VSAN

Zum Thema VSAN gibt es einige sehr schöne Demos zum Durchklicken unter

https://storagehub.vmware.com/#!/vmware-vsan/vmware-vsan-demonstrations

TBD https://via.vmw.com/tchzcoreno1848 https://youtu.be/nvrCShwWMWY

TRIM/UNMAP

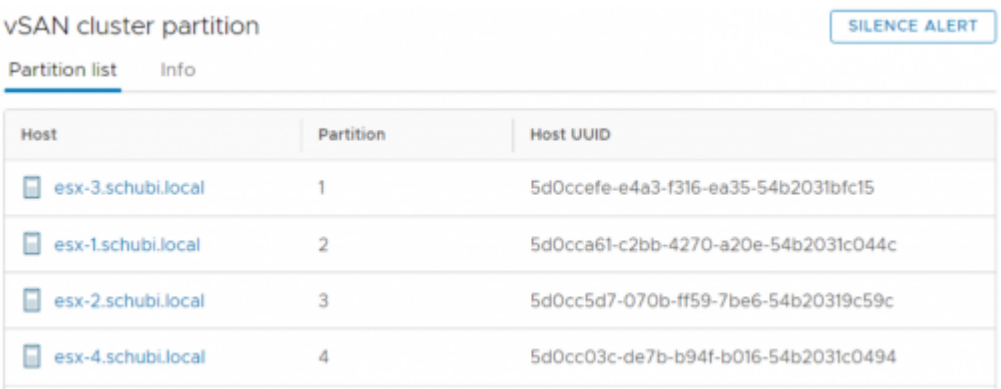
https://knowledge.broadcom.com/external/article/326595/procedure-to-enable-trimunmap.html

vSAN mit CLI ohne vCenter

https://blog.rylander.io/2017/01/19/configure-2-node-vsan-on-esxi-free-using-cli-without-vcenter/

vSAN Cluster Partitions beheben

Im vSAN Cluster kann es dazu kommen, dass die einzelnen Nodes partitioniert sind. Meist ist nur eine einzelne Node weg, manchmal aber auch alle ☹ Auch wenn auf den vSAN Kernelport alle Hosts sich sehen können, finden sich die Partitions nicht zusammen.



Im schlimmsten Fall sieht das so aus.

Erst mal sollte man alle vSAN Ports testen, ob die sich untereinander pingen lassen.

```
[root@esx-4:~] vmkping -d -s 7000 -I vmk2 192.168.4.11
PING 192.168.4.11 (192.168.4.11): 7000 data bytes
7008 bytes from 192.168.4.11: icmp_seq=0 ttl=64 time=0.615 ms
7008 bytes from 192.168.4.11: icmp_seq=1 ttl=64 time=0.854 ms
7008 bytes from 192.168.4.11: icmp_seq=2 ttl=64 time=0.765 ms
```

```
--- 192.168.4.11 ping statistics ---
3 packets transmitted, 3 packets received, 0% packet loss
round-trip min/avg/max = 0.615/0.745/0.854 ms
```

Was sagt der Cluster?

```
[root@esx-4:~] esxcli vsan cluster get
Cluster Information
   Enabled: true
   Current Local Time: 2020-11-15T12:18:22Z
   Local Node UUID: 5d0cc03c-de7b-b94f-b016-54b2031c0494
   Local Node Type: NORMAL
   Local Node State: MASTER
   Local Node Health State: HEALTHY
   Sub-Cluster Master UUID: 5d0cc03c-de7b-b94f-b016-54b2031c0494
   Sub-Cluster Backup UUID:
   Sub-Cluster UUID: 52c9d6db-b533-a36c-6030-98d37f927e6a
   Sub-Cluster Membership Entry Revision: 2
   Sub-Cluster Member Count: 1
   Sub-Cluster Member UUIDs: 5d0cc03c-de7b-b94f-b016-54b2031c0494
   Sub-Cluster Member HostNames: esx-4.schubi.local
   Sub-Cluster Membership UUID: af16b15f-9a9b-99f7-989e-54b2031c0494
   Unicast Mode Enabled: true
   Maintenance Mode State: OFF
   Config Generation: 27b68356-c3e0-4c64-ab66-3b80741ab493 14
2020-10-19T20:29:54.524
```

Sieht auf jedem Cluster so aus. Jeweils nur ein Member. Aber sie “fühlen” sich dem gleichen Cluster zugehörig. Überall ist `Sub-Cluster UUID: 52c9d6db-b533-a36c-6030-98d37f927e6a` gleich.

Mal sehen wir die UniCast Agenten sich fühlen. Da gibts bei mir eine Überraschung

```
[root@esx-1:~] esxcli vsan cluster unicastagent list
NodeUuid                                IsWitness  Supports Unicast  IP
Address      Port  Iface Name  Cert Thumbprint
SubClusterUuid
------------------------------------  ---------  ---------------  ---------
----  -----  ----------  -----------------------------------------------------
--------  --------------
5d0cc03c-de7b-b94f-b016-54b2031c0494          0             true
192.168.4.232  12321
EB:EF:07:00:20:7B:09:97:AD:EB:34:27:3D:B0:A7:32:8B:CF:57:CE  52c9d6db-b533-
a36c-6030-98d37f927e6a
5d0cc5d7-070b-ff59-7be6-54b20319c59c          0             true
192.168.4.149  12321
09:C6:4C:7B:AD:5F:4B:7C:29:7D:BE:DC:4C:95:04:7E:02:32:35:B5  52c9d6db-b533-
a36c-6030-98d37f927e6a
5d0ccefe-e4a3-f316-ea35-54b2031bfc15          0             true
192.168.4.151  12321
B6:89:41:DC:9B:20:72:1B:19:E3:8A:C6:F8:49:AC:6D:93:8C:A3:76  52c9d6db-b533-
a36c-6030-98d37f927e6a
```

```
[root@esx-2:~] esxcli vsan cluster unicastagent list
NodeUuid                                IsWitness  Supports Unicast  IP
Address      Port  Iface Name  Cert Thumbprint
SubClusterUuid
------------------------------------  ---------  ---------------  ---------
----  -----  ----------  -----------------------------------------------------
--------  --------------
5d0cc03c-de7b-b94f-b016-54b2031c0494          0              true
192.168.4.232  12321
EB:EF:07:00:20:7B:09:97:AD:EB:34:27:3D:B0:A7:32:8B:CF:57:CE  52c9d6db-b533-
a36c-6030-98d37f927e6a
5d0cca61-c2bb-4270-a20e-54b2031c044c          0              true
192.168.4.219  12321
59:E9:AA:76:F8:14:55:F2:5D:DE:B8:AA:0E:0D:D0:30:4C:39:77:21  52c9d6db-b533-
a36c-6030-98d37f927e6a
5d0ccefe-e4a3-f316-ea35-54b2031bfc15          0              true
192.168.4.151  12321
B6:89:41:DC:9B:20:72:1B:19:E3:8A:C6:F8:49:AC:6D:93:8C:A3:76  52c9d6db-b533-
a36c-6030-98d37f927e6a

[root@esx-3:~] esxcli vsan cluster unicastagent list
NodeUuid                                IsWitness  Supports Unicast  IP
Address      Port  Iface Name  Cert Thumbprint
SubClusterUuid
------------------------------------  ---------  ---------------  ---------
----  -----  ----------  -----------------------------------------------------
--------  --------------
5d0cc03c-de7b-b94f-b016-54b2031c0494          0              true
192.168.4.232  12321
EB:EF:07:00:20:7B:09:97:AD:EB:34:27:3D:B0:A7:32:8B:CF:57:CE  52c9d6db-b533-
a36c-6030-98d37f927e6a
5d0cc5d7-070b-ff59-7be6-54b20319c59c          0              true
192.168.4.149  12321
09:C6:4C:7B:AD:5F:4B:7C:29:7D:BE:DC:4C:95:04:7E:02:32:35:B5  52c9d6db-b533-
a36c-6030-98d37f927e6a
5d0cca61-c2bb-4270-a20e-54b2031c044c          0              true
192.168.4.219  12321
59:E9:AA:76:F8:14:55:F2:5D:DE:B8:AA:0E:0D:D0:30:4C:39:77:21  52c9d6db-b533-
a36c-6030-98d37f927e6a

[root@esx-4:~] esxcli vsan cluster unicastagent list
NodeUuid                                IsWitness  Supports Unicast  IP
Address      Port  Iface Name  Cert Thumbprint
SubClusterUuid
------------------------------------  ---------  ---------------  ---------
----  -----  ----------  -----------------------------------------------------
--------  --------------
5d0cc5d7-070b-ff59-7be6-54b20319c59c          0              true
192.168.4.149  12321
09:C6:4C:7B:AD:5F:4B:7C:29:7D:BE:DC:4C:95:04:7E:02:32:35:B5  52c9d6db-b533-
```

```
a36c-6030-98d37f927e6a
5d0cca61-c2bb-4270-a20e-54b2031c044c          0              true
192.168.4.219  12321
59:E9:AA:76:F8:14:55:F2:5D:DE:B8:AA:0E:0D:D0:30:4C:39:77:21  52c9d6db-b533-
a36c-6030-98d37f927e6a
5d0ccefe-e4a3-f316-ea35-54b2031bfc15          0              true
192.168.4.151  12321
B6:89:41:DC:9B:20:72:1B:19:E3:8A:C6:F8:49:AC:6D:93:8C:A3:76  52c9d6db-b533-
a36c-6030-98d37f927e6a
```

Die haben die aktuellen IOs *.4.11 bis *.4.14 gar nicht “gefressen”. Zum Thema UniCast Agent gibt es einen VMware KB https://kb.vmware.com/s/article/2150303.

Gehen wir mal die Liste durch.

```
[root@esx-1:~] esxcli vsan network list
Interface
   VmkNic Name: vmk2
   IP Protocol: IP
   Interface UUID: fbf7215f-49cd-05cd-f9d8-54b2031c044c
   Agent Group Multicast Address: 224.2.3.4
   Agent Group IPv6 Multicast Address: ff19::2:3:4
   Agent Group Multicast Port: 23451
   Master Group Multicast Address: 224.1.2.3
   Master Group IPv6 Multicast Address: ff19::1:2:3
   Master Group Multicast Port: 12345
   Host Unicast Channel Bound Port: 12321
   Data-in-Transit Encryption Key Exchange Port: 0
   Multicast TTL: 5
   Traffic Type: vsan
```

Überall gut. Die Kernelport IP Config sieht auch gut aus.

```
[root@esx-1:~]  esxcli network ip interface ipv4 get | grep vmk2
vmk2  192.168.4.11   255.255.255.0  192.168.4.255   STATIC
192.168.4.1      false
```

Für das weitere Vorgehen benötigen wir die UUID.

```
[root@esx-1:~]  cmmds-tool whoami
5d0cca61-c2bb-4270-a20e-54b2031c044c
[root@esx-2:~]  cmmds-tool whoami
5d0cc5d7-070b-ff59-7be6-54b20319c59c
[root@esx-3:~]  cmmds-tool whoami
5d0ccefe-e4a3-f316-ea35-54b2031bfc15
[root@esx-4:~] cmmds-tool whoami
5d0cc03c-de7b-b94f-b016-54b2031c0494
```

Damit unsere folgenden manuellen Änderungen nicht durch ein Update des vCenters gestört wird, ignorieren wir die vCenter Settings temporär.

```
[root@esx-1:~] esxcfg-advcfg -s 1 /VSAN/IgnoreClusterMemberListupdates
Value of IgnoreClusterMemberListUpdates is 1
```

Wichtig! Niemals auf einem Host die eigene Agent IP hinzufügen! VMware: *“When an ESXi host has its own IP address in its Unicast agent list, many networking problems can arise and potentially lead to the host encountering a PSOD.”* Also “Augen auf” bei den folgenden Schritten!

Für den esx1 Host müssen 2, 3 und 4 Hinzugefügt werden

```
esxcli vsan cluster unicastagent add -t node -u 5d0cc5d7-070b-
ff59-7be6-54b20319c59c -U true -a 192.168.4.12 -p 12321
esxcli vsan cluster unicastagent add -t node -u 5d0ccefe-e4a3-f316-
ea35-54b2031bfc15 -U true -a 192.168.4.13 -p 12321
esxcli vsan cluster unicastagent add -t node -u 5d0cc03c-de7b-b94f-
b016-54b2031c0494 -U true -a 192.168.4.14 -p 12321
```

Ein Check des Hosts bringt noch nicht ganz die 100%ige Überzeugung.

```
[root@esx-1:~] esxcli vsan cluster unicastagent list
NodeUuid                              IsWitness  Supports Unicast  IP
Address      Port  Iface Name  Cert Thumbprint
SubClusterUuid
------------------------------------  ---------  ---------------  ---------
----  -----  ----------  -----------------------------------------------
--------  --------------
5d0cc03c-de7b-b94f-b016-54b2031c0494          0              true
192.168.4.232  12321
EB:EF:07:00:20:7B:09:97:AD:EB:34:27:3D:B0:A7:32:8B:CF:57:CE  52c9d6db-b533-
a36c-6030-98d37f927e6a
5d0cc03c-de7b-b94f-b016-54b2031c0494          0              true
192.168.4.14   12321
52c9d6db-b533-a36c-6030-98d37f927e6a
5d0cc5d7-070b-ff59-7be6-54b20319c59c          0              true
192.168.4.149  12321
09:C6:4C:7B:AD:5F:4B:7C:29:7D:BE:DC:4C:95:04:7E:02:32:35:B5  52c9d6db-b533-
a36c-6030-98d37f927e6a
5d0cc5d7-070b-ff59-7be6-54b20319c59c          0              true
192.168.4.12   12321
52c9d6db-b533-a36c-6030-98d37f927e6a
5d0ccefe-e4a3-f316-ea35-54b2031bfc15          0              true
192.168.4.151  12321
B6:89:41:DC:9B:20:72:1B:19:E3:8A:C6:F8:49:AC:6D:93:8C:A3:76  52c9d6db-b533-
a36c-6030-98d37f927e6a
5d0ccefe-e4a3-f316-ea35-54b2031bfc15          0              true
192.168.4.13   12321
52c9d6db-b533-a36c-6030-98d37f927e6a
```

Aber zum Glück kann man “überflüssige” Einträge löschen...

```
esxcli vsan cluster unicastagent remove -a 192.168.4.232
```

```
esxcli vsan cluster unicastagent remove -a 192.168.4.149
esxcli vsan cluster unicastagent remove -a 192.168.4.151
```

So sieht es nun gut aus.

```
[root@esx-1:~]  esxcli vsan cluster unicastagent list
NodeUuid                              IsWitness  Supports Unicast  IP
Address     Port  Iface Name  Cert Thumbprint  SubClusterUuid
------------------------------------  ---------  ----------------  ---------
---  -----  ----------  ---------------  --------------
5d0cc03c-de7b-b94f-b016-54b2031c0494          0              true
192.168.4.14  12321                                52c9d6db-b533-
a36c-6030-98d37f927e6a
5d0cc5d7-070b-ff59-7be6-54b20319c59c          0              true
192.168.4.12  12321                                52c9d6db-b533-
a36c-6030-98d37f927e6a
5d0ccefe-e4a3-f316-ea35-54b2031bfc15          0              true
192.168.4.13  12321                                52c9d6db-b533-
a36c-6030-98d37f927e6a
```

Exemplarisch noch für die weiteren Hosts

ESX2:

```
esxcli vsan cluster unicastagent add -t node -u 5d0cca61-c2bb-4270-
a20e-54b2031c044c -U true -a 192.168.4.11 -p 12321
esxcli vsan cluster unicastagent add -t node -u 5d0ccefe-e4a3-f316-
ea35-54b2031bfc15 -U true -a 192.168.4.13 -p 12321
esxcli vsan cluster unicastagent add -t node -u 5d0cc03c-de7b-b94f-
b016-54b2031c0494 -U true -a 192.168.4.14 -p 12321
```

ESX3:

```
esxcli vsan cluster unicastagent add -t node -u 5d0cca61-c2bb-4270-
a20e-54b2031c044c -U true -a 192.168.4.11 -p 12321
esxcli vsan cluster unicastagent add -t node -u 5d0cc5d7-070b-
ff59-7be6-54b20319c59c -U true -a 192.168.4.12 -p 12321
esxcli vsan cluster unicastagent add -t node -u 5d0cc03c-de7b-b94f-
b016-54b2031c0494 -U true -a 192.168.4.14 -p 12321
```

ESX4:

```
esxcli vsan cluster unicastagent add -t node -u 5d0cca61-c2bb-4270-
a20e-54b2031c044c -U true -a 192.168.4.11 -p 12321
esxcli vsan cluster unicastagent add -t node -u 5d0cc5d7-070b-
ff59-7be6-54b20319c59c -U true -a 192.168.4.12 -p 12321
esxcli vsan cluster unicastagent add -t node -u 5d0ccefe-e4a3-f316-
ea35-54b2031bfc15 -U true -a 192.168.4.13 -p 12321
```

Danach natürlich wieder Kontrolle mit

```
 esxcli vsan cluster unicastagent list
```

und ggf. Löschen mit

```
esxcli vsan cluster unicastagent remove -a
```

Kaum macht man es richtig, schon geht's. Ich habe einen Master, einen Backup, der Rest Agent. Hier mal ein Auszug.

```
[root@esx-1:~]  esxcli vsan cluster get
Cluster Information
   Enabled: true
   Current Local Time: 2020-11-15T13:10:16Z
   Local Node UUID: 5d0cca61-c2bb-4270-a20e-54b2031c044c
   Local Node Type: NORMAL
   Local Node State: BACKUP
   Local Node Health State: HEALTHY
   Sub-Cluster Master UUID: 5d0cc5d7-070b-ff59-7be6-54b20319c59c
   Sub-Cluster Backup UUID: 5d0cca61-c2bb-4270-a20e-54b2031c044c
   Sub-Cluster UUID: 52c9d6db-b533-a36c-6030-98d37f927e6a
   Sub-Cluster Membership Entry Revision: 4
   Sub-Cluster Member Count: 4
   Sub-Cluster Member UUIDs: 5d0cc5d7-070b-ff59-7be6-54b20319c59c, 5d0cca61-
c2bb-4270-a20e-54b2031c044c, 5d0ccefe-e4a3-f316-ea35-54b2031bfc15, 5d0cc03c-
de7b-b94f-b016-54b2031c0494
   Sub-Cluster Member HostNames: esx-2.schubi.local, esx-1.schubi.local,
esx-3.schubi.local, esx-4.schubi.local
   Sub-Cluster Membership UUID: b216b15f-40f2-2395-1900-54b20319c59c
   Unicast Mode Enabled: true
   Maintenance Mode State: OFF
   Config Generation: 27b68356-c3e0-4c64-ab66-3b80741ab493 20
2020-11-15T12:57:11.0
```

Zum Schluss nicht vergessen, den Advanced Parameter zu ersetzen:

```
esxcfg-advcfg -s 0 /VSAN/IgnoreClusterMemberListupdates
```

vSAN Services checken und starten

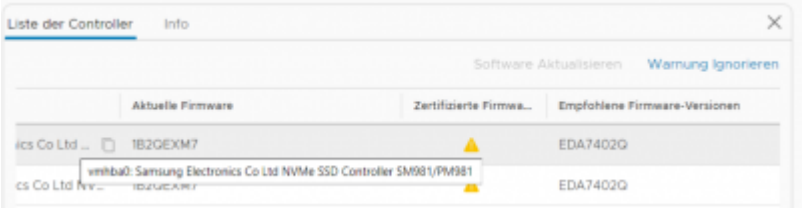
```
/etc/init.d/cmmdsd status && /etc/init.d/epd status && /etc/init.d/clomd
status

/etc/init.d/cmmdsd restart && /etc/init.d/epd restart && /etc/init.d/clomd
restart
```

ESXi Infos in der RVC

```
/localhost/DC/computers/vSAN-Cluster> table -f name -s name -f
state.connection:state.maintenancemode:build:uptime:num.vms:num.poweredonvms
:cpuusage:memusage hosts/*
```

Samsung 970 EVO PlusFirmware

Bei den 970er Evos wird bei neueren vSAN Releases evtl. die Firmware angemeckert.




```
[root@esx-1:~] esxcli nvme device get -A vmhba1 | egrep "Serial Number|Model
Number|Firmware Revision"
   Serial Number: S4EUNG0M219966P
   Model Number: Samsung SSD 970 EVO Plus 250GB
   Firmware Revision: 1B2QEXM7
```

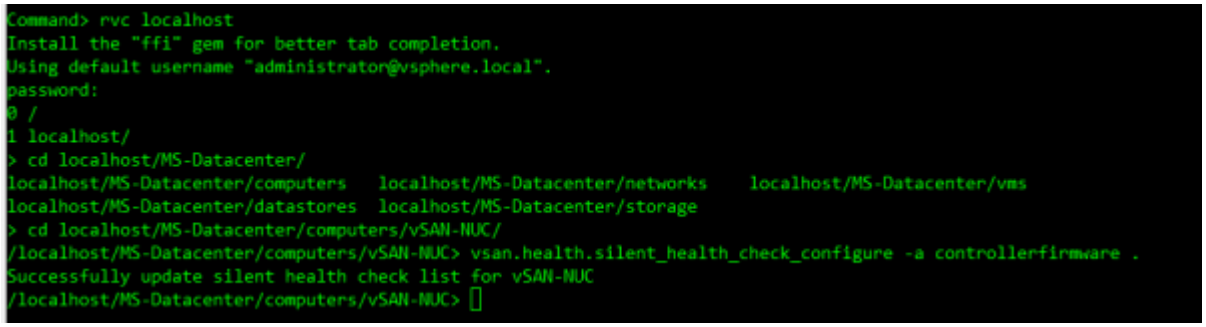
Richtig wäre hier aber FW EDA7402Q.

Die Firmware findet man unter https://www.samsung.com/semiconductor/minisite/ssd/download/tools/

Aber da ist keine EDA7402Q zu finden 😕 Die höchste Version ist 2B2QEXM7. Weitere Infos für den firmware Download können auch unter https://tinkertry.com/ssd-magician-and-firmware-download-limits-exceeded-samsung-seriously gefunden werden. Das ISO findet man unter http://ssd.samsungsemi.com/ecomobile/ssd/update14.do?fname=/Samsung_SSD_960_PRO_2B6QCXP7.iso

Also wird der Healthcheck nach https://www.virten.net/2017/04/how-to-silence-vsan-health-checks/ abgeschaltet. RVC:

```
vsan.health.silent_health_check_configure -a controllerfirmware .
```



Netzwerk checken

```
esxcli network ip interface set  -m 9000 -i vmk3

esxcli vsan health cluster list

esxcli network ip interface list

esxcli vsan network ip add -i vmk0 -T=witness
```

Durchsatz mit iperf checken:

- Firewall ausschalten auf dem Ziel -

  ```
  esxcli network firewall set --enabled false
  ```

- iperf im Listen Modus unter Angabe der horchenden IP Angeben -

  ```
  /usr/lib/vmware/vsan/bin/iperf3.copy -s -B <Kernelport IP>
  ```

- auf der Quelle die Firewall ausschalten -

  ```
  esxcli network firewall set --enabled false
  ```

- Auf der Quelle iperf unter Angabe der Ziel-IP starten -

  ```
  /usr/lib/vmware/vsan/bin/iperf3.copy -c <Kernelport IP>
  ```

- nach dem Test nicht vergessen, die Firewall wieder einzuschalten -

  ```
  esxcli network firewall set --enabled true
  ```

nützliche Links

https://www.vmware.com/try-vmware/vsan-new-hol-labs.html

https://blogs.vmware.com/virtualblocks/2017/04/05/m-2-ssd-boot-device-vsan/

https://blogs.vmware.com/virtualblocks/2017/01/18/designing-vsan-disk-groups-cache-ratio-revisited/

https://nolabnoparty.com/en/virtual-san-2-node-cluster-installtion-robo-pt1/

https://blogs.vmware.com/virtualblocks/2018/02/12/microsoft-sql-server-database-on-vmware-vsan-day-2-operations-and-management/

https://github.com/equelin/vsanmetrics

https://storagehub.vmware.com/t/vmware-vsan/vsan-poc-performance-checklist/

https://www.virten.net/2017/04/how-to-silence-vsan-health-checks/

https://www.altaro.com/vmware/how-to-generate-vsan-html-report-powercli/

https://storagehub.vmware.com/t/vmware-vsan/vsan-cluster-design-large-clusters-versus-small-clusters/

https://blogs.vmware.com/performance/2018/12/hcibench-specific-issues-recommendations-vsan.html

https://storagehub.vmware.com/section-assets/powercli-cookbook-for-vsan

https://storagehub.vmware.com/t/vmware-vsan/vmworld/vmworld-vsan-sessions/1

VSAN Cluster verschieben

http://www.virtuallyghetto.com/2014/09/how-to-move-a-vsan-cluster-from-one-vcenter-server-to-another.html

Default Repair Delay Time anpassen

https://kb.vmware.com/selfservice/microsites/search.do?language=en_US&cmd=displayKC&externalId=2075456

```
esxcli system settings advanced set -o /VSAN/ClomRepairDelay -i <value in minutes>

/etc/init.d/clomd restart
```

2-Node Cluster with esxcli

https://blog.rylander.io/2017/01/19/configure-2-node-vsan-on-esxi-free-using-cli-without-vcenter/

VSAN Fehlerscenarien und Reaktion darauf

VSAN Cluster lokal

VSAN stretched Cluster/Fault Domains

Ich betrachte hier mal vorrangig eine VSAN Cluster mit zwei Seiten und einer Witness Appliance.

Hintergrund und Konfiguration von Fault Domains ist am einfachsten bei Cormac Hogan nachzulesen.

http://cormachogan.com/2015/04/20/vsan-6-0-part-8-fault-domains/

http://cormachogan.com/2017/03/10/2-node-vsan-topologies-review/

http://cormachogan.com/2015/09/11/a-closer-look-at-the-vsan-witness-appliance/

http://cormachogan.com/2015/09/14/step-by-step-deployment-of-the-vsan-witness-appliance/

Ausfall einer Platte

In einem Stretched Cluster wird von der lokalen Seite gelesen und auf beiden Seiten geschrieben. Bei Ausfall einer Platte erfolgt:

- Lesen und Schreiben der betroffenen Objekte von der Remote Site
- Die Objekte werden auf anderen Disks der lokalen Seite neu gespiegelt.
    - nach 60 Minuten, wenn die Platte als “nicht da” markiert wird
    - sofort, wenn die Platte als defekt markiert wird
- im Hybrid Cluster muss der Cache auf der Remote Site der betroffenen VMs erst einmal angewärmt werden (Advanced Parameter existiert, um beide Seiten für Lese- und **Schreib**operationen im Normalbetrieb zu nutzen (Warmer Cache)

Ausfall eines Hosts

Ein Ausfall deines Hosts entspricht dem Ausfall einer Disk, nur das wesentlich mehr Componenten betroffen sind HA startet die VMs neu und die Objekte werden neu auf Disks anderer Hosts synchronisiert.

Ausfall Netzverbindungen zu Hosts einer Seite

Wenn auf einer Seite Netzwerkprobleme mit ein oder mehreren Hosts auftreten, werden folgende Aktionen durchgeführt:

- bei einem Host
    - Host wird als isoliert erkannt
    - Der VSAN Host hat kein Mehrheitsprinzip und die VMs können auf dem Host nicht mehr zu VSAN Objekten schreiben
    - HA stoppt VMs und startet diese auf anderen Knoten neu (“Shutdown” als Isolation Response ist wichtig!)
    - der isolierte Host hat keine VSAN Aktivität mehr
    - Die betroffenen Objekte werden auf anderen Hosts neu synchronisiert
- wenn mehrere Hosts einer Seite vom Rest des Clusters isoliert werden
    - Es wird eine neue Network Group in diesen Hosts etabliert (in der vCenter WebGUI sieht man Gruppen mit fortlaufenden Nummern)
    - Die Components dieser neuen Network Group haben aber keine N+1 Mehrheit mit der Witness und verlieren die Vote
    - HA startet die VMs auf anderen Hosts neu

Ausfall einer kompletten Seite

Wenn eine komplette Seite ausfällt, gibt's folgende Reaktion:

- Die Components der verblieben Seite bilden mit der Witness zusammen die Mehrheit. Somit wird diese Seite als “aktiv” erkannt
- die VMs der verbliebenen Seite laufen weiter
- die VMs der Crash-Seite werden mittels HA auf der anderen Seite (Fault Domain) neu gestartet.
- ist die Seite wieder da, werden die Components wieder aussynchronisiert und alles läuft normal weiter. Dabei kann mit DRS Regeln Einfluss genommen werden, wie die Maschinen wieder zurück wandern.

Ausfall der Netzverbindung zwischen den Seiten

Gibt es einen Netzausfall zwischen den Seiten, aber die Witness ist noch erreichbar, kommt zum ersten mal die “Preferred Site” zum Einsatz.

- Beide Seiten können nicht mehr synchron schreiben
- Beide Seiten haben über die Witness “eigentlich” ein N+1 Verhältnis
- Dadurch, das beide Seiten die Witness sehen, wird die “Preferred Site” als einzige aktive Seite markiert.
- HA schaltet auf der Secondary Site die VMs aus und startet diese auf der “Preferred Site”
- Ist die Netzverbindung wieder da, werden die Components neu aussynchronisiert und VSAN läuft normal weiter.
- Mit DRS Regeln können die sekundären VMs automatisch wieder auf ihre Seite gebracht werden

Ausfall der Witness

Wenn die Witness komplett ausfällt oder die Erreichbarkeit über Netz von der primären Seite und der sekundären Seite nicht mehr gegeben ist, hat dies erst einmal keinen direkten Einfluss auf das VSAN.

Was passiert:

- Die Witness hält nur Metadaten
- Beide Seiten haben die Mehrheit, kein Ausfall, alles geht wie gewohnt weiter
- Kommt jedoch jetzt ein weiter Ausfall (Disk, Host) gibt es Ausfall für die entsprechenden VMs, die Components auf diesen Geräten hatte (da die verbliebenen gespiegelten Components keine Mehrheit mehr haben.
- kommt die Witness zurück, werden die Metadaten aktualisiert und alles geht weiter
- ist die Witness nicht mehr wiederherstellbar, muss eine neue Witness ausgerollt und dem VSAN Cluster zugewiesen werden
- Die Witnessdaten werden wiederhergestellt und alles läuft weiter

Die Witness muss beide Seiten sehen, um in einen VSAN Cluster verbunden zu werden. Erst nach der Aussynchronisation kann eine Seite “fehlen”.

kompletter Netzausfall - alle Hosts isoliert

Wenn alle Hosts in einem VSAN Cluster untereinander isoliert sind, passiert folgends:

- Es entstehen für jeden Host “Single Node Cluster”, in dem der Host Master ist
- Der Host kann aber keinen Backup Host oder Agent Hosts finden, außerdem hat er für seine Components keine Mehrheit. Somit werden keine Schreibzugriffe mehr zugelassen
- HA findet in diesem Scenario keine Failover Hosts und kann nichts machen
- Die VMs auf dem Host werden zu “Zombie” VMs, da sie laufen, aber keine Schreiboperationen mehr ausführen können
- Falls VMs mit FTT=0 auf einem Host laufen und durch Zufall alle zugehörigen Components ebenfalls auf diesem Host laufen, wird diese VM weiterhin ausgeführt
- werden die Netzverbindungen wiederhergestellt, läuft alles nach einer Sync-Zeit weiter. Die Zombie VMs müssen meist resetet werden und es kann vorkommen, das der Original-VM Name im Inventory durch die interne ID ersetzt wurde.

Netzausfall zwischen allen Standorten

Wenn das Netz an allen Standorten noch funktioniert, aber zwischen den Standorten nichts mehr geht entsteht ei ähnliche Scenario wir “Alle Hosts sind isoliert”.

- Auf beiden Seiten wird jeweils eine eigene Netzwerk Partition erstellt
- Jede Seite wählt einen eigenen Master und Backup Host
- Jedoch können die Components keine Mehrheit gewinnen
- HA kann nicht auf die andere Seite Schalten
- Die Components werden für Schreibzugriffe gesperrt
- Es entstehen Zombie VMs
- Ausnahme: VMs mit FTT=0 laufen weiter, aber da der Standort keine Ausenanbindung hat, nutzt das wenig...
- Wird die Netzverbindung wiederhergestellt, wird alles aussynchronisiert und es geht weiter

Storage Policy mit PowerCLI ausrollen

```
$ds = Get-Datastore vsanDatastore
$sp = "thickProvisioned"
$vms = Get-VM -Datastore $ds
foreach ($vm in $vms) { $vm, (Get-HardDisk -VM $vm) | Set-
SpbmEntityConfiguration -StoragePolicy $sp }
```

Löschen einer Diskgroup

Um eine Diskgruppe zu löschen, muss die entsprechende Cachedisk gelöscht werden.

```
esxcli vsan storage remove --ssd=naa.xxxxxxx
```

Adv. Params

```
esxcli system settings advanced set -o /Net/Vmxnet3HwLRO -i 0
esxcli system settings advanced set -o /Net/UseHwTSO -i 0
esxcli system settings advanced set -o /Net/UseHwTSO6 -i 0
esxcli system settings advanced set -o /Net/TcpipDefLROEnabled -i 0

esxcfg-advcfg -s 2047 /LSOM/heapSize
esxcfg-advcfg -s 110000 /LSOM/diskIoTimeout
esxcfg-advcfg -s 4 /LSOM/diskIoRetryFactor
esxcfg-advcfg -s 512 /VSAN/DomClientheapsize
esxcfg-advcfg -s 48 /LSOM/lsomLogCongestionHighLimitGB

vsish
get /vmkModules/vsan/dom/MaxNumResyncCopyInFlight
Default: 50

vsish -e set /vmkModules/vsan/dom/MaxNumResyncCopyInFlight 25


esxcfg-advcfg -s 1 /VSAN/SwapThickProvisionDisabled
Value of SwapThickProvisionDisabled is 1
```

Cluster Shutdown

https://lifeofbrianoc.com/2017/09/19/shutdownpower-up-a-vsan-cluster-with-powercli/

https://www.isjw.uk/post/vmware/vsan/vsan-driver-not-certified-copy/

https://4sysops.com/archives/startup-and-shutdown-a-vmware-cluster-with-powercli-and-powershell/