VMware Virtual SAN

Quick Monitoring & Troubleshooting Reference Guide

TECHNICAL MARKETING DOCUMENTATION
March 2014

vmware

Introduction

The VMware Virtual SAN Quick Monitoring and Troubleshooting Reference Guide provide troubleshooting guidance for issues encountered during the configuration and monitoring of VMware Virtual SAN. This document identifies the most common issues that can be encounter and how to quickly resolve them.

Tools

Aside from the vSphere Web Client, the vCenter Server and vSphere Hypervisor provide a set of monitoring and troubleshooting tools that are bundled with the software and can be used to monitor and troubleshoot VMware Virtual SAN (VSAN).

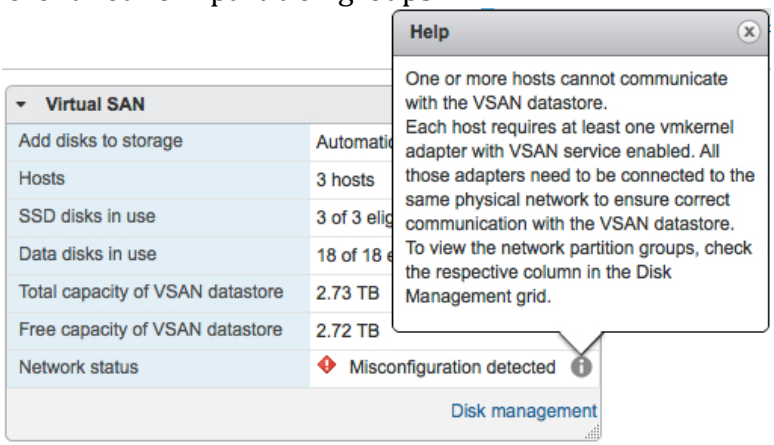
- vSphere Web Client
- ESXCLI
- Ruby vSphere Console (RVC)
- VSAN Observer

Details on these tools can be found in the Virtual SAN documentation.

Known Issues and Solutions

Network Status reports "Misconfiguration Detected"

Issue: The Virtual SAN cluster fails to form correctly due to one or more members of the cluster not being able to communicate over the Virtual SAN network as a result of being in different network partition groups.



Solution: Ensure the physical switch and the ports used for Virtual SAN are active and have multicast enabled. Enabling multicast can be done in one of two way on your physical switches:

- Disabling IGMP snooping.
- Configure IGMP snooping for selective traffic

Also, validate the virtual switch configuration for correct uplink, VLAN, NIC team failover policy configuration and the Virtual SAN traffic service is enabled on the VMkernel interfaces.

Virtual SAN requires a VMkernel network interface with the Virtual SAN traffic enabled. All members of the clusters must communicate on the same L2 network segment with multicast enabled, and all members of the cluster should be able to ping each other. Failing to meet this requirement will prevent Virtual SAN from being successfully configured, as it will prevent the hosts from communicating.

Use the tools and examples below to monitor and troubleshoot this type of issue.

VMKPING
Use the "vmkping" command to validate network accessibility for the Virtual SAN network. All of the host in the cluster should be able to successfully ping each others VSAN network interface.

```
~ # vmkping 10.20.196.84
PING 10.20.196.84 (10.20.196.84): 56 data bytes
64 bytes from 10.20.196.84: icmp_seq=0 ttl=64 time=0.057 ms
64 bytes from 10.20.196.84: icmp_seq=1 ttl=64 time=0.081 ms
64 bytes from 10.20.196.84: icmp_seq=2 ttl=64 time=0.071 ms

--- 10.20.196.84 ping statistics ---
3 packets transmitted, 3 packets received, 0% packet loss
round-trip min/avg/max = 0.057/0.070/0.081 ms
~ #
```

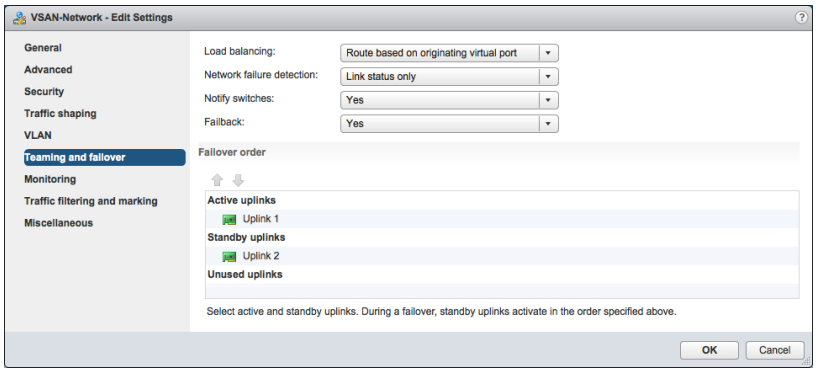
ESXCLI

Use the "esxcli vsan network" namespace to examine and modify VSAN network configurations. Monitor and validate hosts multicast configuration details.

```
~ # esxcli vsan network list
Interface
   VmkNic Name: vmk1
   IP Protocol: IPv4
   Interface UUID: fe63bc52-846c-6815-2a1c-ac162d7444e8
   Agent Group Multicast Address: 224.2.3.4
   Agent Group Multicast Port: 23451
   Master Group Multicast Address: 224.1.2.3
   Master Group Multicast Port: 12345
   Multicast TTL: 5
~ #
```

vSphere Web Client

The use of multiple network adapters for the VSAN network is recommended from an availability perspective, the team failover policy adapter configuration should be in active/standby configuration in order to avoid possible VSAN network partition issues.



Automatic "Add Disk to Storage Mode"

Issue: Automatic disk claiming operation fails to claim disks.

Solution: The disk groups will have to be created manually. Disks are automatically claimed when ESXi flags them as "local". Many SAS controllers allow disks to be

shared. If ESXi determines that the disks are shared, it does not report them as "local". Shared disks are currently not supported in Virtual SAN 1.0 and are reported as not local or "Is local: false". However, some disks are reported as shared but are actually not shared, in that case you will need to mark them "as local". This applies to both magnetic disks (HDD) as well as solid state disks (SSD).

ESXCLI

Use the "esxcli storage core" namespace to examine whether the disks are flagged as local or not, by looking at the "Is Local" attribute. This can be done before or after enabling Virtual SAN.

```
~ # esxcli storage core device list
naa.5000c500643e03a3
   Display Name: ATA Serial Attached SCSI Disk (naa.5000c500643e03a3)
   Has Settable Display Name: true
   Size: 953869
   Device Type: Direct-Access
   Multipath Plugin: NMP
   Devfs Path: /vmfs/devices/disks/naa.5000c500643e03a3
   Vendor: ATA
   Model: ST91000640NS
   Revision: SN03
   SCSI Level: 6
   Is Pseudo: false
   Status: degraded
   Is RDM Capable: true
   Is Local: false
   Is Removable: false
   Is SSD: false
   Is Offline: false
   Is Perennially Reserved: false
   Queue Full Sample Size: 0
   Queue Full Threshold: 0
   Thin Provisioning Status: unknown
   Attached Filters:
   VAAI Status: unknown
   Other UIDs: vml.020000000005000c500643e03a3535439313030
   Is Local SAS Device: false
   Is Boot USB Device: false
   No of outstanding IOs with competing worlds: 32
```

RVC

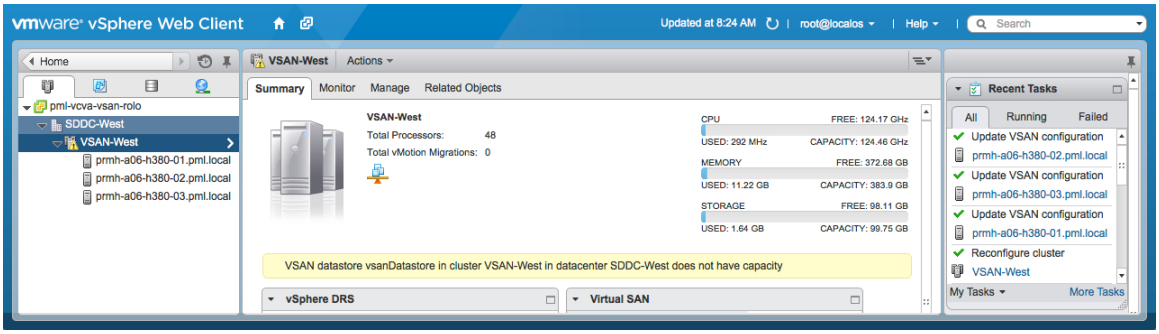
Use "vsan.disks_info" to gather detailed disks capabilities and characteristics such as size, disk type, manufactures, model, as well as identify if the disks are flagged as local or non-local.

```
> cd 1
/10.144.106.90> vsan.disks_info SDDC-East/computers/VSAN-DC01/hosts/w3r6c5-tm-stw07.eng.vmware.com/
Disks on host w3r6c5-tm-stw07.eng.vmware.com:
+-----------------------------------------------+------+-------+-----------------------------------------------------------------------------------------------+
| DisplayName                                   | isSSD | Size  | State                                                                                         |
+-----------------------------------------------+------+-------+-----------------------------------------------------------------------------------------------+
| ATA Serial Attached SCSI Disk (naa.5000c500643e2225) | MD   | 931 GB | inUse (The disk resides on a non-local storage transport: 'naa.5000c500643e2225'.)          |
| ATA ST91000640NS                              |      |       |                                                                                               |
+-----------------------------------------------+------+-------+-----------------------------------------------------------------------------------------------+
| ATA Serial Attached SCSI Disk (naa.5000c50064425545) | MD   | 931 GB | inUse (The disk resides on a non-local storage transport: 'naa.5000c50064425545'.)          |
| ATA ST91000640NS                              |      |       |                                                                                               |
+-----------------------------------------------+------+-------+-----------------------------------------------------------------------------------------------+
| Local ATA Disk (naa.55cd2e404b4cff48)         | SSD  | 279 GB | inUse                                                                                         |
| ATA INTEL SSDSC2BB30                          |      |       |                                                                                               |
+-----------------------------------------------+------+-------+-----------------------------------------------------------------------------------------------+
| ATA Serial Attached SCSI Disk (naa.5000c500643d99b0) | MD   | 931 GB | inUse (The disk resides on a non-local storage transport: 'naa.5000c500643d99b0'.)          |
| ATA ST91000640NS                              |      |       |                                                                                               |
+-----------------------------------------------+------+-------+-----------------------------------------------------------------------------------------------+
| Local USB Direct-Access (mpx.vmhba32:C0:T0:L0) | MD   | 14 GB  | ineligible (Existing partitions found on disk 'mpx.vmhba32:C0:T0:L0'.)                       |
| TOSHIBA TransMemory                           |      |       |                                                                                               |
|                                               |      |       | Partition table:                                                                              |
|                                               |      |       | 5: 0.24 GB, type = vfat                                                                       |
|                                               |      |       | 6: 0.24 GB, type = vfat                                                                       |
|                                               |      |       | 7: 0.11 GB, type = coredump                                                                   |
|                                               |      |       | 8: 0.28 GB, type = vfat                                                                       |
|                                               |      |       | 9: 2.50 GB, type = coredump                                                                   |
+-----------------------------------------------+------+-------+-----------------------------------------------------------------------------------------------+
| ATA Serial Attached SCSI Disk (naa.5000c50063cb01ad) | MD   | 931 GB | inUse (The disk resides on a non-local storage transport: 'naa.5000c50063cb01ad'.)          |
| ATA ST91000640NS                              |      |       |                                                                                               |
+-----------------------------------------------+------+-------+-----------------------------------------------------------------------------------------------+
| ATA Serial Attached SCSI Disk (naa.5000c500643e62c6) | MD   | 931 GB | inUse (The disk resides on a non-local storage transport: 'naa.5000c500643e62c6'.)          |
| ATA ST91000640NS                              |      |       |                                                                                               |
+-----------------------------------------------+------+-------+-----------------------------------------------------------------------------------------------+
```

```
/10.20.181.200> vsan.disks_info SDDC-West/computers/VSAN-PA/hosts/prmh-a06-h380-01.pml.local/
Disks on host prmh-a06-h380-01.pml.local:
+-----------------------------------------------------------+-------+--------+-----------------------------------------------------------------------------------------------------+
| DisplayName                                               | isSSD | Size   | State                                                                                               |
+-----------------------------------------------------------+-------+--------+-----------------------------------------------------------------------------------------------------+
| ATA Serial Attached SCSI Disk (naa.5000c5004d51a960)      | MD    | 465 GB | ineligible (Existing partitions found on disk 'naa.5000c5004d51a960'.)                              |
| ATA MM0500GBKAK                                           |       |        |                                                                                                     |
|                                                           |       |        | Partition table:                                                                                    |
|                                                           |       |        | 2: 4.00 GB, type = vfat                                                                             |
|                                                           |       |        | 3: 458.38 GB, type = vmfs ()                                                                        |
|                                                           |       |        | 5: 0.24 GB, type = vfat                                                                             |
|                                                           |       |        | 6: 0.24 GB, type = vfat                                                                             |
|                                                           |       |        | 7: 0.11 GB, type = coredump                                                                         |
|                                                           |       |        | 8: 0.28 GB, type = vfat                                                                             |
|                                                           |       |        | 9: 2.50 GB, type = coredump                                                                         |
+-----------------------------------------------------------+-------+--------+-----------------------------------------------------------------------------------------------------+
| ATA Serial Attached SCSI Disk (naa.5000c5004d51bcc0)      | MD    | 465 GB | inUse (The disk resides on a non-local storage transport: 'naa.5000c5004d51bcc0'.)                  |
| ATA MM0500GBKAK                                           |       |        |                                                                                                     |
+-----------------------------------------------------------+-------+--------+-----------------------------------------------------------------------------------------------------+
| Local ATA Disk (naa.5002538250006d92)                     | SSD   | 93 GB  | inUse                                                                                               |
| ATA MO0100EBTJT                                           |       |        |                                                                                                     |
+-----------------------------------------------------------+-------+--------+-----------------------------------------------------------------------------------------------------+
| ATA Serial Attached SCSI Disk (naa.5000c5004d51bff0)      | MD    | 465 GB | inUse (The disk resides on a non-local storage transport: 'naa.5000c5004d51bff0'.)                  |
| ATA MM0500GBKAK                                           |       |        |                                                                                                     |
+-----------------------------------------------------------+-------+--------+-----------------------------------------------------------------------------------------------------+
| ATA Serial Attached SCSI Disk (naa.5000c5004e91495a)      | MD    | 465 GB | inUse (The disk resides on a non-local storage transport: 'naa.5000c5004e91495a'.)                  |
| ATA MM0500GBKAK                                           |       |        |                                                                                                     |
+-----------------------------------------------------------+-------+--------+-----------------------------------------------------------------------------------------------------+
| DGC Fibre Channel Disk (naa.600601608261280 0166c8b743078e211) | MD    | 200 GB | ineligible (Existing partitions found on disk 'naa.600601608261280 0166c8b743078e211'.)             |
| DGC VRAID                                                 |       |        |                                                                                                     |
|                                                           |       |        | Partition table:                                                                                    |
|                                                           |       |        | 2: 4.00 GB, type = vfat                                                                             |
|                                                           |       |        | 5: 0.24 GB, type = vfat                                                                             |
|                                                           |       |        | 6: 0.24 GB, type = vfat                                                                             |
|                                                           |       |        | 7: 0.11 GB, type = coredump                                                                         |
|                                                           |       |        | 8: 0.28 GB, type = vfat                                                                             |
|                                                           |       |        | 9: 2.50 GB, type = coredump                                                                         |
+-----------------------------------------------------------+-------+--------+-----------------------------------------------------------------------------------------------------+
| ATA Serial Attached SCSI Disk (naa.5000c5004e9d68df)      | MD    | 465 GB | inUse (The disk resides on a non-local storage transport: 'naa.5000c5004e9d68df'.)                  |
| ATA MM0500GBKAK                                           |       |        |                                                                                                     |
+-----------------------------------------------------------+-------+--------+-----------------------------------------------------------------------------------------------------+
| ATA Serial Attached SCSI Disk (naa.5000c5004e9d46cb)      | MD    | 465 GB | inUse (The disk resides on a non-local storage transport: 'naa.5000c5004e9d46cb'.)                  |
| ATA MM0500GBKAK                                           |       |        |                                                                                                     |
+-----------------------------------------------------------+-------+--------+-----------------------------------------------------------------------------------------------------+
| ATA Serial Attached SCSI Disk (naa.5000c5004d48a3ca)      | MD    | 465 GB | inUse (The disk resides on a non-local storage transport: 'naa.5000c5004d48a3ca'.)                  |
| ATA MM0500GBKAK                                           |       |        |                                                                                                     |
+-----------------------------------------------------------+-------+--------+-----------------------------------------------------------------------------------------------------+
```

Disk Groups Creation Fails

Issue: After manually selecting a desired group of HDD and an SSD for the creation of a disk group, the operation is successfully completed but there are no disk groups created.

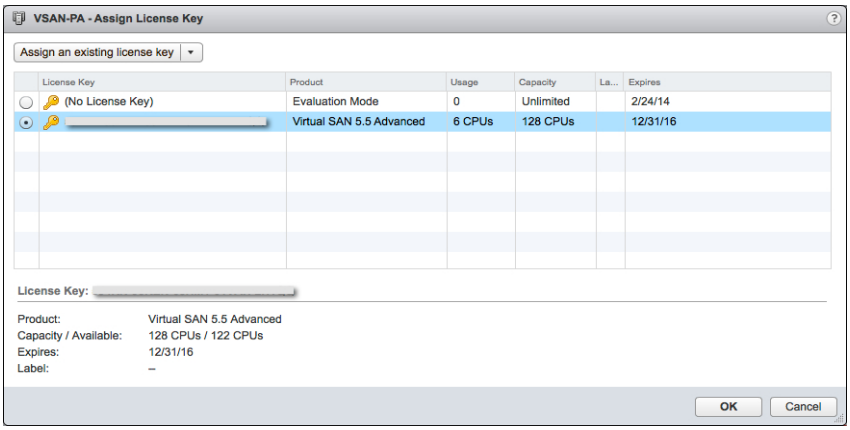


Solution: There are currently two possible issues that can be causing this behavior. In most cases, it has to do with Virtual SAN not being licensed correctly. Assign a Virtual SAN license to the cluster via the vSphere Web Client. The Virtual SAN feature is not automatically added to the cluster when a Virtual SAN license is added to the vCenter Server license catalog.

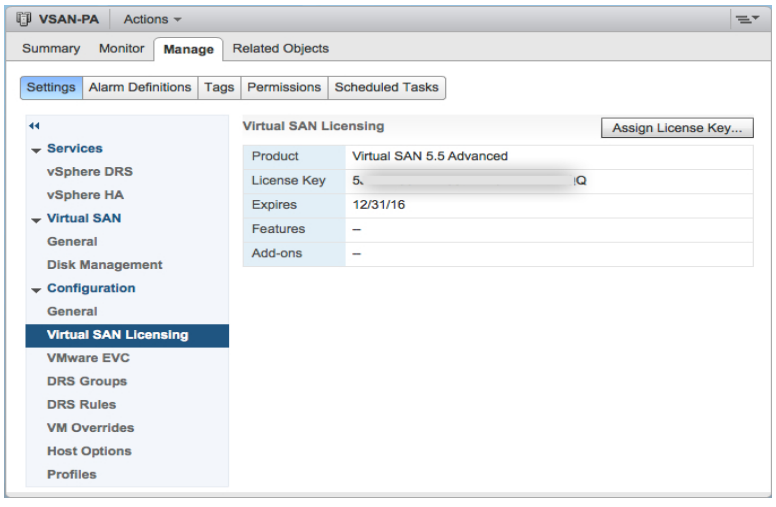
The second possibility is the vSphere Web Client refresh time out. Depending on the number of disk groups and the number of disks, the completion of the operation can take some time. Logout of they system and log back on.

vSphere Web Client

After enabling Virtual SAN in a vSphere Cluster, assigned the license to the Virtual SAN enabled cluster. From the Home screen go to > licenses > Cluster tab > Select cluster object > Assign License Key. The license key has to be assigned to the cluster in order to avoid any possible issues related to licensing.

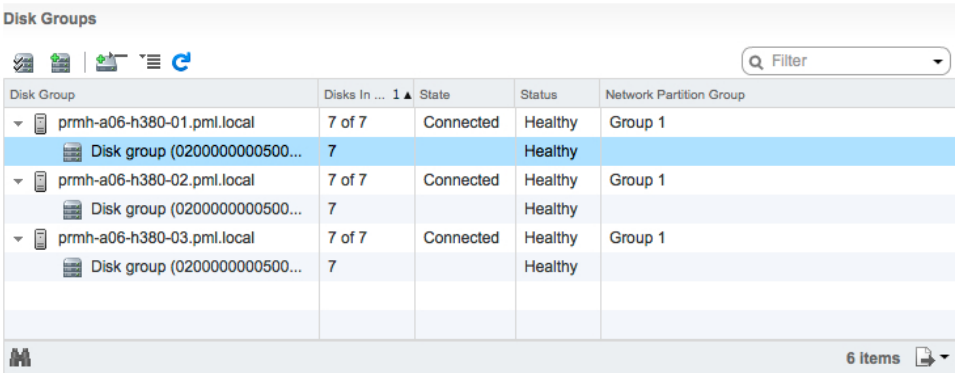


Verify the license has been successfully assigned to the cluster in the vSphere Web Client. Navigate to Host and Cluster > Cluster > Manage Tab > Settings > Virtual SAN Licensing.



Unable to delete Disk Group

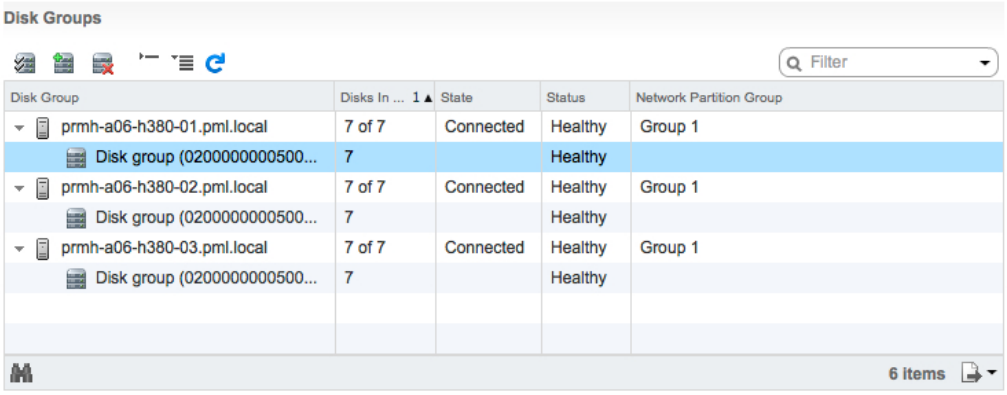
Issue: Unable to delete disk groups from the vSphere Web Client user interface.



Solution: The inability to delete disk groups is the result of Virtual SAN disk claiming operation being set to automatic. In order to be able to delete disk groups, modify the disk claiming operation and change it to manual.

vSphere Web Client

After the changing the add disk to storage setting to manual, the delete disk group icon will be displayed next to the "Add Disk" icon when a disk group is selected.



RVC

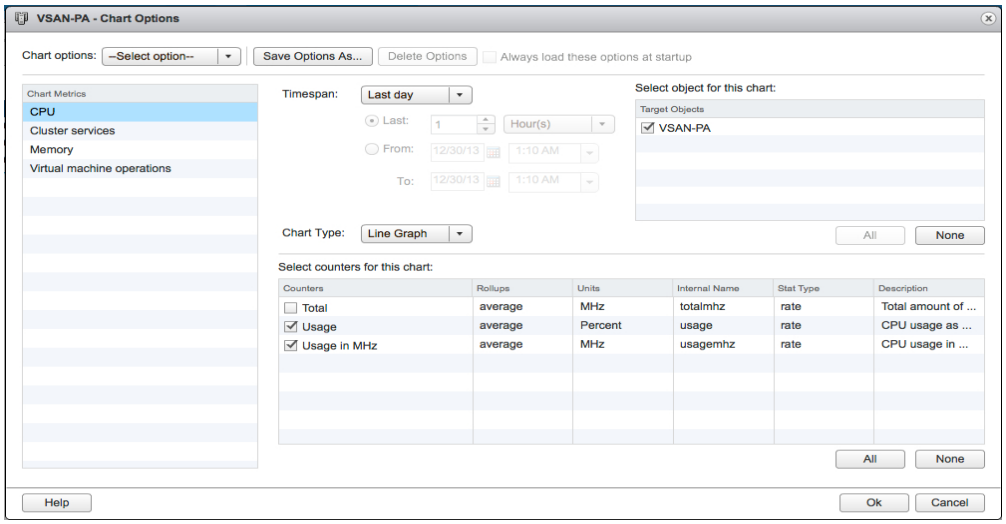
Use the "vsan.host_wipe_vsan_disks" command to wipe disks being used by Virtual SAN. The operation is to be performed individually in a per host basis. Use the **--force** command option to start the disk wiping process.

```
/10.20.181.200> vsan.host_wipe_vsan_disks SDDC-West/computers/VSAN-PA/hosts/prmh-a06-h380-01.pml.local/
Would wipe disk ATA Serial Attached SCSI Disk (naa.5000c5004d51bcc0) (ATA MM0500GBKAK, ssd = false)
Would wipe disk Local ATA Disk (naa.5002538250006d92) (ATA MO0100EBTJT, ssd = true)
Would wipe disk ATA Serial Attached SCSI Disk (naa.5000c5004d51bff0) (ATA MM0500GBKAK, ssd = false)
Would wipe disk ATA Serial Attached SCSI Disk (naa.5000c5004e91495a) (ATA MM0500GBKAK, ssd = false)
Would wipe disk ATA Serial Attached SCSI Disk (naa.5000c5004e9d68df) (ATA MM0500GBKAK, ssd = false)
Would wipe disk ATA Serial Attached SCSI Disk (naa.5000c5004e9d46cb) (ATA MM0500GBKAK, ssd = false)
Would wipe disk ATA Serial Attached SCSI Disk (naa.5000c5004d48a3ca) (ATA MM0500GBKAK, ssd = false)

NO ACTION WAS TAKEN. Use --force to actually wipe.
CAUTION: Wiping disks means all user data will be destroyed!
/10.20.181.200>
```

Unable to identify Virtual SAN Performance Stats

Issue: Unable to identify Virtual SAN performance stats in the vSphere Web Client user interface.



Solution: Use RVC's VSAN Observer to monitor and gather in-depth performance metrics for Virtual SAN. The vSphere Web Client is not currently able to display performance counters for Virtual SAN.

VSAN Observer

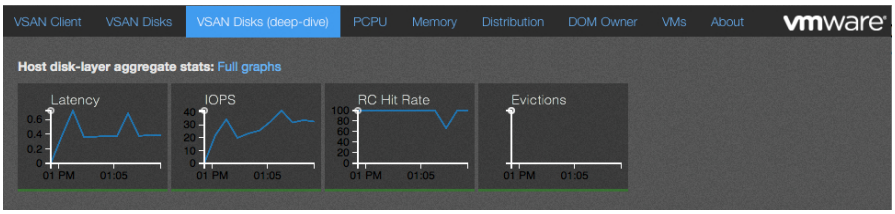
Launch the VSAN Observer to start collecting performance metrics. Typing the command "vsan.observer ~/computers/<cluster name> --run-webserver --force" to start the tool and use a modern web browser and access the metrics portal. VSAN Observer provides in-depth monitoring of Virtual SAN's physical disk layer performance, cache hit rates, latencies, etc.

Monitoring Virtual SAN Read Cache Utilization

Issue: Unable to monitor Virtual SAN read cache layer from the vSphere Web Client.

Solution: To monitor Virtual SAN's caching layer use the VSAN Observer. The VSAN Disks (deep-dive) screen displays hosts aggregated stats of all disks and disk groups as well as details about every physical disk.

VSAN Observer
Launch the VSAN Observer to start collecting performance metrics. Typing the command "vsan.observer ~/computers/<cluster name> --run-webserver --force" to start the tool and use a modern web browser and access the metrics portal. Go to the VSAN Disk (deep-dive) page and select a host from the "host to show list" to view the stats.



Monitoring Virtual SAN Related Logs

Issue: Location of Virtual SAN related logs.

Solution: Virtual SAN related logs as well as respective traces could be found under the /var/log directory on each host.

```
/var/log # ls -l
lrwxrwxrwx    1 root     root            19:33 clomd.log -> /scratch/log/clomd.log
lrwxrwxrwx    1 root     root            19:33 hostd.log -> /scratch/log/hostd.log
lrwxrwxrwx    1 root     root            19:33 osfsd.log -> /scratch/log/osfsd.log
lrwxrwxrwx    1 root     root            19:33 vpxa.log -> /scratch/log/vpxa.log
lrwxrwxrwx    1 root     root            16:37 vsantraces -> /scratch/vsantraces
lrwxrwxrwx    1 root     root            19:33 vsanvpd.log -> /scratch/log/vsanvpd.log
/var/log #
```

Cluster Level Object Manager (CLOM) Logs

- /var/log/clomd.log

Object Storage File System (OSFS)

- /var/log/osfsd.log

Hostd /vpxa / disklib / objectlib

- /var/log/hostd.log
- /var/log/vpxa.log

Disk Capacity

Issue: Identification of used and reserved disk capacity in Virtual SAN

```
/10.144.106.90> vsan.disks_stats SDDC-East/computers/VSAN-DC01/
+---------------------+------------------------------+-------+-----+----------+------+----------+--------+
|                     |                              |       | Num | Capacity |      |          | Status |
| DisplayName         | Host                         | isSSD | Comp | Total    | Used | Reserved | Health |
+---------------------+------------------------------+-------+-----+----------+------+----------+--------+
| naa.55cd2e404b46754b | w3r6c5-tm-stw01.eng.vmware.com | SSD   | 0    | 195.62 GB | 0 %  | 0 %      | OK     |
| naa.5000c5006439c3b4 | w3r6c5-tm-stw01.eng.vmware.com | MD    | 3    | 931.25 GB | 0 %  | 0 %      | OK     |
| naa.5000c500643e03a3 | w3r6c5-tm-stw01.eng.vmware.com | MD    | 4    | 931.25 GB | 0 %  | 0 %      | OK     |
| naa.5000c500643e1dee | w3r6c5-tm-stw01.eng.vmware.com | MD    | 3    | 931.25 GB | 0 %  | 0 %      | OK     |
| naa.5000c500643eb6a4 | w3r6c5-tm-stw01.eng.vmware.com | MD    | 4    | 931.25 GB | 0 %  | 0 %      | OK     |
| naa.5000c500643e36b5 | w3r6c5-tm-stw01.eng.vmware.com | MD    | 13   | 931.25 GB | 0 %  | 0 %      | OK     |
+---------------------+------------------------------+-------+-----+----------+------+----------+--------+
```

Solution: use the RVC "disks_stats" command for complete output of hosts and disks in a Virtual SAN cluster. The percentage of used and reserved capacity is displayed

Monitoring Virtual SAN Component Limits

Issue: Determination of Virtual SAN component count against the maximum number of components allowed per host.

```
/10.144.106.90> vsan.check_limits SDDC-East/computers/VSAN-DC01/
2014-01-15 06:23:45 +0000: Gathering stats from all hosts ...
2014-01-15 06:23:47 +0000: Gathering disks info ...
+---------------------------------+-------------------+-------------------------+
| Host                            | RDT               | Disks                   |
+---------------------------------+-------------------+-------------------------+
| w3r6c5-tm-stw01.eng.vmware.com | Assocs: 38/20000  | Components: 27/3000     |
|                                 | Sockets: 37/10000 | naa.5000c500643e03a3: 0% |
|                                 | Clients: 0        | naa.5000c500643eb6a4: 0% |
|                                 | Owners: 2         | naa.5000c500643e36b5: 0% |
|                                 |                   | naa.5000c5006439c3b4: 0% |
|                                 |                   | naa.5000c500643e1dee: 0% |
|                                 |                   | naa.55cd2e404b46754b: 0% |
```

Solution: Use the RVC command "check_limits" in order to identify components count per host.

Host Failure Impact

Issue: Impact of a host failure in Virtual SAN.

```
/10.144.106.90> vsan.whatif_host_failures SDDC-East/computers/VSAN-DC01/
Simulating 1 host failures:
+-----------------+-----------------------------+-----------------------------------+
| Resource        | Usage right now             | Usage after failure/re-protection |
+-----------------+-----------------------------+-----------------------------------+
| HDD capacity    |   0% used (37147.79 GB free) |   0% used (32491.54 GB free)      |
| Components      |   1% used (23771 available)  |   1% used (20771 available)       |
| RC reservations |   0% used (1564.96 GB free)  |   0% used (1369.34 GB free)       |
+-----------------+-----------------------------+-----------------------------------+
```

Solution: use the RVC command "whatif_host_failure" in order to identify the difference between the current available resources and the results if a failure were to occur.

Conclusion

In summary, this troubleshooting and monitoring reference guide was developed to provide quick guidance around some of the most common configuration issues and performance monitoring inquiries related to Virtual SAN.

Acknowledgements

I would like to thank Christian Dickmann & Christos Karamanolis of VMware R&D, whose deep knowledge & understanding of Virtual SAN was leveraged throughout this paper. I would also like to thank Cormac Hogan and Duncan Epping from VMware R&D, Wade Holmes, Senior Architect in the Storage and Availability Technical Marketing group and Charu Chaubal, Senior Manager of Storage and Availability Technical Marketing group for reviewing and contributing to this paper.

About the Author

Rawlinson is a Senior Technical Marketing Architect in the Cloud Infrastructure Technical Marketing Group at VMware, Inc. focused on Storage Virtualization technologies. Previously he was an Architect in VMware's Cloud Infrastructure & Management Professional Services Organization focused on vSphere and Cloud enterprise architectures for VMware's fortune 100 and 500 customers.

Rawlinson is amongst the first VMware Certified Design Experts (VCDX#86) and the author of multiple books based on VMware and other technologies.

• Follow Rawlinson’s blogs:

- http://blogs.vmware.com/vsphere/storage
- http://www.punchingclouds.com

• Follow Rawlinson on Twitter:

- @PunchingClouds

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