure A.6 Navigate to VM Storage Policies

Click on the "Create New Policy" icon. Give it a name and an optional description.

🚰 Create New VM Storage Policy		? »
 Create New VM Storage Policy 1 Name and description 2 Rule-Sets 2a Rule-Set 1 3 Storage compatibility 4 Ready to complete 	Name and description Enter a name and (optional) description vCenter Server: ie-vcsa-09.ie.local Name: StripeWidth=3 Description: Fault Domains Test	
		Back Next Finish Cancel

Figure A.7 Give a name and description to the policy

Click through the rule-set description.

Create New VM Storage Policy					(?))
 1 Name and description 2 Rule-Sets 2a Rule-Set 1 3 Storage compatibility 4 Ready to complete 	Rule-Sets AVM storage policy consists of m sets describing the storage reso The VM storage policy will match Gold Storage Policy	policy can include multiple rule-			
	Rule-set 1 Datastore type 1	Rule-set 2 Datastore type 2	or	Rule-set 3 Gold tags	
				Back	Next Finish Cancel

Figure A.8 Rule-set description

In the Rule-Set 1 window, select Virtual SAN as the "Rules based on data service". Then add the rule "Number of disk stripes per object" and set the value to 3. There is no need to add "Number of failures to tolerate" as this is automatically set to 1 for every policy unless you explicitly set it to a value of 0.

Create New VM Storage Policy		(?)»
 1 Name and description 2 Rule-Sets 	Rule-Set 1 Select rules specific for a datastore type. Rules can be based on data services provided by data: The VM storage policy will match datastores that satisfy all the rules in at least one of the rule-se	store or based on tags. ets.
2a Rule-Set 1 3 Storage compatibility 4 Ready to complete	Rules based on data services VSAN Number of disk stripes per object 3 <add rule=""> Rules based on tags Add tag-based rule Add another</add>	Storage Consumption Model A virtual disk with size 100 GB would consume: Storage space 200.00 GB Initially reserved storage space 0.00 B Reserved flash space 0.00 B
	Back	Next Finish Cancel

Figure A.9 Number of disk stripes per object

The Virtual SAN datastore should appear as compatible, in other words it understands the policy settings.

🚰 Create New VM Storage Policy											? H
1 Name and description 2 Rule-Sets	Storage compatibility As defined, this VM storage policy is compatible with the following storage:										
✓ 2a Rule-Set 1											
3 Storage compatibility	Storage Compatibility	Total Capa	city	Virtual SAN	Capacity	Vir	rtual Volumes Cap	VMFS Capacit	ty	NFS Capacity	
 4 Ready to complete 	Compatible	1.06 TB		1.06 TB		0.	.00 B	0.00 B		0.00 B	
	Incompatible	52.50 TB		0.00 B		0.	.00 B	273.00 GB		52.23 TB	
	Compatible storage										
	Q Filter					r	•				
	Name		Datacenter		Туре		Free Space	Capacity	Warnings		
	🗐 vsanDatastore		I VSAN	16-DC	vsan		901.48 GB	1.06 TB			
							Bac	k Next	F	inish C	Cancel

Figure A.10: vsanDatastore shows as compatible

The final step is to click on Finish and create the policy.

😤 Create New VM Storage Policy			? »
 1 Name and description 2 Rule-Sets 	Ready to complete Confirm that the inform	nation below is correct and click Finish to create the VM storage policy.	
✓ 2a Rule-Set 1	General		
✓ 3 Storage compatibility	Name	StripeWidth=3	
4 Ready to complete	Description	Fault Domains Test	
	vCenter Server	ie-vcsa-09.ie.local	
	Rule-Set 1: VSAN		
	Number of disk object	stripes per 3	
		Back Next	Finish Cancel

Figure A.11 Finish creating the policy

We can now go ahead and deploy a VM with this policy, and afterwards we shall examine the layout and see if it is taking Fault Domains into account.

A3. Create a VM and Check the Fault Domains

At this point, a new VM can be deployed. The only inputs required for this VM are to provide it with a name and to choose the newly created policy with a StripeWidth = 3.



Figure A.12 Create a new VM

1 Select creation type	Select storage Select the datastore in which to sto	ore the configuratio	n and disk files			
1a Select a creation type 2 Edit settings 2a Select a name and folder 2b Select a compute resource	VM Storage Policy. StripeWidth= The following datastores are acce virtual machine configuration files	3 essible from the de and all of the virtua	stination resource that	t you selected. Select	the destination d	atastore for the
2c Select storage	Name	Capacity	Provisioned	Free	Туре	Storage DRS
2d Select compatibility	Compatible					
2e Selecta guest OS	🗐 vsanDatastore	1.06 TB	361.48 GB	901.48 GB	vsan	
2f Customize hardware	Incompatible					
3 Ready to complete	NFS-Isilon	50.46 TB	379.50 GB	50.19 TB	NFS v3	
o noug to complete	corkisos.csl.vmware.com	1.77 TB	1.62 TB	160.36 GB	NFS v3	
	cs-ie-h04-scratch	136.50 GB	3.23 GB	133.27 GB	VMFS	
	scie-h03-scratch	136.50 GB	3.33 GB	133.17 GB	VMFS	
	4	::				
	Compatibility:					
	Compatibility checks succee	eded.				

Figure A.13 Select the new VM Storage Policy

The rest of the VM creation options can be left at the default. Once the virtual machine has been deployed, check the Manage tab > Policies and verify that the VM is compliant with the policy. It should be compliant as shown below.

_					
6	FD-SW3-TEST-VM Actions -				±.
(Getting Started Summary Monitor Manage Relat	ted Objects			
0	Settings Aarm Definitions Tags Permissions Polic	Scheduled Tasks			
	Storage				Edit VM Storage Policies
	C 🍖 👼				Q Filter 🔹
	Name	VM Storage Policy	Compliance Status	Last Checked	
	🗂 VM home	StripeWidth=3	✓ Compliant	16/04/2015 09:19	
	🚈 Hard disk 1	StripeWidth=3	✓ Compliant	16/04/2015 09:19	

Figure A.14 VM Storage Policy is Compliant

Finally check the distribution of VM components under the Monitor tab > Policies.

	-					
ng Started Summary Mc	onitor Manage Related C	bjects				
use Porformance Policies	Taske Events Utilization	1				
les renormance roncies	Tasks Events Oulization	J				
😜 👘						Q Filter
ne	VM	Storage Policy	Complia	nce Status	Last Checked	
VM home	5	StripeWidth=3	🗸 Co	mpliant	16/04/2015 09:19	
Hard disk 1	6	StripeWidth=3	✓ C0	npliant	16/04/2015 09:19	
						2 items
compliance Failures Physic	cal Disk Placement					
SW3-TEST-VM - Hard disk 1	cal Disk Placement 1 : Physical Disk Placement					
>SW3-TEST-VM - Hard disk 1	cal Disk Placement 1 : Physical Disk Placement					Q Filter
Dompliance Failures Physic Dompliance Failures Physic Provide the second	cal Disk Placement 1 : Physical Disk Placement Component State	Host	Flash Disk Name	Flash Disk Uuid	HDD Disk Name	Q Filter HDD Disk Uuid
20mpliance Failures Physie Physie PSW3-TEST-VM - Hard disk 1 T T T T T T T T T T T T T T T T T T T	1 : Physical Disk Placement	Host	Flash Disk Name	Flash Disk Uuid	HDD Disk Name	Q Filter HDD Disk Uuid
Pompliance Failures Physie Ph	1 : Physical Disk Placement Component State	Host	Flash Disk Name	Flash Disk Uuid	HDD Disk Name	Q Filter HDD Disk Uuid
ompliance Failures Physi > SW3-TEST-VM - Hard disk 1 ▼ ▼ PP r RAID 1 ▼ RAID 0 Component	Component State Active	Host	Flash Disk Name	Flash Disk Uuid 52fff7ee-ecd4-2c7d-47a4-5a125	HDD Disk Name	Q Filter HDD Disk Uuid 523d0aeb-4863-cb62-1a
ompliance Failures Physis >SW3-TEST-VM - Hard disk \ ▼ = ↓ RAID 1 ↓ RAID 0 Component Component	Component State Active Active	Host	Flash Disk Name	Flash Dia Uuid 521177ee-ec04-2c70-47a4-5a125 524a1c33-5570-caa0-01277d5	HDD Disk Name	Q Filter HDD Disk Uuid 523d0aeb-4863-cb62-1a 522e849c-a461-4141-116
Compliance Failures Physis OSW3-TEST-VM - Hard disk :	Component State Component State Active Active Active Active	Hoat Cs-ie-h03.ie.l Cs-ie-h04.ie.l Cs-ie-h04.ie.l	Flash Disk Name HP Serial Attached SCSI Dis	Flash Diak Uuid 52ff7ee-ecd4-2c7d-47a4-5a125 524a1c38-5b7d-caad-0127-fd5 524a1c38-5b7d-caad-0127-fd5	HCD Disk Name HCD Disk Name HC Serial Attached SCSI Dis HP Serial Attached SCSI Dis HC Serial Attached SCSI Dis	Q. Filter HDD Disk Uuld 523d0aeb-4863-cb62-1a 5228849c-a461-4441-116 52a9719d-d754-289e-e1
vompliance Failures Physis 0-SW3-TEST-VM - Hard disk ' vpe RAID 1 RAID 0 Component Component Component RAID 0	Component State	Host Cs-ie-h03.ie.l Cs-ie-h04.ie.l Cs-ie-h04.ie.l	Flash Disk Name HP Serial Attached SCSI Dis. HP Serial Attached SCSI Dis. HP Serial Attached SCSI Dis.	Flash Diat Uuid 52077ee-ecd4-2c7d-47a4-5af25 524a1c38-5b7d-caad-0127-fd5 524a1c38-5b7d-caad-0127-fd5	HDD Disk Name HD Serial Attached SCSI Dis HP Serial Attached SCSI Dis HP Serial Attached SCSI Dis	Q. Filter HOD Diak Usia 523d0aeb-4803-cb62-1a 522e649c-a461-441-1f6 52a9719d-d754-289e-e1
compliance Failures Physis DSW3-TEST-VM - Hard disk ' * * * RAID 1 * RAID 0 Component Component * RAID 0 Component Component Component	Cal Disk Placement	Host Cs-ie-h03.ie.l Cs-ie-h04.ie.l Cs-ie-h04.ie.l Cs-ie-h01.ie.l	Flash Disk Name HP Serial Attached SCSI Dis	Flash Disk Uuid 52fff7ee-ecd4-2c7d-47a4-5af25 524a1c38-5b7d-caad-0127-fd5 524a1c38-5b7d-caad-0127-fd5 5269d5e7-8630-9951-9962-dd	HDD Disk Name HP Serial Attached SCSI Dis	Q. Filter HDD Diak Uuld 523d0aeb-4863-cb62-1a 522e849c-a461-441-1t6 52a9719d-d754-289e-e1 526581bb-1118-6e56-cel
ompliance Failures Physis DSW3-TEST-VM - Hard disk ' ▼ ■ ▼ RAID 1 ♥ RAID 0 Component Component Component Component Component	Cal Disk Placement	Host CS-Ie-h03.Ie.I CS-Ie-h04.Ie.I CS-Ie-h04.Ie.I CS-Ie-h01.Ie.I CS-Ie-h01.Ie.I CS-Ie-h02.Ie.I	Flash Diak Name HP Serial Attached SCSI Dis	Fleah Dia Uuid 521177 ee-ecd4-2c7d-47a4-5a125 524a1 c38-567 d-caa d-0127-fd5 524a1 c38-567 d-caa d-0127-fd5 5269 d667-8630-9951-9962-dd 5265 138e-8650-3947-ed6f-4df6	HDD Disk Name HD Serial Attached SCSI Dis HP Serial Attached SCSI Dis	Q. Filter NOD Die Uvis 523d0aeb-4863-cb62-1a 522e849c-a461-4441-1f6i 52a9719d-0754-289e-e1 52f881bb-1118-666-ce0 52c28847-ae01-7afe-59e

rigure A.15 Component distribution

The questions that need to be asked now are related to rack failures and fault domain failures. For example, if rack 1 were to fail, is there still a full copy of the data? The answer is yes. What about rack 2? Yes, there is still a full copy of the data. What about rack 3, which houses hosts 3 and 4? The answer is yes, once again there would be a fully copy of the data even if rack 3 failed.

One additional item to highlight here is the lack of witnesses. This is something new in Virtual SAN 6.0. Certain configurations do not need witnesses as a new voting mechanism has been introduced which gives components extra votes. Therefore in some configurations, such as this one, witnesses are not needed, reducing the overall component count.

Appendix B—Migrating from Standard vSwitch to Distributed

Before we begin, this procedure is rather complicated, and can easily go wrong. The only real reason why one would want to migrate from VSS (standard vSwitches) to a DVS (Distributed vSwitch) is to make use of the Network I/O Control feature that is only available with DVS. This will then allow you to place QoS (Quality of Service) on the various traffic types such as Virtual SAN traffic.

Warning: Ensure that you have console access to the ESXi hosts during this exercise. All going well, you will not need it. However, should something go wrong, you may well need to access the console of the ESXi hosts.

B.1 Create Distributed Switch

To begin with, create the distributed switch. This is a relatively straight forward exercise.



Figure B.1 Create a new distributed switch

Provide it with a name.

Level New Distributed Switch			(?) }>
1 Name and location	Name:	VSAN-DSwitch	
2 Select version	Location:	VSAN6-DC	
3 Edit settings			
4 Ready to complete			
		Park Next Finish Co	ncol
		Back Next Prinsi Ca	

Figure B.2: Provide a name for the new distributed switch

Select the version of the DVS. In this example, we shall use the latest version, 6.0.0.



Figure B.3: Select the distributed switch version

At this point, we get to add the settings. First, you will need to determine how many uplinks you are currently using for networking. In our POC, we are using six; one for management, one for vMotion, one for virtual machines and three for Virtual SAN. Therefore, when we are prompted for the number of uplinks, we select "6". This may differ in your environment but you can always edit it later on.

Level 2 New Distributed Switch		())
 1 Name and location 2 Select version 3 Edit settings 4 Ready to complete 	Edit settings Specify number of uplinks: Number of uplinks: Network I/O Control: Default port group: Port group name:	k ports, resource allocation and default port group.
		Back Next Finish Cancel

Figure B.4: Select the number of uplinks

Another point to note here is that a default portgroup can be created. You can certainly create a port group at this point, but there will be additional port groups that need to be created shortly. At this point, the distributed switch can be completed.

Les New Distributed Switch				(4 €
1 Name and location2 Select version	Ready to complete Review your settings selection	s before finishing the wizard.		
 3 Edit settings 4 Ready to complete 	Name: Version: Number of uplinks: Network I/O Control: Default port group: Suggested next actions Mew Distributed Port C Add and Manage Host	VSAN-DSwitch 6.0.0 6 Enabled MGMT-DPG Sroup is available in the Actions menu of the new d	listributed switch.	
			Back Next Finish	Cancel

Figure B.5: Complete the creation of the DVS

As alluded to earlier, configure and create the additional port groups.

B.2 Create Port Groups

In the previous exercise, a single default port group was created for the management network. There was little in the way of configuration that could be done at that time. It is now important to edit this port group to make sure it has all the characteristics of the management port group on the VSS, such as VLAN and NIC teaming and failover settings. Select the distributed port group, and click on the Edit button shown below.

vmware vSphere Web Client _ ♠ =				0	Administrator@VSPHERE.LOCAL •	l Help - I Q Search
Navigator I	Actions -					E.
Hosts and Clusters	Getting Started Summary Mr	nitor Manage Related Object	ts			
Image: Windows Image:	Settings Narm Definitions Tag	ps Permissions Network Proto	col Profile Ports			Edit
VM Network	Properties					Luit
± urant-anc > ≣VS44+05wtdr-0VUplants-277	Policies	Security Promiscuous mode: NAC address changes: Forged transmits: Ingress traffic shaping Status: Average bandwidth: Burst size Egress traffic shaping Status: Average bandwidth: Pest bandwidth: Pest bandwidth: Pest bandwidth: Pest bandwidth: Pest bandwidth: Type:	Reject Reject Disabled Disabled Disabled Disabled - None			

Figure B.6: Edit the distributed port group

For some port groups it may be necessary to change the VLAN. Since the management VLAN in this POC is on 51, we need to tag the distributed port group accordingly.

🍰 MGMT-DPG - Edit Settings			?
MMMT-DPG - Edit Settings General Advanced Security Traffic shaping VLAN Teaming and failover Monitoring Traffic filtering and marking Miscellaneous	VLAN type: VLAN ID:	MAN ▼ 51 ▼	•
			OK Cancel

Figure B.7: Tag the distributed port group with a VLAN

That is the management distributed port group taken care of. You will also need to create distributed port groups for vMotion, virtual machine networking and of course Virtual SAN networking. In the "Getting Started" tab of the distributed switch, there is a basic task link called "Create a new port group".



Figure B.8: Create a new distributed port group

In this exercise, we shall create a port group for the vMotion network.

2 New Distributed Port Group			?	**
 Select name and location Configure settings 	Name: Location:	Vilotion-DPG VSA4-DSwitch		
3 Ready to complete	20020001.			
		Bark Next Finish C	ancel	
		Back Next Finish C	ancel	

Figure B.9: Provide a name for the new distributed port group

A New Distributed Port Group					(?)
1 Select name and location 2 Configure settings	Configure settings Set general properties of the	e new port group.			
3 Ready to complete	Port binding: Port allocation:	Static binding Elastic Blastic port groups		ase or decrease the number of ports as needed	L.
	Number of ports: Network resource pool:	8 (default)	•		
	VLAN VLAN type:	VLAN	•		
	VLAN ID: Advanced	51	*		
	Customize default polic	cies configuration			
				Back Next Finish	Cancel

Figure B.10: Configure distributed port group settings, such as VLAN

😫 New Distributed Port Group						(€
1 Select name and location 2 Configure settings	Ready to complete Review the changes before proceed	ding.				
3 Ready to complete	Distributed port group name: Port binding: Number of ports: Port allocation: Network resource pool: VLAN ID:	VMotion-DPG Static binding 8 Elastic (default) 51				
			Back	Next	Finish	Cancel

Figure B.11: Finish creating the new distributed port group

Once all the distributed port groups are created on the distributed switch, the uplinks, VMkernel networking and virtual machine networking can be migrated to the distributed switch and associated distributed port groups.

Warning: While the migration wizard allows many uplinks and many networks to be migrated concurrently, we recommend migrating the uplinks and networks step-by-step to proceed smoothly and with caution. For that reason, this is the approach we use here.

B.3 Migrate Management Network

To begin, let's migrate just the management network (vmk0) and its associated uplink, which in this case is vmnic0 from VSS to DVS. To begin, select "Add and manage hosts" from the basic tasks in the Getting started tab of the DVS.



Figure B.12: Add and manage hosts

The first step is to add hosts to the DVS.



Figure B.13: Add hosts

Click on the green + and add all four hosts from the cluster.

1	Add and Manage Hosts		(?)
~	1 Select task	Select hosts Select hosts to add to this distributed switch.	
	3 Select network adapter tasks	💠 New hosts 🛛 🗶 Remove	
	4 Manage physical network	Host	Host Status
	- Manage VMkernel network	1 (New) cs-ie-h01.ie.local	Connected
	5 adapters	(New) cs-ie-h02.ie.local	Connected
	6 Analyze impact	(New) cs-ie-h03.ie.local	Connected
	7 Ready to complete	(New) cs-ie-h04.ie.local	Connected
		Configure identical network settings on multiple hosts (template	e mode). 🚯
			Back Next Finish Cancel

Figure B.14: Select all hosts in the cluster

The next step is to manage both the physical adapters and VMkernel adapters. To repeat, what we wish to do here is migrate both vmnic0 and vmk0 to the DVS.



Figure B.15: Select physical adapters and VMkernel adapters

Next, select an appropriate uplink on the DVS for physical adapter vmnic0. In this example we chose Uplink1.

🕼 Add and Manage Hosts					?
 ✓ 1 Select task ✓ 2 Select hosts 	Manage physical network adapters Add Select an Uplink for vmnic0		۲		
✓ 3 Select network adapter tasks	Uplink	Assigned Adapter			
4 Manage physical network	Host Uplink 1			Uplink Port Group	
Manage VMkernel network	Uplink 2				-
5 adapters	v 🕈 Uplink 3				
6 Analyze impact	Uplink 4				
7 Ready to complete	Uplink 5	-		VSAN-DSwitch-DVU	pli
	Uplink 6	-			
	(Auto-assign)			-	
	- 4				
		OK	Cancel		
	vmnic0	vSwitch0			
	📻 vmnic1	vSwitch0			
	pm vmnic2	vSwitch0			
	vmnic3	vSwitch1		-	
	vmnic/	vSwitch1			*
			Back	Next Finish	Cancel

Figure B.16: Assign uplink (uplink1) to physical adapter vmnic0

With the physical adapter selected and an uplink chosen, the next step is to migrate the management network on vmk0 from the VSS to the VDS. We are going to leave vmk1 and vmk2 for the moment and just migrate vmk0.

Select vmk0, and then click on the "Assign port group" as shown below. The port group assigned should be the newly created distributed port group created for the management network earlier. Remember to do this for each host.

🕼 Add and Manage Hosts					?	
 1 Select task 2 Select hosts 	Manage VMkernel network adapters Manage and assign VMkernel network	s k adapters to the distributed switcl	h.			
 Select network adapter tasks 4 Manage physical network adapters 	Vulkernel network adapters with the warning sign might lose network connectivity unless they are migrated to the distributed switch. Select a destination port group to migrate them.					
5 Manage VMkernel network adapters	Host/VMkernel Network Adapters	1 A In Use by Switch	Source Port Group	Destination Port Group		
6 Analyze impact					-	
7 Ready to complete	On this switch					
	vmk0	vSwitch0	VMkernel-mgmt	Do not migrate		
	vmk1	vSwitch0	VMkernel-vmotion	Do not migrate		
	ymk2	vSwitch 1	VSAN	Do not migrate	::	
	On this switch					
	 On other switches 					
	🗾 vm k0	vSwitch0	VMkernel-mgmt	Do not migrate		
	m vmk1	vSwitch0	VMkernel-vmotion	Do not migrate		
	ymk2	vSwitch1	VSAN	Do not migrate		
	👻 👕 cs-ie-h03.ie.local					
	On this switch					
	 On other switches 					
	wmk0	vSwitch0	Management Network	Do not migrate	-	
			Back Nex	t Finish C	Cancel	

Figure B.17: Assign port group for vmk0

Click through the analyze impact screen since it only checks iSCSI and is not relevant to the Virtual SAN POC.



Figure B.18: Impact on iSCSI (not relevant)

At the finish screen, you can examine the changes. We are adding 4 hosts, 4 uplink s (vmnic0 from each host) and 4 VMkernel adapters (vmk0 from each host).

🕼 Add and Manage Hosts		?
✓ 1 Select task	Ready to complete Review your settings selections before finishing the wizard.	
 2 Select hosts 		
✓ 3 Select network adapter tasks	Number of managed hosts	
 4 Manage physical network adapters 	Hosts to add: 4	
Manage VMkernel network	Number of network adapters for update	
^b adapters	Physical network adapters: 4	
 6 Analyze impact 7 Ready to complete 	Reassigned VMkernel network 4 adapters:	
	Back Next Finish	Cancel

Figure B.19: Ready to complete

When the networking configuration of each host is now examined, you should observe the new DVS, with one uplink (vmnic0) and the vmk0 management port on each host.



Figure B.20: Management network migration to DVS complete

You will now need to repeat this for the other networks.

B.4 Migrate vMotion

Migrating the vMotion network takes the exact same steps as the management network. Before you begin, ensure that the distributed port group for the vMotion network has all the same attributes as the port group on the standard (VSS) switch. Then it is just a matter of migrating the uplink used for vMotion (in this case vmnic1) along with the VMkernel adapter (vmk1). As mentioned already, this takes the same steps as the management network.

When the migration completes, the individual host network configuration should look similar to the following diagram.

cs-ie-h01.ie.local Actions	,		=*
Getting Started Summary Mo	nitor Manage Related Objects		
Settings Networking Storage	Aarm Definitions Tags Permissions		
44	Virtual switches		
Virtual switches	92 92 💷 🐺 🙀 🗙 🚯		
VMkernel adapters	Switch	Discovered Issues	
Physical adapters	C VSAN-DSwitch		
TCP/IP configuration	1 vSwitch0	**	
Advanced	1 vSwitch1	-	
	Distributed switch: VSAN-DSwitch (MGMT	-DPG)	
	<u>9</u>	(assigned p	iort groups) 🛛 🔻 🧲
	Assigned port groups filter applied, show	ving: 🚨 2/4	
		A MGMT-DPG	
		VLAN ID: 51	
		VMkernel Ports (1)	
		Virtual Machines (0)	
		La VMotion-DPG	
		VLAN ID: 51	
		Vintual Machines (0)	

Figure B.21: vMotion network migration to DVS complete

B.5 Migrate Virtual SAN Network

If you are using a single uplink for the Virtual SAN network, then the process becomes the same as before.

However, if you are using more than one uplink, then there are additional steps to be taken. If the Virtual SAN network is using a feature such as Link Aggregation (LACP), or it is on a different VLAN to the other VMkernel networks, then you will need to place some of the uplinks into an unused state for certain VMkernel adapters.

For example, in this scenario, VMkernel adapter vmk2 is used for Virtual SAN. However uplinks vmnic3, 4 and 5 are used for Virtual SAN and they are in a LACP configuration. Therefore for vmk2, all other vmnics (0, 1 and 2) must be placed in an unused state. Similarly, for the management adapter (vmk0) and vMotion adapter (vmk0), the Virtual SAN uplinks/vmnics should be placed in an unused state.

Modifying the settings of the distributed port group and changing the path policy/failover appropriately do this.

In the manage physical network adapter, the steps are similar as before except that now you are doing this for multiple adapters.

🗊 Add and Manage Hosts				(
 1 Select task 2 Select hosts 	Manage physical network adapters Add or remove physical network adapters to	this distributed switch.		
✓ 3 Select network adapter tasks	📾 Assign uplink 🗙 Unassign adapter 👔	👩 Reset changes 👔 View	settings	
Manage physical network adapters Manage VMkernel network adapters	Host/Physical Network Adapters	1 🛦 In Use by Switch	Uplink	Uplink Port Group
6 Analyze impact 7 Ready to complete	vmnic0	VSAN-DSwitch VSAN-DSwitch	Uplink 1 Uplink 2	VSAN-DSwitch-DVUpli VSAN-DSwitch-DVUpli
	vmnic3 (Assigned)	vSwitch1 vSwitch1	Uplink 4 Uplink 5	VSAN-DSwitch-DVUpli
	 On other switches/unclaimed mnic1 	vSwitch0	-	-
	 cs-ie-h02.ie.local On this switch wmnic0 	VSAN-DSwitch	Uplink 1	VSAN-DSwitch-DVUplin
	vmnic2	VSAN-DSwitch vSwitch1	Uplink 2 Uplink 4	VSAN-DSwitch-DVUpli VSAN-DSwitch-DVUpli
	vmnic4 (Assigned) vmnic5 (Assigned) • On other switches/unclaimed	vSwitch1 vSwitch1	Uplink 6	VSAN-DSwitch-DVUplin VSAN-DSwitch-DVUplin
	vmpic1	vSwitch0	- Rack	Vart Sinish Canael

Figure B.22: Multiple uplinks used by the Virtual SAN network

As before, vmk2 (the Virtual SAN VMkernel adapter) should be assigned to the distributed port group for Virtual SAN.

D	Add and Manage Hosts					?
`	1 Select task 2 Select hosts	Manage VMkernel network adapters Manage and assign Wikernel networ	s k adapters to the distributed switch	ı.		
č	A Manage physical network adapter adapters Manage VMkernel network	VMkernel network adapters with switch. Select a destination port Assign port group + New adapters	the warning sign might lose netwo group to migrate them.	rk connectivity unless they a	re migrated to the distributed	
	^o adapters	Host/VMkernel Network Adapters	1 A In Use by Switch	Source Port Group	Destination Port Group	
	6 Analyze impact	🛤 vmk1	VSAN-DSwitch	VMotion-DPG	Do not migrate	^
	7 Ready to complete	(mk2 (Reassigned)	vSwitch1	VSAN	VSAN-DPG	
		On other switches				
		→ On this switch				
		📖 vm k0	VSAN-DSwitch	MGMT-DPG	Do not migrate	
		📖 vmk1	VSAN-DSwitch	VMotion-DPG	Do not migrate	
		vmk2 (Reassigned)	vSwitch1	VSAN	VSAN-DPG	
		On other switches				
		 On this switch 				::
		📖 vmk0	VSAN-DSwitch	MGMT-DPG	Do not migrate	
		📷 vm k1	VSAN-DSwitch	VMotion-DPG	Do not migrate	
		(Image: with the set of the set	vSwitch1	VSAN	VSAN-DPG	
		On other switches				*
				Back	Next Finish Ca	ncel

Figure B.23: Assign distributed port group for Virtual SAN networking

Note: If you are only now migrating the uplinks for the Virtual SAN network, you may not be able to change the distributed port group settings until after the migration. During this time, Virtual SAN may have communication issues. After the migration, move to the distributed port group settings and make any policy changes and mark any uplinks that should be unused. Virtual SAN networking should then return to normal when this task is completed. Use the Health Check plugin to verify that everything is functional once the migration is completed.



Figure B.24: Change distributed port group settings

🍰 VSAN-DPG - Edit Settings				?
General	Load balancing:	Route based on IP hash	• 0	
Advanced	Network failure detection:	Link status only	•	
Security	Notify switches:	Yes	•	
Traffic shaping	Failback:	Ves	•	
VLAN	- difedole	105		
Teaming and failover	Failover order			
Monitoring	+ +			
Traffic filtering and marking	Active uplinks			
Miscellaneous	对 Uplink 4			
	Uplink 5			
	🜉 Uplink 6			
	Standby uplinks			
	Unused uplinks			
	pe Uplink 1			
	Uplink 2			
	🗾 Uplink 3			
	Select active and standby up	links. During a failover, standby uplinks activ	ate in the order specified above.	
			OK Cance	<u>اللہ</u>

Figure B.25: Showing load balancing and unused uplinks

That completed the VMkernel adapter migrations. The final step is to move the VM networking.

B.6 Migrate VM Network

This is the final step of migrating the network from a standard vSwitch (VSS) to a distributed switch (DVS). Once again, we use the "Add and manage hosts", the same link used for migrating the VMkernel adapters. The task is to manage host networking.

🕞 Add and Manage Hosts	0
1 Select task	Select task Select a task to perform on this distributed switch.
 2 Select hosts 3 Select network adapter tasks 4 Manage physical network adapters 5 Manage VMkernel network adapters 6 Analyze impact 7 Ready to complete 	 Add hosts Add new hosts to this distributed switch. Manage host networking Manage networking Manage networking distributed switch. Remove hosts Remove hosts Remove hosts from this distributed switch. Add host and manage host networking (advanced) Add new hosts and manage networking of hosts already attached to this distributed switch. Use this option to unify the network configuration of new and existing hosts.
	Back Next Finish Cancel

Figure B.26: Manage host networking

Select all the hosts in the cluster, as all hosts will have their virtual machine networking migrated to the distributed switch.

12 4	🔂 Add and Manage Hosts			
✓ 1	Select task	Select hosts Select hosts to manage their networking on this distribut	ed switch	
	Select hosts			
3	Select network adapter tasks	+ Attached hosts 🗶 Remove		
4	4 Manage physical network adapters 5 Manage VMkernel network adapters	Host	Host Status	
		cs-ie-h01.ie.local	Connected	
5		cs-ie-h02.ie.local	Connected	
6	Analyze impact	cs-ie-h03.ie.local	Connected	
7	Ready to complete	cs-ie-h04.ie.local	Connected	
		Configure identical network settings on multiple hosts	s (template mode). 🕕	
			Back	Finish Cancel

Figure B.27: Select all hosts

On this occasion, we do not need to move any uplinks. However, if the VM networking on your hosts used a different uplink, then this of course would also need to be migrated from the VSS. In this example, the uplink has already been migrated.

🕼 Add and Manage Hosts	0
 1 Select task 2 Select hosts 	Select network adapter tasks Select the network adapter tasks to perform.
3 Select network adapter tasks 4 Migrate VM networking	Manage physical adapters Add physical network adapters to the distributed switch, assign them to uplinks, or remove existing ones.
5 Ready to complete	Manage VMkernel adapters Add or migrate VMkernel network adapters to this distributed switch, assign them to distributed port groups, configure VMkernel adapter settings, or remove existing ones.
	Migrate virtual machine networking Migrate VM network adapters by assigning them to distributed port groups on the distributed switch.
	Manage advanced host settings Set the number of ports per legacy host proxy switch.
	Migrate virtual machines
	Back Next Finish Cancel

Figure B.28: Migrate virtual machine networking

Select the VMs that you wish to have migrated from a virtual machine network on the VSS to the new virtual machine distributed port group on the DVS. Click on the "Assign port group" option like we have done many times before, and select the distributed port group, name VM-DPG here.

🔂 Add and Manage Hosts				
1 Select task 2 Select hosts 3 Select network adapter tasks 4 Migrate VM networking 5 Ready to complete	Migrate VM networking Select virtual machines or network adapters to migrate to the distributed switch. Assign VMs or network adapters to a destination port group to migrate them. Press and hold down the CTRL key, and then click the VMs to select multiple items. Some network adapters are not connected to a valid port group. Migrate these adapters to prevent networking issues. Assign port group Reset changes VM we settings			
	HottVirtual Machine/Network Ada 1	NRC Count 1 1	Source Port Group VM Network VM Network	Destination Port Group VM-DPG VM-DPG Do not migrate
			Back	Next Finish Cancel

Figure B.29: Assign port groups for the VMs

Reviewing the final screen. In this case we are only moving to VMs. Note that any templates using the original VSS virtual machine network will need to be converted

to virtual machines, edited and the new distributed port group for virtual machines will need to be selected as the network. This step cannot be achieved through the migration wizard.

🕼 Add and Manage Hosts		?
 Add and manage nosis 1 Select task 2 Select hosts 3 Select network adapter tasks 4 Migrate VM networking 5 Ready to complete 	Ready to complete Review your settings selections before finishing the wizard. Number of managed hosts Hosts to update: 4 Number of network adapters for update Virtual machine adapters: 2	
	Back Next Finish	Cancel

Figure B.30: Finish

The VSS should no longer have any uplinks of port groups and can be safely removed.

Getting Started Summary Monitor Manage Related Objects						
Settings Networking Storage Alarm Definitions Tags Permissions						
	Virtual switches					
Virtual switches	😟 😡 🛤 🗁 🦯 🗙 🚯					
viikernei adapters	Switch	Discovered Issues				
Physical adapters	VSAN-DSwitch	-				
TCP/IP configuration	1 vSwitch0					
Advanced	1 vSwitch1	-				
		=				
	Standard switch: vSwitch1 (no item selected)					
		G				
		No associated port gro				

Figure B.31: VSS no longer in use

This completes the migration from a standard vSwitch (VSS) to a distributed vSwitch (DVS).