



Proof of Concept

Starwind vSAN

Installation, configuration, testing

De Vos Brent

Academiejaar 2020-2021

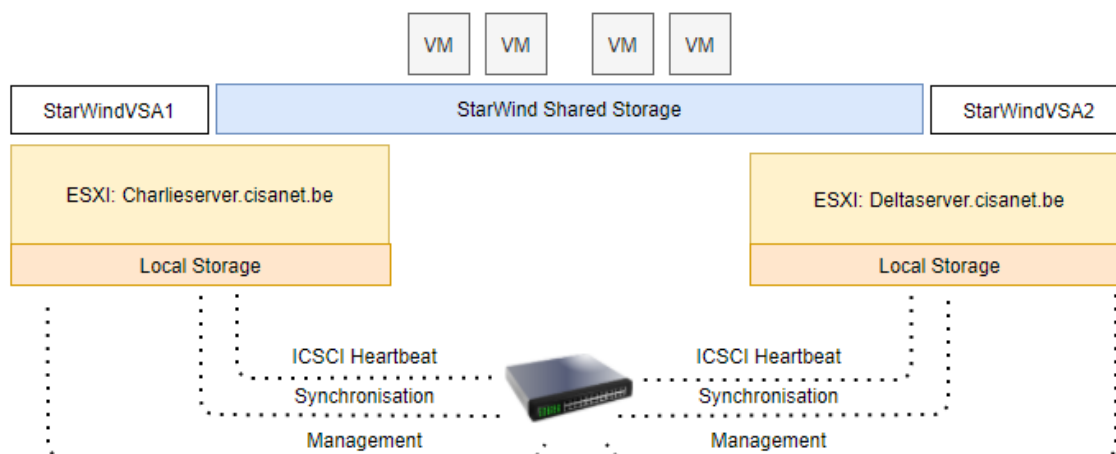
Voorwoord

In dit document staan alle installatiestappen beschreven hoe ik de proof of concept heb opgebouwd in de vSphere cluster van CISA.

Ook de testresultaten zullen hierin te vinden zijn.

Gevoelige informatie zoals belangrijke ip adressen en hostnames zijn verborgen voor veiligheidsredenen.




Architectuur 2-node setup



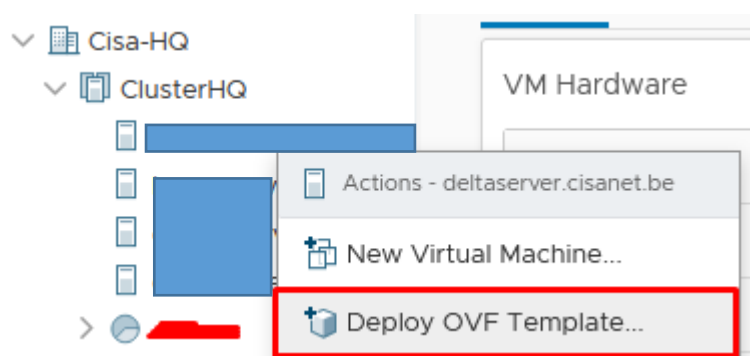
Prerequisites per ESXI server:

4x	Virtual 1.7 Ghz processor
1x	4GB RAM (If using cache -> min. 8GB)
1x	Network interface for Synchronization traffic
1x	Network interface for ISCSI / Heartbeat traffic
1x	Network interface for management
1x	Hypervisor with vSphere 5.5 or later for each server
1x	20 GB Virtual disk for Starwind OS
1x	Virtual disks on separate partition (Thick provisioned Eager Zeroed) To create shared storage pool
1x	Starwind vSAN license for 3 nodes
1x	Windows server (vm) to run Starwind management console

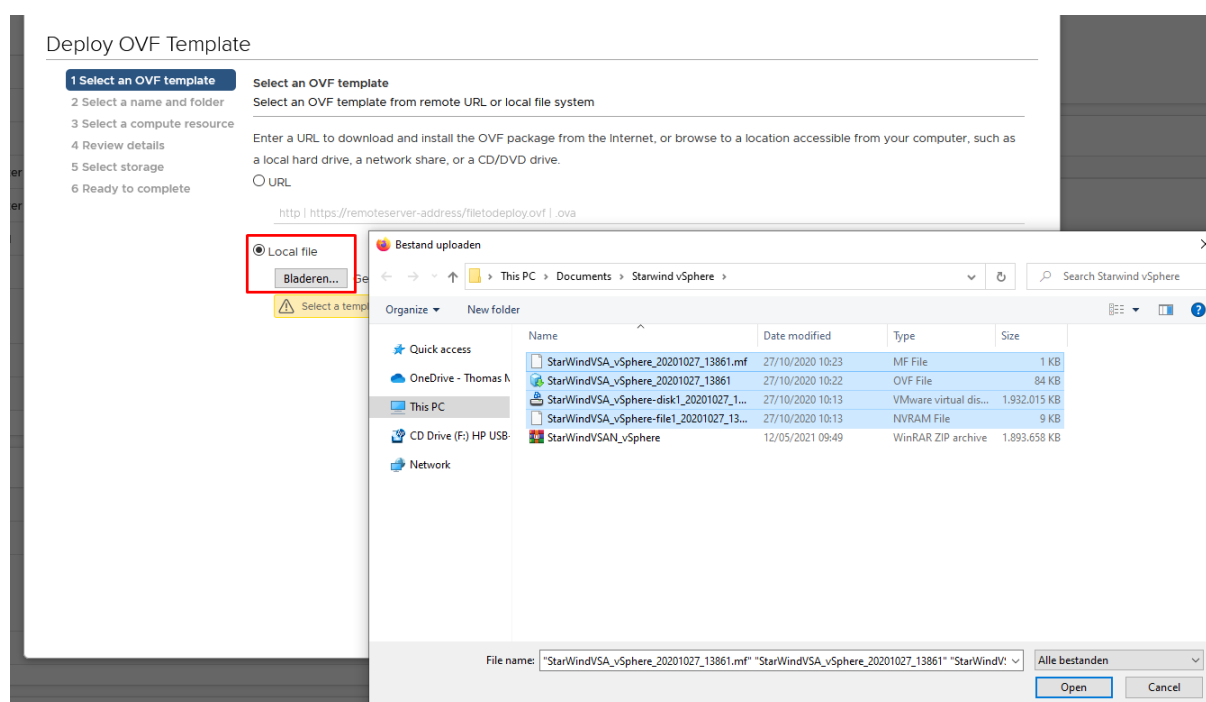
Download OVA files: [Download](#)

	StarWindVSA_vSphere_20201027_13861.mf	27/10/2020 10:23	MF File	1 KB
	StarWindVSA_vSphere_20201027_13861	27/10/2020 10:22	OVF File	84 KB
	StarWindVSA_vSphere-disk1_20201027_13861	27/10/2020 10:13	VMware virtual dis...	1.932.015 KB
	StarWindVSA_vSphere-file1_20201027_13861.nvram	27/10/2020 10:13	NVRAM File	9 KB
	StarWindVSAN_vSphere	12/05/2021 09:49	WinRAR ZIP archive	1.893.658 KB

Selecteer een ESXI host > right klik > Deploy OVG Template



Kies voor local File > Selecteer de 4 installatiebestanden



Kies een ESXI server waar de VSA op gedeployed mag worden
!! elke VSA op een andere ESXI host !!

Deploy OVF Template

- ✓ 1 Select an OVF template
- ✓ 2 Select a name and folder
- 3 Select a compute resource**
- 4 Review details
- 5 Select storage
- 6 Ready to complete

Select a compute resource

Select the destination compute resource for this operation

Cisa-HQ
ClusterHQ
charlieserver.cisanet.be
deltaserver.cisanet.be

Compatibility

✓ Compatibility checks succeeded.

CANCEL

BACK

NEXT

Kies een opslaglocatie

Deploy OVF Template

- ✓ 1 Select an OVF template
- ✓ 2 Select a name and folder
- ✓ 3 Select a compute resource
- ✓ 4 Review details
- ✓ 5 License agreements
- 6 Select storage**
- 7 Select networks
- 8 Ready to complete

Select storage

Select the storage for the configuration and disk files

☐ Encrypt this virtual machine (Requires Key Management Server)

Select virtual disk format:

Thick Provision Lazy Zeroed

VM Storage Policy:

Datastore Default

Name	Capacity	Provisioned	Free	Type	Cluster
Deltaserver-Datastore02	764 TB	4.99 TB	2.65 TB	VMFS 6	
deltaserver-datastore1	129.25 GB	1.41 GB	127.84 GB	VMFS 6	

Compatibility

✓ Compatibility checks succeeded.

CANCEL

BACK

NEXT

Stel de juiste networks in

Deploy OVF Template

- ✓ 1 Select an OVF template
- ✓ 2 Select a name and folder
- ✓ 3 Select a compute resource
- ✓ 4 Review details
- ✓ 5 License agreements
- ✓ 6 Select storage
- 7 Select networks**
- 8 Ready to complete

Select networks

Select a destination network for each source network.

Source Network	Destination Network
ISCSI	ISCSI_Network_1
Management	VI-Management
Sync	ISCSI_Network_2

IP Allocation Settings

IP allocation:

Static - Manual

IP protocol:

IPv4

Boot beide vm's

Connect naar vm via remote console

vmstorage1 - VMware Workstation 16 Player (Non-commercial use only)

Player |   

```
StarWind Virtual Storage Appliance
Codename: Silent Assassin

Web console: [redacted]

vmstorage1 login: _
```

Login: [redacted]

Passwd: [redacted]

Vervolgens gaan we de netwerk config instellen met volgende ip adressen:

Starwind1: [redacted]

Starwind2: [redacted]

Default Gateway: [redacted]

Subnet mask: [redacted]

In de remote console gebruik je commando 'nmtui'

```
StarWind Virtual Storage Appliance
Codename: Silent Assassin

Web console: [redacted]

vmstorage1 login: [redacted]
Password:
Last login: Wed May 12 05:42:41 on tty1
[root@vmstorage1 ~]# nmtui_
```

Edit a connection



Selecteer ens224 > NIC1 > Edit



Geef de juiste ip configuratie in & enable automatically connect:

Profile name **ens224**
Device **ens224 (00:50:56:A4:DF:66)**

= ETHERNET <Show>

IPv4 CONFIGURATION <Manual> <Hide>

Addresses <Remove>

<Add...>

Gateway

DNS servers <Add...>

Search domains <Add...>

Routing (No custom routes) <Edit...>

☐ Never use this network for default route

☐ Ignore automatically obtained routes

☐ Ignore automatically obtained DNS parameters

☐ Require IPv4 addressing for this connection

= IPv6 CONFIGURATION <Automatic> <Show>

☒ Automatically connect

☐ Available to all users

<Cancel> **<OK>**

Doe hetzelfde voor de andere netwerk interface:

| Edit Connection |

Profile name **ens256**
Device **ens256 (00:50:56:A4:DF:71)**

= ETHERNET <Show>

IPv4 CONFIGURATION <Manual> <Hide>

Addresses <Remove>

<Add...>

Gateway

DNS servers <Add...>

Search domains <Add...>

Routing (No custom routes) <Edit...>

☐ Never use this network for default route

☐ Ignore automatically obtained routes

☐ Ignore automatically obtained DNS parameters

☐ Require IPv4 addressing for this connection

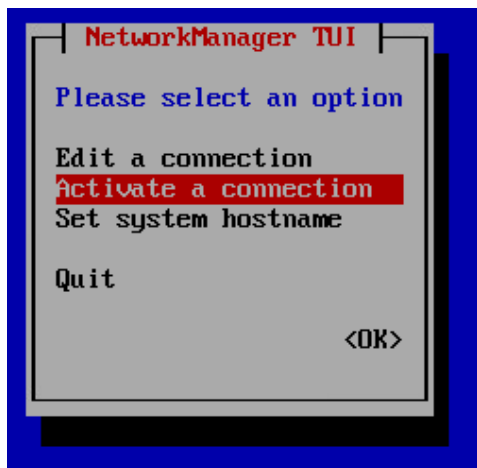
= IPv6 CONFIGURATION <Automatic> <Show>

☒ Automatically connect

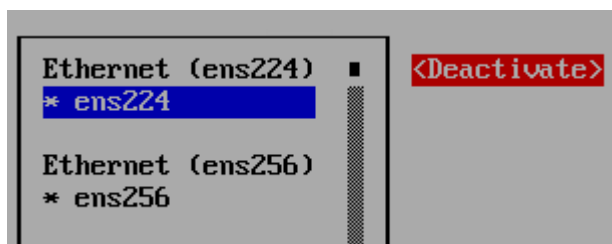
☐ Available to all users

<Cancel> **<OK>**

Ga terug naar het hoofdmenu en kies voor 'Activate a connection'



Activeer beide netwerk interfaces



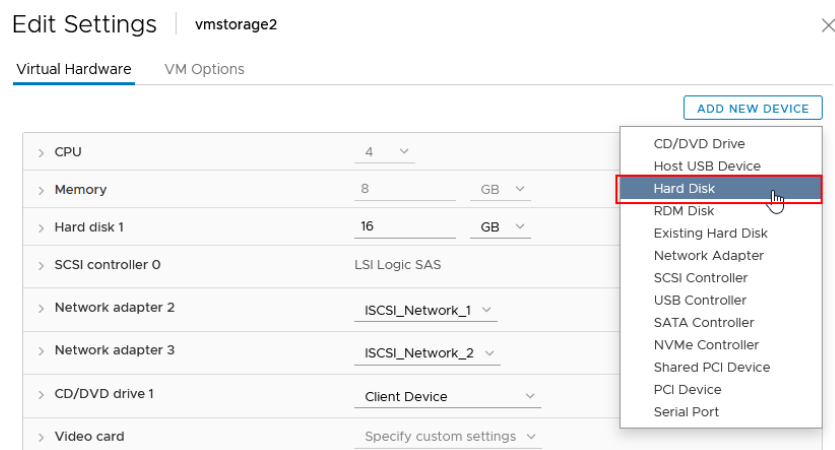
Reboot de server met het commando 'reboot'

```
[root@vmstorage1 ~]# reboot_
```

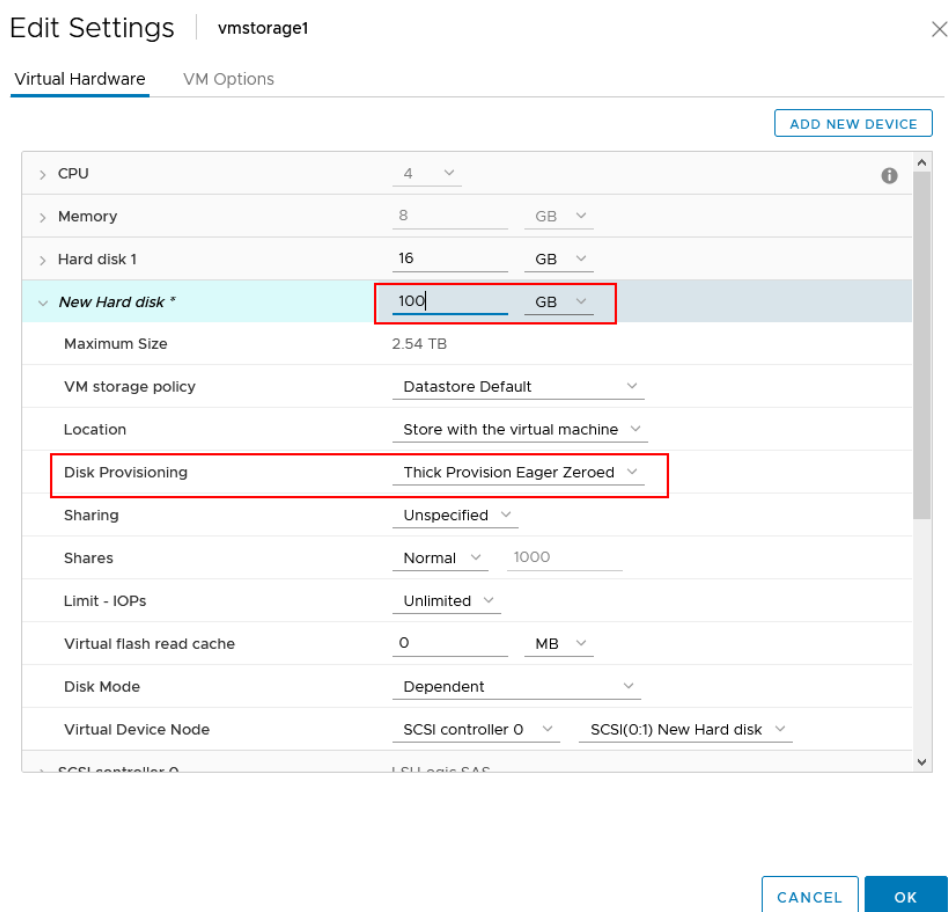
Herhaal deze stappen voor de 2^{de} Starwind VM

Bij beide StarWind machines voegen we een Hard Disk toe van x aantal GB
Deze zullen we gebruiken om shared storage te creëren voor virtuele servers

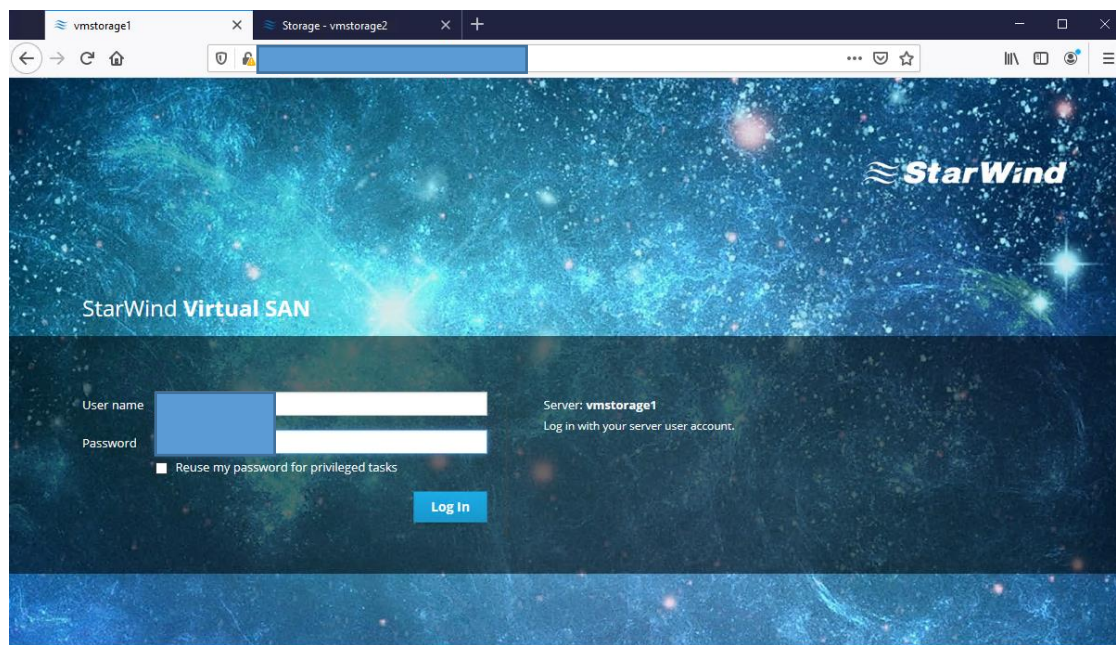
Right klik op VM > Edit settings > Add new device > Hard disk



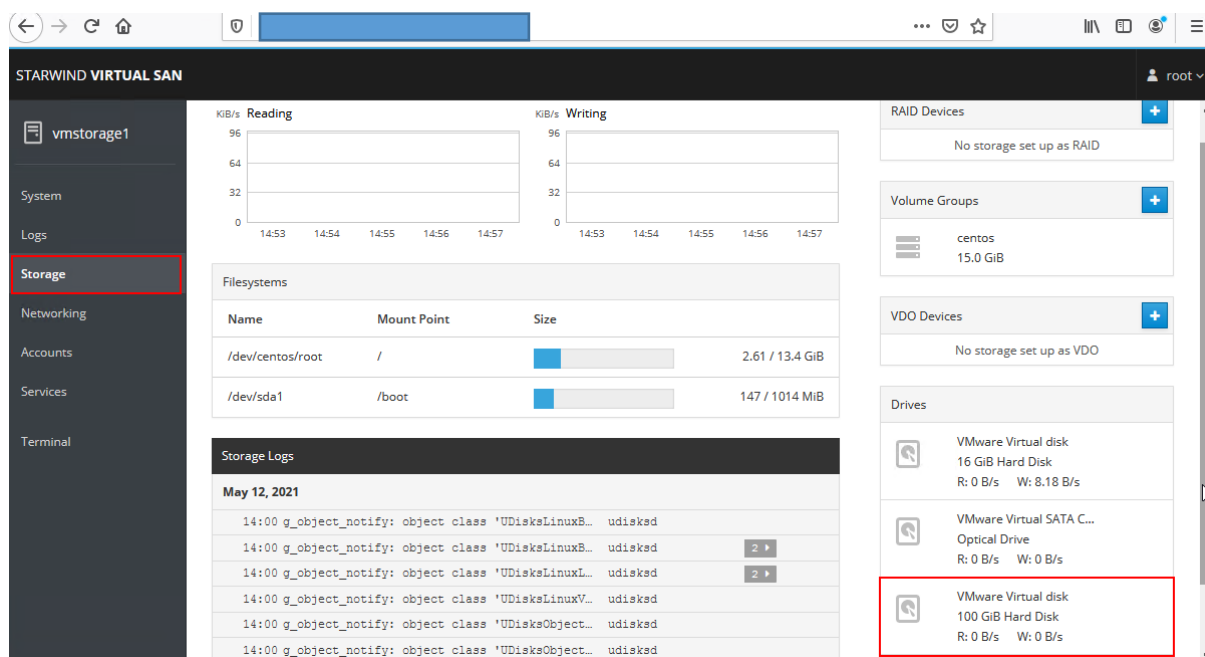
Belangrijk is dat je de provisioning insteld op 'Thick Provision Eager Zeroed'



Dankzij onze eerdere netwerk config kunnen we de VSA verder instellen via de web GUI
Hiernaar kan je connecteren via <https://ip-adress:9090>



In het tabblad Storage zien we onze pas gecreëerde disk staan



Deze disk dient geformatteerd te worden vooraleer deze gebruikt kan worden

klik op de Disk > Format en gebruik volgende settings:

Format /dev/sdb

Erase

Don't overwrite existing data

Type

XFS - Recommended default

Name

disk1

☐ Encrypt data

Mounting

Custom

Mount Point

/mnt/disk1

Mount Options

☒ Mount at boot
☐ Mount read only
☐ Custom mount options

Formatting a storage device will erase all data on it.

Cancel

Format

Format

En vervolgens Mount je de drive

Content

Create Partition Table

100 GiB xfs File System

/dev/sdb

Filesystem

Format

Name

disk1

Mount Point

/mnt/disk1

Mount

Mount Options

defaults

Used

-

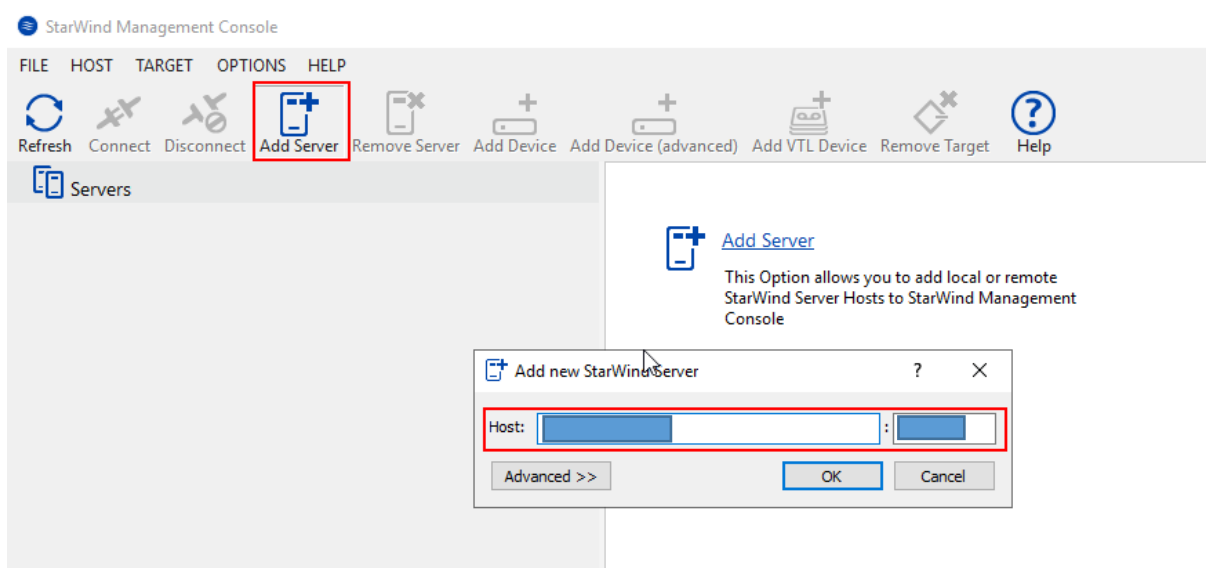
Deze acties herhaal je ook voor de andere Starwind machine.

Best practices zeggen dat je daar best dezelfde naamgevingen gebruikt qua disken

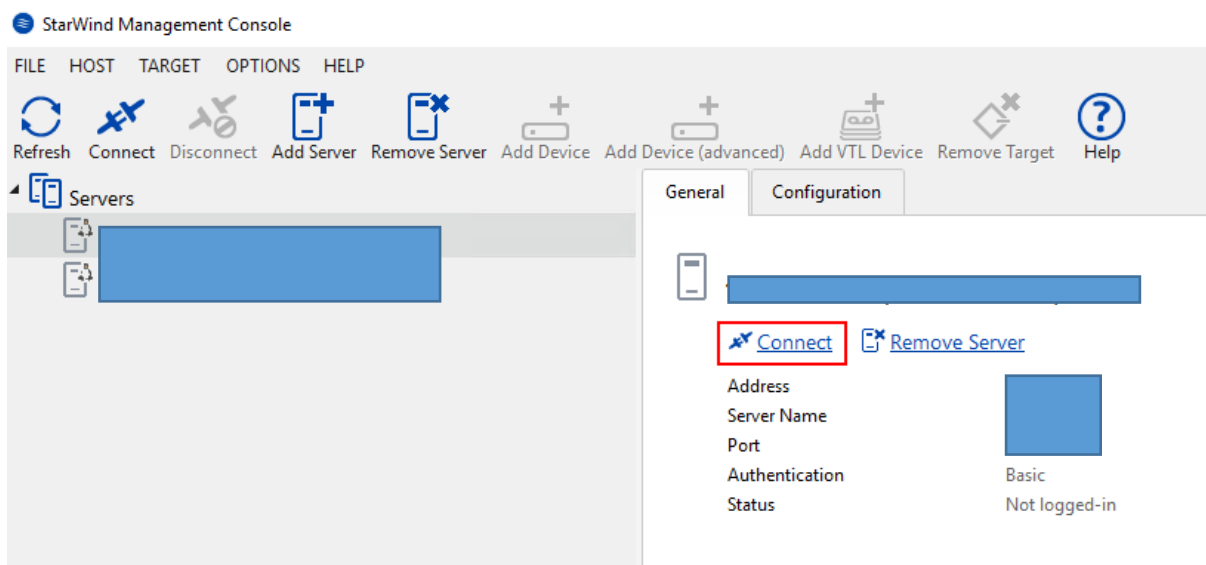
Om nu verdere configuratie te doen zoals high availability, replication, ... moet je StarWind Management console gebruiken

De eerste stap is om de 2 Starwind vm's te connecteren met de management console

Add server > Host IP > OK

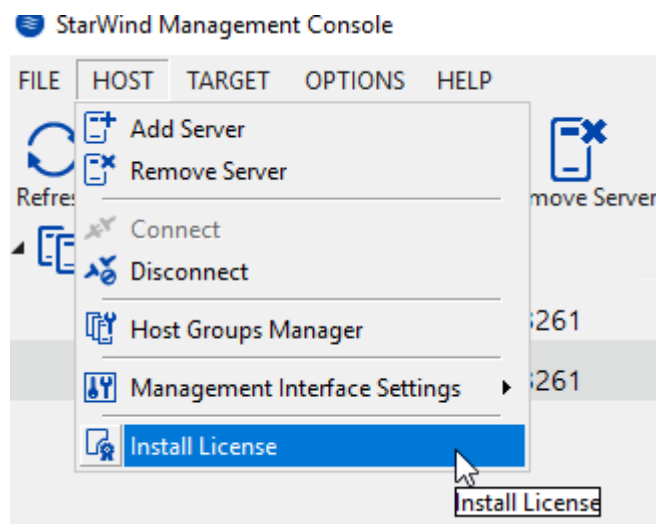


Connect beide servers (de melding van storage pool mag je gewoon sluiten)

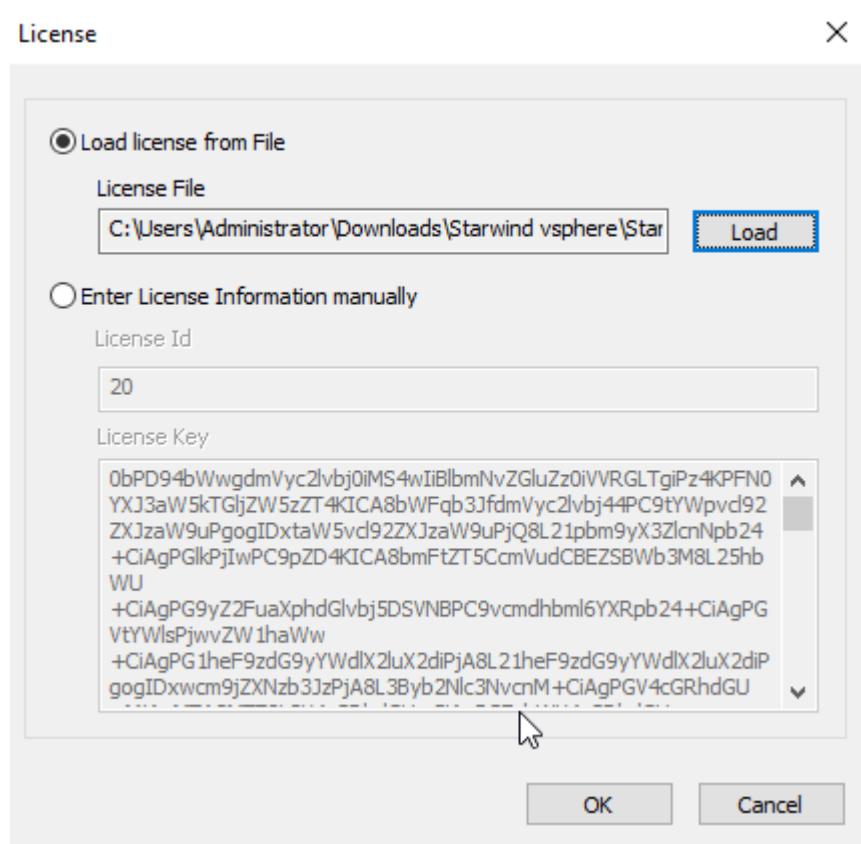


Vervolgens kennen we een licentie toe aan beide servers

Selecteer de server > HOST > Install License

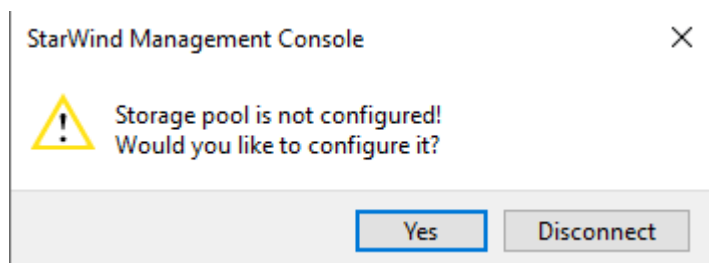


En laad de juiste licentie file in:

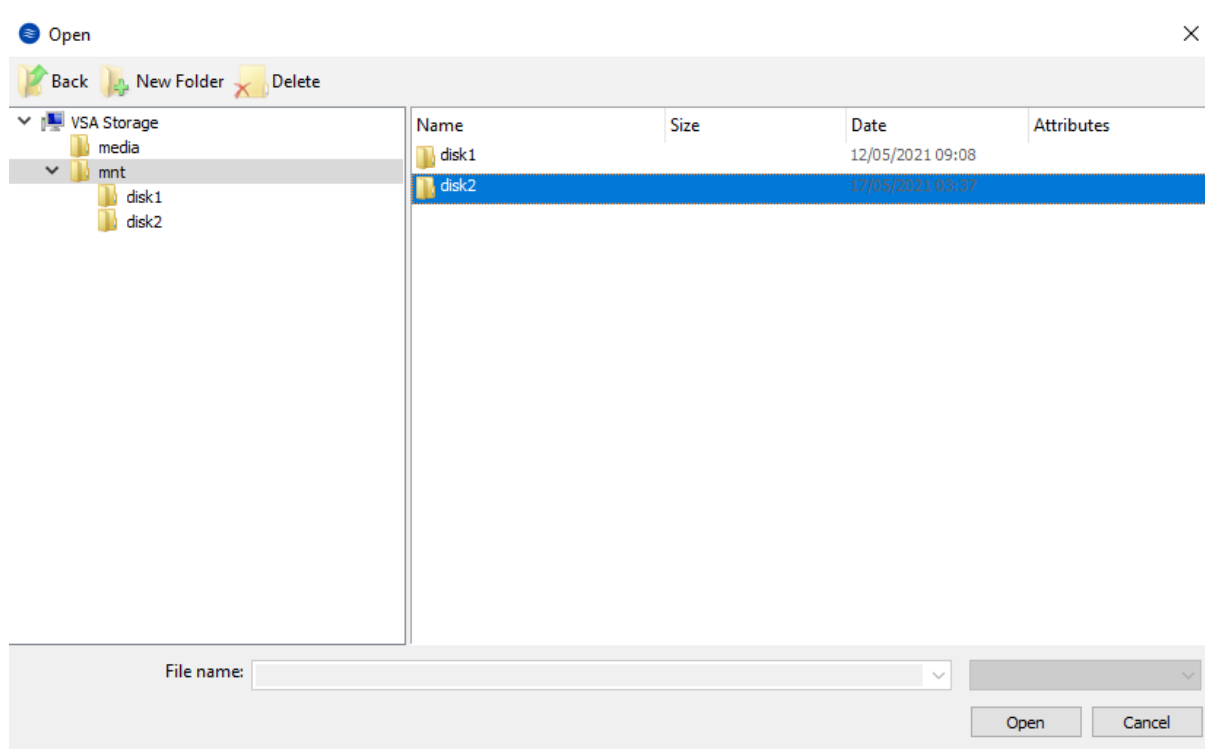


Nu alle servers de nodige licentie hebben gekregen kunnen we verder met de configuratie

Wanneer je opnieuw naar de servers connecteerd krijg je terug volgende melding:

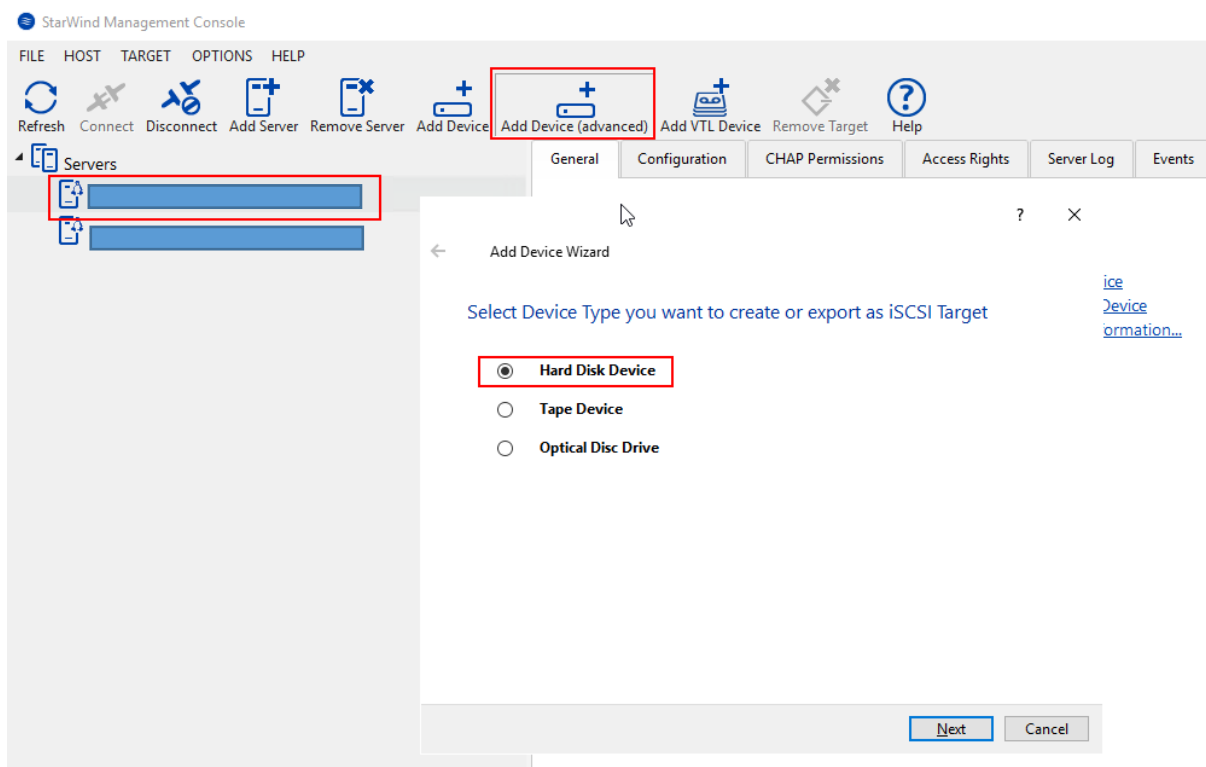


Kies voor 'Yes' en selecteer de eerder aangemaakte gemounte disk:

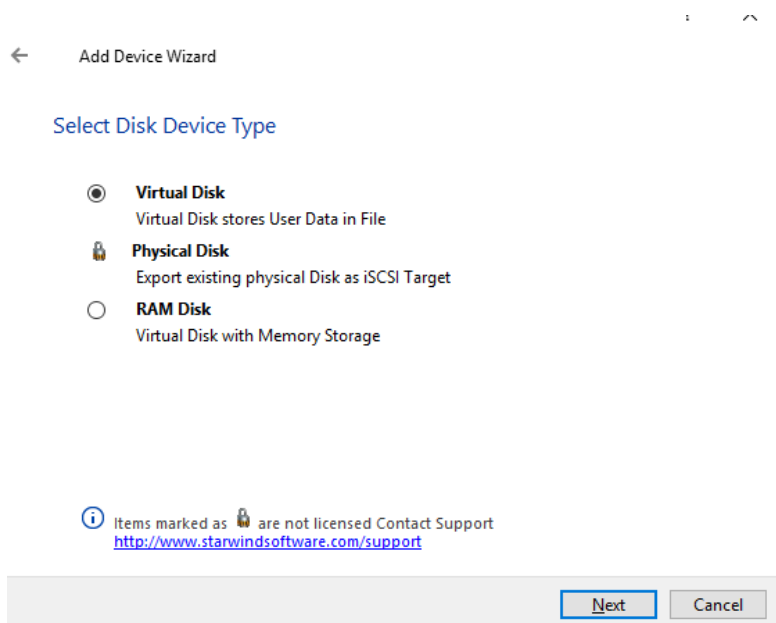


Aanmaken van Starwind devices:

Selecteer een server > Add device (advanced) > Hard Disk Device



Kies voor Virtual Disk



Kies een gepaste naam en grootte van de Virtual Disk

? X

← Add Device Wizard

Virtual Disk Location

☒ Create a New Virtual Disk

Name:	<input type="text" value="HA1"/>
Location:	<input type="text" value="VSA Storage\mnt\disk2\"/> ...
Size:	<input type="text" value="740"/> GB

☐ Use an Existing Virtual Disk

Location:	<input type="text"/> ...
<input type="checkbox"/> Read-Only Mode	

Next

Cancel

Kies Thick-provisioned

? X

← Add Device Wizard

Virtual Disk Options

☒ Thick-provisioned

☐ Thick-provisioned with Log-structured Write-Back Cache (experimental)

☐ LSFS

☐ Deduplication

StarPack Cache Size: MB

Block Size

☒ Use 512 bytes sector size

☐ Use 4096 bytes sector size. May be incompatible with some clients

Next

Cancel

Default settings zijn voldoende voor RAM Cache Parameters

← Add Device Wizard

Specify Device RAM Cache Parameters

Mode

☐ **Write-Back**
Writes are performed asynchronously, actual Writes to Disk are delayed, Reads are cached

☐ **Write-Through**
Writes are performed synchronously, Reads are cached

☒ **N/A**
Reads and Writes are not cached

☐ Set Maximum available Size

Size:

Next

Cancel

? X

← Add Device Wizard

Specify Flash Cache Parameters

☒ No Flash Cache

☐ Use Flash Cache

Name:

Location: ...

Size:

Next

Cancel

geef een gepaste naam voor het iSCSI target

? X

← Add Device Wizard

Target Parameters

Choose a Target Attachment Method

Create new Target

Target Alias

HA1

☐ Target Name

☒ Allow multiple concurrent iSCSI Connections

Next Cancel

Create het target:

? X

← Add Device Wizard

Device Creation completed

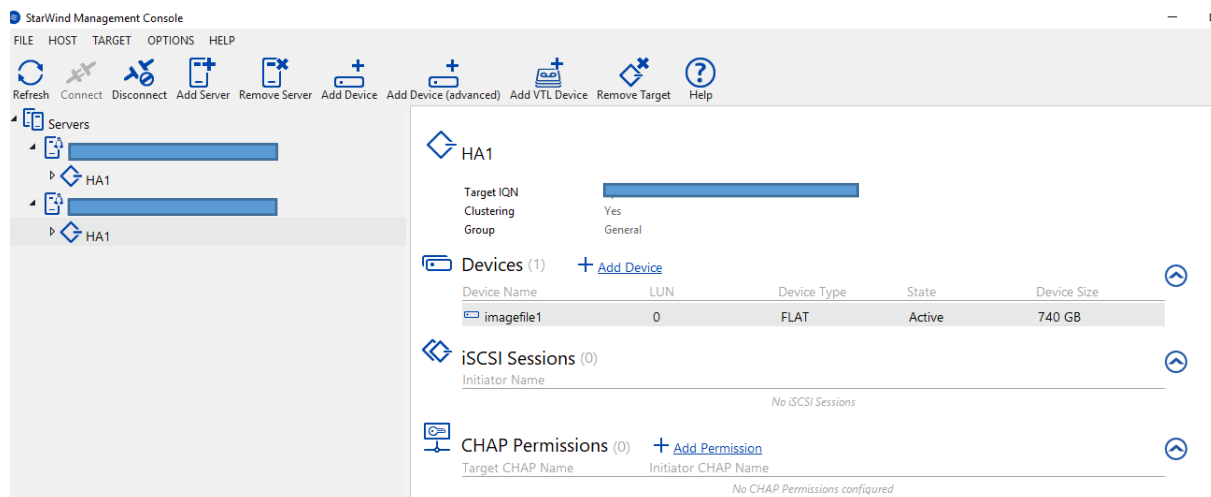
Following Device was created

imagefile1

Target Name:

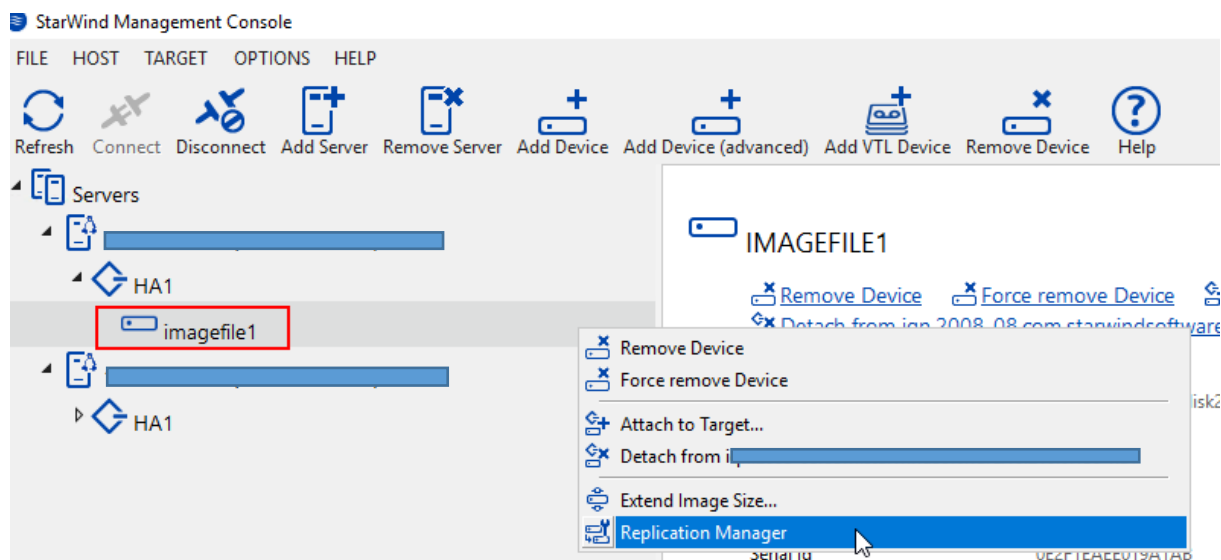
Close Cancel

Vorige stappen zouden volgend resultaat moeten geven:

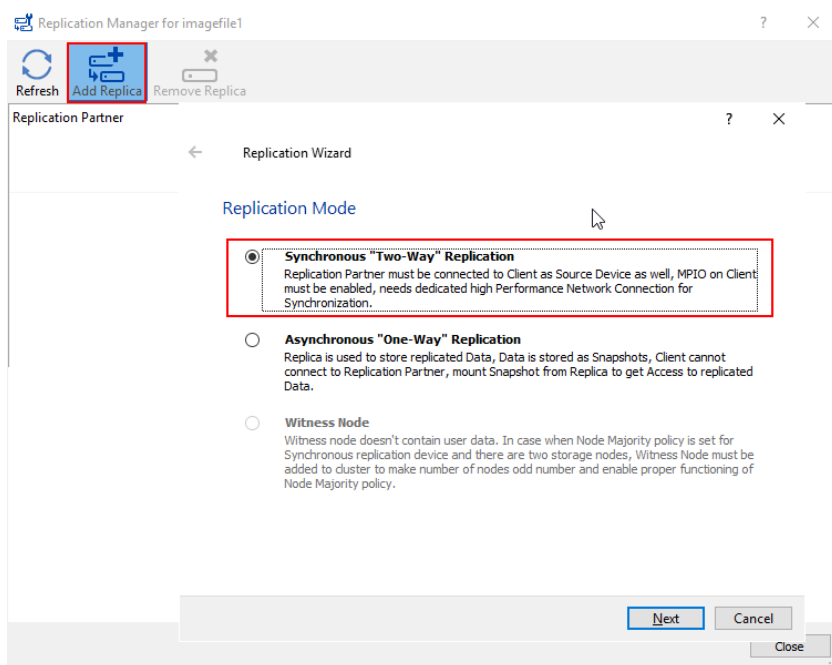


Volgende stap is om 2-way synchronous replication in te stellen. Dit zorgt ervoor dat alles gerepliceerd wordt naar de 2^{de} node & andersom; Zo creëren we high available shared storage.

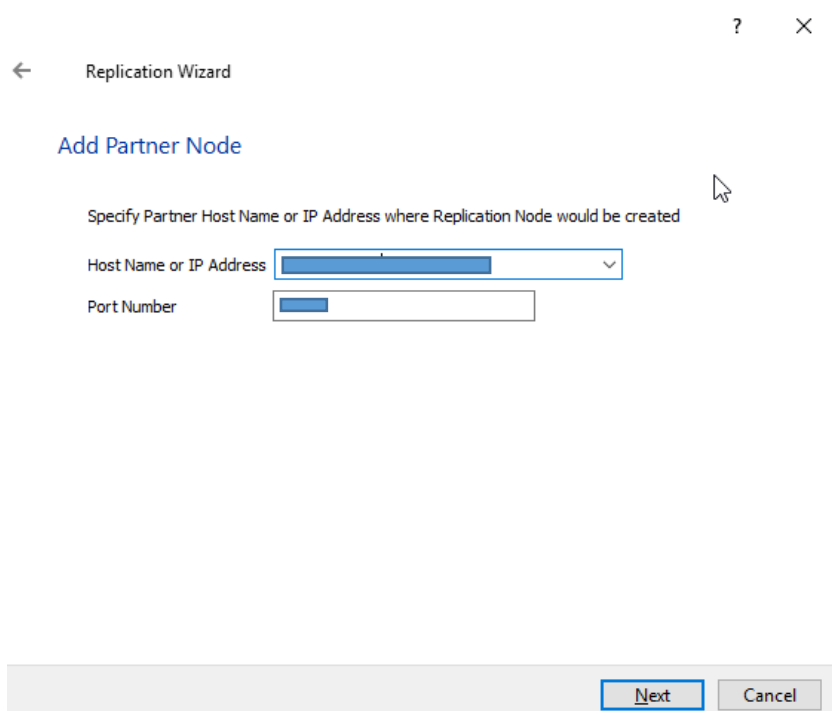
Open rechts het net aangemaakte device > rechterklik op imagefile > Replication Manager



Klik op Add Replica > Kies voor Synchronous “Two-Way” Replication



Geef het ip adres van de 2^{de} node op



Vervolgens krijgen we de Failover Strategy keuze. Even wat uitleg bij beide opties:

Heartbeat

The Heartbeat failover strategy allows avoiding the “split-brain” scenario when the HA cluster nodes are unable to synchronize but continue to accept write commands from the initiators independently. It can occur when all synchronization and heartbeat channels disconnect simultaneously, and the partner nodes do not respond to the node’s requests. As a result, StarWind service assumes the partner nodes to be offline and continues operations on a single-node mode using data written to it.

If at least one heartbeat link is online, StarWind services can communicate with each other via this link. The device with the lowest priority will be marked as not synchronized and get subsequently blocked for the further read and write operations until the synchronization channel resumption. At the same time, the partner device on the synchronized node flushes data from the cache to the disk to preserve data integrity in case the node goes down unexpectedly. It is recommended to assign more independent heartbeat channels during the replica creation to improve system stability and avoid the “split-brain” issue.

With the heartbeat failover strategy, the storage cluster will continue working with only one StarWind node available.

Node Majority

The Node Majority failover strategy ensures the synchronization connection without any additional heartbeat links. The failure-handling process occurs when the node has detected the absence of the connection with the partner. The main requirement for keeping the node operational is an active connection with more than half of the HA device’s nodes. Calculation of the available partners is based on their “votes”.

In case of a two-node HA storage, all nodes will be disconnected if there is a problem on the node itself, or in communication between them. Therefore, the Node Majority failover strategy requires the addition of the third Witness node or file share (SMB) which participates in the nodes count for the majority, but neither contains data on it nor is involved in processing clients’ requests. In case an HA device is replicated between 3 nodes, no Witness node is required.

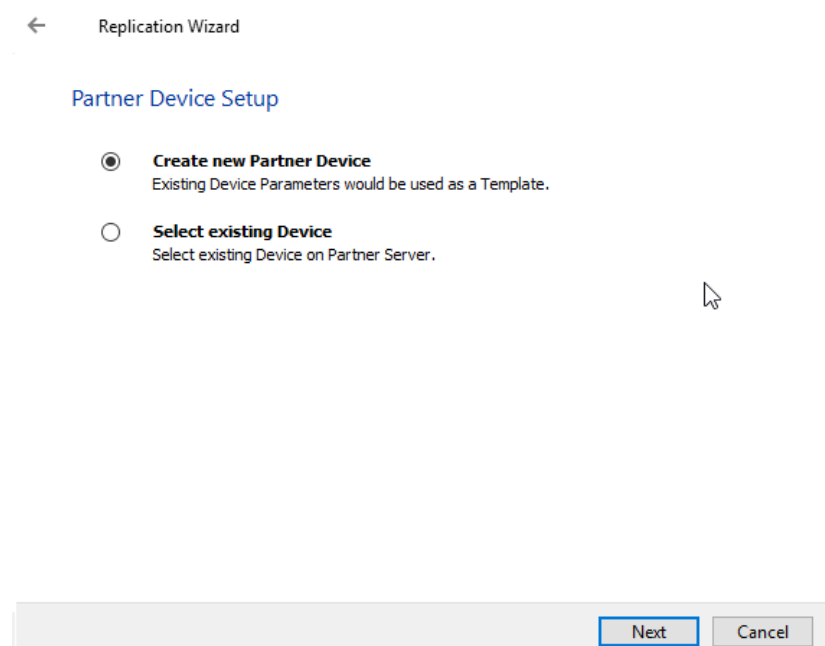
With Node Majority failover strategy, failure of only one node can be tolerated. If two nodes fail, the third node will also become unavailable to clients’ requests.

Please select the required option:

Bron: <https://www.starwindsoftware.com/resource-library/starwind-virtual-san-for-vsphere/>

Voor onze proof of concept kiezen we voor 'Node majority' failover strategy. Deze werkt met de hulp van een witness

Kies voor 'Create new Partner Device'



← Replication Wizard

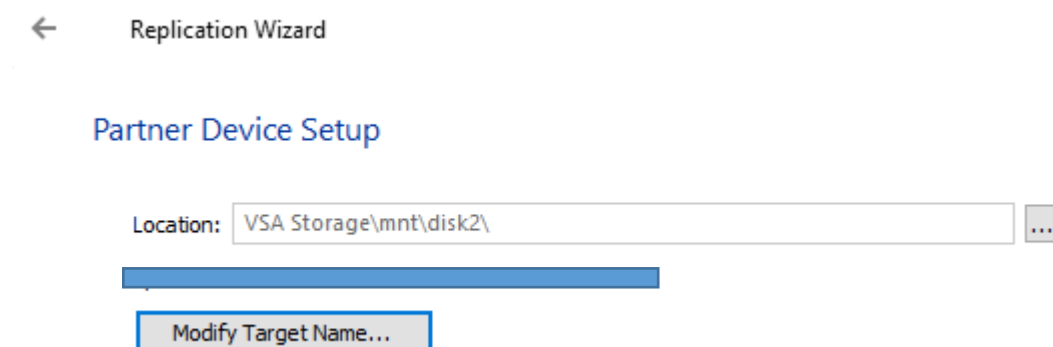
Partner Device Setup

☒ **Create new Partner Device**
Existing Device Parameters would be used as a Template.

☐ **Select existing Device**
Select existing Device on Partner Server.

Next Cancel

Default Partner Device setup is normal voldoende:



← Replication Wizard

Partner Device Setup

Location: VSA Storage\mnt\disk2\ ...

Modify Target Name...

Kijk zeker na of het target effectief de 2 de node is. Dit kan je zien aan het ip adres in de Target Name.

Change Network Settings:

Replication Partner

? X



Replication Wizard

Network Options for Replication

Networks for Synchronization and Heartbeat

Press "Change Network Settings..." to configure Interfaces

Networks for Heartbeat

Press "Change Network Settings..." to configure Interfaces

Change Network Settings

Duid vervolgens aan welke NIC's welke functie hebben:
Duid zeker ook 'Allow Free Select Interfaces' aan

Specify Interfaces for Synchronization Channels X

Select synchronization channel

Interfaces	Networks	Synchronization and Heartbeat	Heartbeat
[-] Host Name: []			
[]		<input checked="" type="checkbox"/>	<input type="checkbox"/>
[]		<input type="checkbox"/>	<input checked="" type="checkbox"/>
[-] Host Name: []			
[]		<input checked="" type="checkbox"/>	<input type="checkbox"/>
[]		<input type="checkbox"/>	<input checked="" type="checkbox"/>

☒ Allow Free Select Interfaces
 OK Cancel

Kies voor: Synchronize from existing Device

Replication Partner



Replication Wizard

Select Partner Device Initialization Mode

- ☒ **Synchronize from existing Device**
All Data from existing Device would be copied to new Device.
- ☐ **Do not Synchronize**
Data on HA Nodes remains unchanged.

Klik in het volgende scherm op create replica & het replication device op de 2^{de} node wordt aangemaakt:

? X



Replication Wizard

Device Creation completed

Following Device was created



HAImage 1

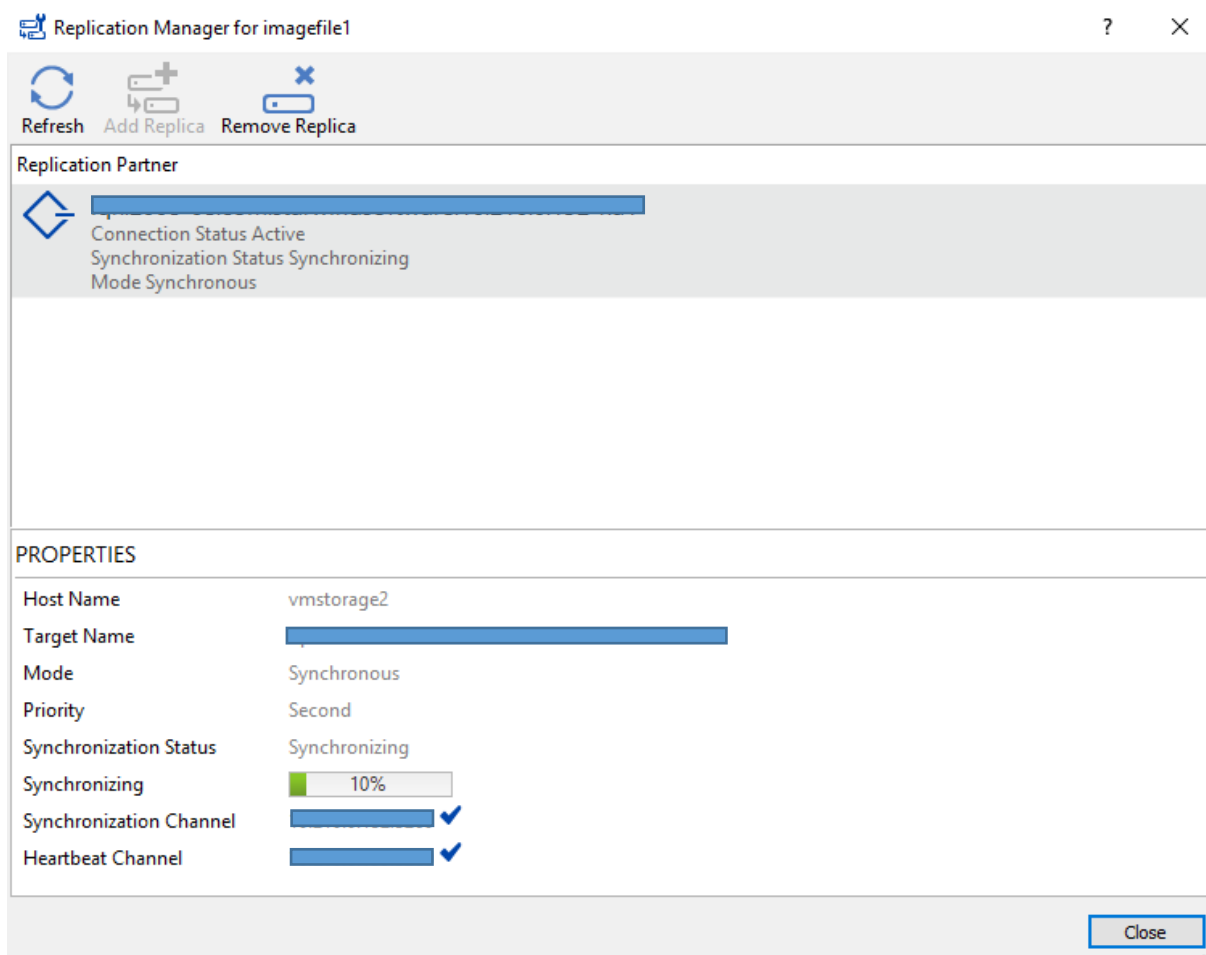
Target Name:



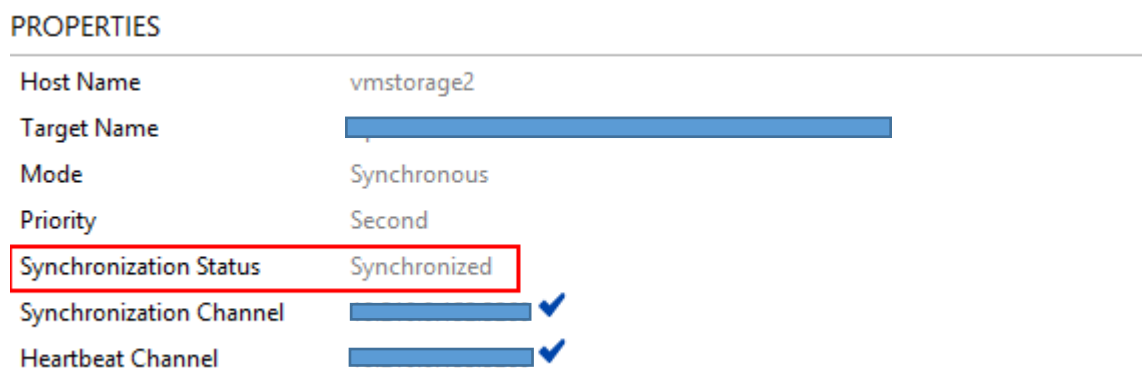
Close

Cancel

In de Replication Manager krijgen we te zien dat beide virtual storage devices aan het Synchroniseren zijn. Dit kan een lange tijd duren afhankelijk van de opslag grootte:



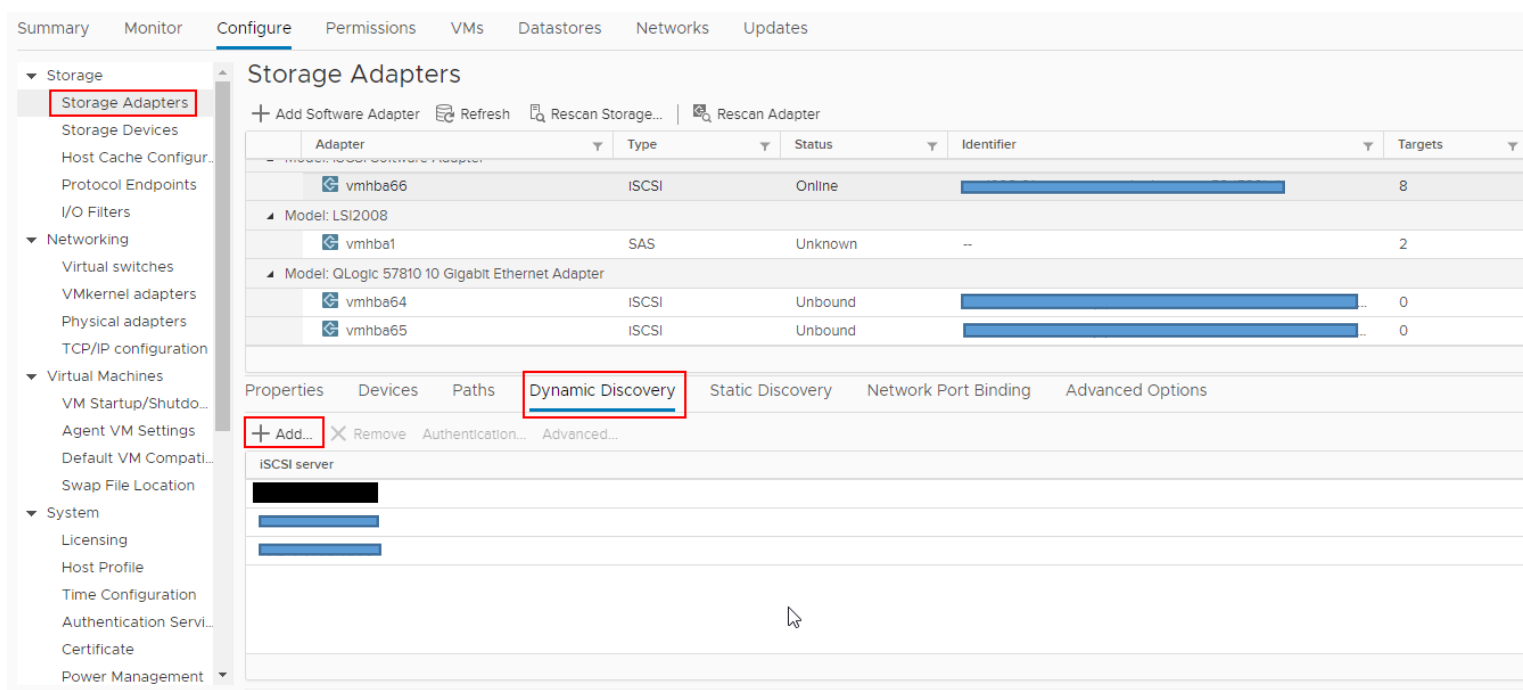
Na Synchroniseren zal de status op Synchronised staan:



Nu de HA storage pool ingesteld is moeten we deze aan de vsphere omgeving toevoegen. Aan elke ESXI host in vSphere moeten we de ip adressen van de ISCSI NIC's van beide nodes toevoegen. Dit gaat als volgt:

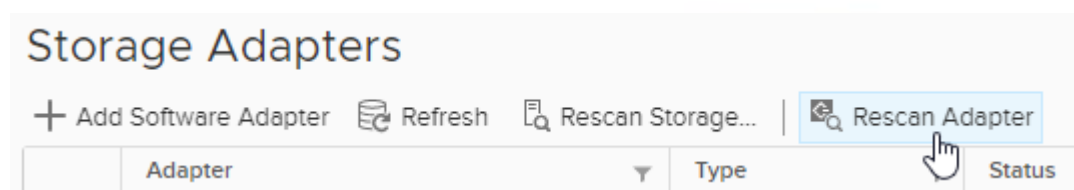
Rechterklik op ESXI host > Settings

Bij Storage Adapters > Dynamic Discovery voegen we de ip adressen van de Starwind nodes toe:



Adapter	Type	Status	Identifier	Targets
vmhba66	ISCSI	Online		8
Model: LSI2008				
vmhba1	SAS	Unknown	--	2
Model: QLogic 57810 10 Gigabit Ethernet Adapter				
vmhba64	ISCSI	Unbound		0
vmhba65	ISCSI	Unbound		0

Vervolgens is een Rescan nodig van de adapter:



Herhaal deze stappen voor alle betrokken ESXI servers

Als we na de rescan bij de Devices van de iSCSI software adapter gaan kijken zien we onze

Storage Adapters

+ Add Software Adapter Refresh Rescan Storage... Rescan Adapter

Adapter	Type	Status	Identifier	Targets	Devices	Paths
Model: HPE P408i-a SR Gen10						
vmhba0	SAS	Unknown	--	3	3	3
Model: iSCSI Software Adapter						
vmhba66	iSCSI	Online		8	3	12
Model: LSI2008						
vmhba1	SAS	Unknown	--	2	9	16

Copy All 5 items

Properties Devices Paths Dynamic Discovery Static Discovery Network Port Binding Advanced Options

Refresh Attach Detach Rename...

Name	LUN	Type	Capacity	Datastore	Operational State	Hardware Acceleration	Drive Type	Transport
STARWIND iSCSI Disk (eul.0e2f1e4ee019a1eb)	0	disk	740.00 GB	Not Consumed	Attached	Supported	HDD	iSCSI

Copy All 3 items

HA Starwind storage in de Datastore lijst staan

Deze Datastore moeten we nog instellen als een nieuwe VMFS Datastore op de betrokken ESXI servers

Rechterklik op betrokken ESXI host > Storage > New Datastore

Kies voor een VMFS volume

New Datastore

1 Type

2 Name and device selection

3 VMFS version

4 Partition configuration

5 Ready to complete

Type

Specify datastore type.



VMFS

Create a VMFS datastore on a disk/LUN.



NFS

Create an NFS datastore on an NFS share over the network.



VVOL

Create a Virtual Volumes datastore on a storage container connected to a storage provider.

Selecteer het Starwind device en geef het een gepaste naam

New Datastore

✓ 1 Type

2 Name and device selection

3 VMFS version

4 Partition configuration

5 Ready to complete

Name and device selection
Select a name and a disk/LUN for provisioning the datastore.

Datastore name:

Name	LUN	Capacity	Hardware...	Drive T...	S
STARWIND ISCSI Disk (eu...	0	740.00 GB	Supported	HDD	5

Kies een VMFS version die je wilt gebruiken

New Datastore

✓ 1 Type

✓ 2 Name and device selection

3 VMFS version

4 Partition configuration

5 Ready to complete

VMFS version
Specify the VMFS version for the datastore.

☒ VMFS 6
VMFS 6 enables advanced format (512e) and automatic space reclamation support.

☐ VMFS 5
VMFS 5 enables 2+TB LUN support.

Default settings zijn hier voldoende

New Datastore

✓ 1 Type

✓ 2 Name and device selection

✓ 3 VMFS version

4 Partition configuration

5 Ready to complete

Partition configuration
Review the disk layout and specify partition configuration details.

Partition Configuration

Datastore Size GB

Block size

Space Reclamation Granularity

Space Reclamation Priority

Empty: 740.0 GB

Next > Finish

Testing 2-Node setup

Dankzij vorige stappen uit te voeren hebben we succesvol een 2-node setup gebouwd met virtual shared storage.

Beide ESXI hebben toegang tot de shared storage pool waar we VM's op kunnen draaien. Met deze opstelling hebben we geprobeerd high available storage op te zetten. Deze hebben we aan volgende testen onderworpen:

VM Migration (compute resource) tussen beide nodes:

VM wordt snel en met succes gemigreerd tussen de nodes

Uitval van 1 VSA:

De opstelling kan de uitval van 1 VSA permiteren. De VM's op de virtuele storage blijven draaien

Failover:

Failover wordt mee afgehandelt door de vCenter HA configuratie. Indien een ESXI server zal uitvallen waar de compute resource draait zal deze automatisch snel herstart worden op de andere beschikbare ESXI servers.

Uitval 2 VSA's:

Bij een uitval van beide VSA's vallen VM's op de virtuele shared storage uit. Deze kunnen niet meer geboot of gemigreerd worden

LSFS bestandssysteem + Deduplicatie:

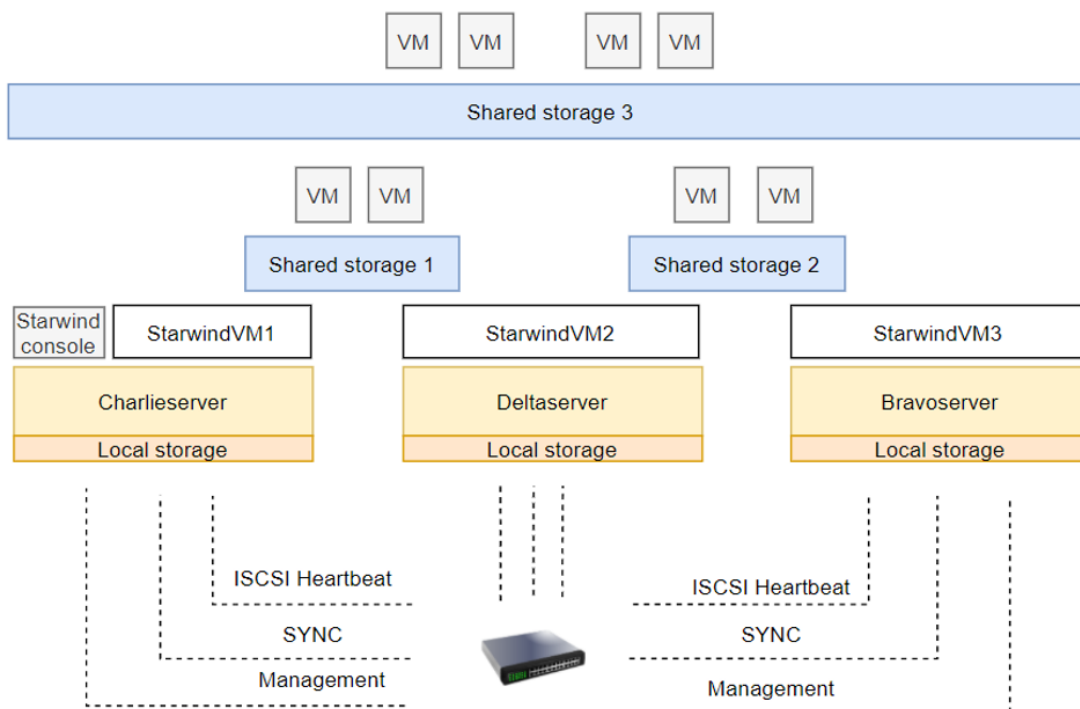
Bij het willen instellen van dit bestandssysteem met bijhorende deduplicatie als deze optie niet beschikbaar. Na het contacteren van hun support geven zij aan dat LSFS niet meer ondersteund wordt door Starwind.

Hun software beschikt dus niet meer over deduplicatie functies en dergelijke.

we vallen hier dus terug op hypervisor of server specifieke storage technieken.

Scale up naar 3-node setup

Architectuur



Prerequisites per node:

4x	Virtual 1.7 Ghz processor
1x	4GB RAM (If using cache -> min. 8GB)
2x	Network interface for Synchronization traffic
2x	Network interface for ISCSI / Heartbeat traffic
1x	Network interface for management
1x	Hypervisor with vSphere 5.5 or later for each server
1x	20 GB Virtual disk for Starwind OS
1x	Virtual disks on separate partition (Thick provisioned Eager Zeroed) To create shared storage pool
1x	Starwind vSAN license for 3 nodes
1x	Windows server (vm) to run Starwind management console

We installeren een nieuwe Starwind vSAN vm op een 3^{de} ESXI server.
Deze VM zal zoals hierboven beschreven 4 netwerkkarten nodig hebben & dus ook 4 ip adressen:

Edit Settings | vmstorage3 >

Virtual Hardware | VM Options

ADD NEW DEVICE

CPU	4	
Memory	8	GB
Hard disk 1	16	GB
Hard disk 2	100	GB
SCSI controller 0	LSI Logic SAS	
Network adapter 1	ISCSI_Network_1	<input checked="" type="checkbox"/> Connected
Network adapter 2	ISCSI_Network_2	<input checked="" type="checkbox"/> Connected
Network adapter 3	ISCSI_Network_1	<input checked="" type="checkbox"/> Connected
Network adapter 4	ISCSI_Network_2	<input checked="" type="checkbox"/> Connected
CD/DVD drive 1	Client Device	<input type="checkbox"/> Connected
Video card	Specify custom settings	
VMCI device	Device on the virtual machine PCI bus that provides support for the virtual machine communication interface	
SATA controller 0	AHCI	

Remote console > commando 'nmtui' > Edit a connection >
De 2 extra NIC's krijgen nummers ens161 & ens192 > Stel deze in met de juiste ip adressen:

Profile name: ens256
Device: ens256 (08:50:56:A4:9B:B6)

= ETHERNET <Show>

IPv4 CONFIGURATION <Manual> <Hide>

Addresses: 192.168.1.100 <Remove> <Add...>
Gateway: 192.168.1.1
DNS servers: <Add...>
Search domains: <Add...>

Routing (No custom routes) <Edit...>

[] Never use this network for default route
[] Ignore automatically obtained routes
[] Ignore automatically obtained DNS parameters
[] Require IPv4 addressing for this connection

= IPv6 CONFIGURATION <Automatic> <Show>

[X] Automatically connect
[X] Available to all users

<Cancel> <OK>

Voeg bij de anderen 2 nodes ook 2 netwerkkarten toe en stel de ip adressen in

Bij alle 3 vm's voegen we een extra disk to van 240GB.

Edit Settings | vmstorage1

×

Virtual Hardware

VM Options

ADD NEW DEVICE

> CPU	4	▼
> Memory	8	GB ▼
> Hard disk 1	16	GB ▼
> Hard disk 2	100	GB ▼
▼ New Hard disk *	240	GB ▼
Maximum Size	2.41 TB	
VM storage policy	Datastore Default ▼	
Location	Store with the virtual machine ▼	
Disk Provisioning	Thick Provision Eager Zeroed ▼	
Sharing	Unspecified ▼	
Shares	Normal ▼	1000
Limit - IOPs	Unlimited ▼	
Virtual flash read cache	0	MB ▼
Disk Mode	Dependent ▼	
Virtual Device Node	SCSI controller 0: New Hard disk	

Deze gaan we gebruiken om extra opslag pools aan te maken op de starwind vm's

Format /dev/sdc

Erase

Don't overwrite existing data ▼

Type

XFS - Recommended default ▼

Name

Disk2

☐ Encrypt data

Mounting

Custom ▼

Mount Point

/mnt/disk2

Mount Options

☒ Mount at boot
☐ Mount read only
☐ Custom mount options

Formatting a storage device will erase all data on it.

Cancel

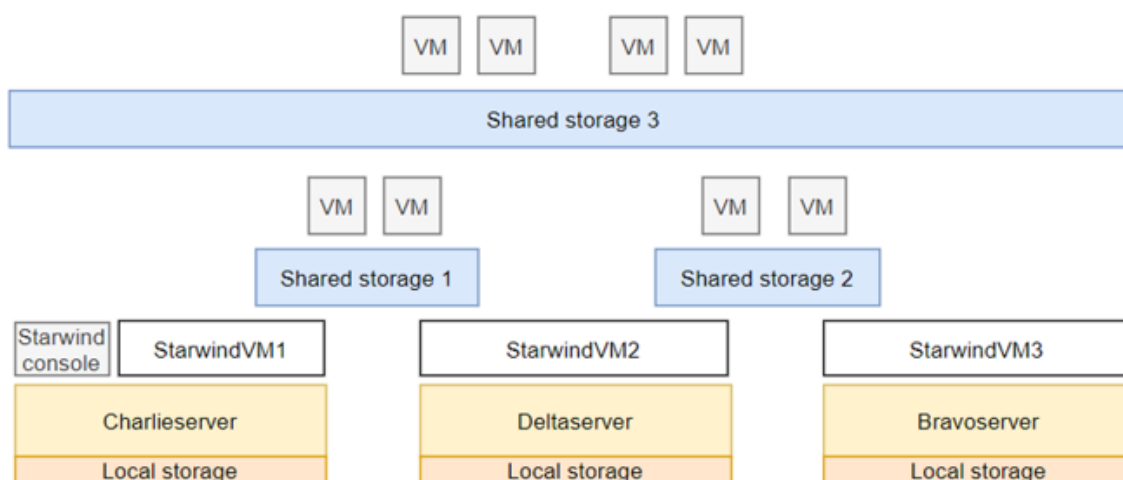
Format

Bij elke Starwind vm beschikken we nu over 2 opslag pools:

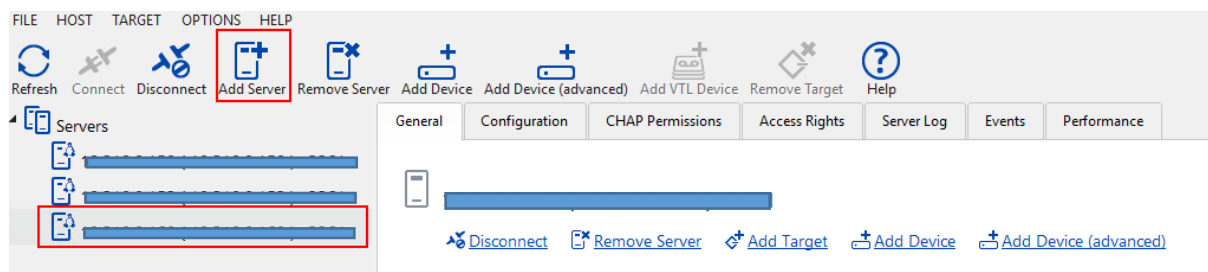
Filesystems			
Name	Mount Point	Size	
/dev/centos/root	/	<div><div></div></div>	2.62 / 13.4 GiB
/dev/sda1	/boot	<div><div></div></div>	147 / 1014 MiB
Disk1	/mnt/disk1	<div><div></div></div>	62.0 / 100 GiB
Disk2	/mnt/disk2	<div><div></div></div>	0.0315 / 240 GiB

Deze gaan we vervolgens gebruiken om High available Virtual Disk's aan te maken in de Starwind console

De Shared storage setup die we in deze proof of concept willen bereiken is de volgende:



Voeg de 3^{de} Starwind node toe in de management console:



Op node 1 maken we de sw-vsan-01 aan. Deze wordt shared storage voor ESX 1 & 2:

Add Device advanced > Hard Disk Device > Virtual Disk

Deze plaatsen we op storage pool disk1

Virtual Disk Location

☒ Create a New Virtual Disk

Name:

Location:

Size: GB

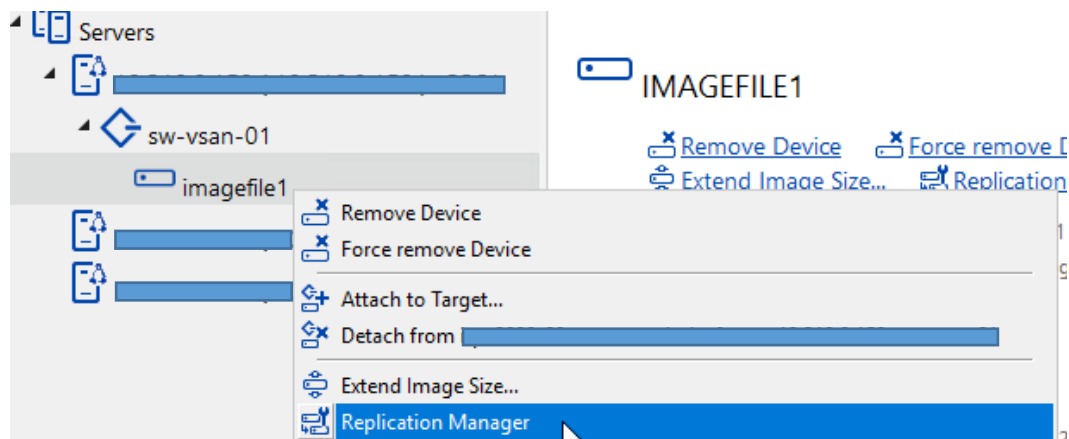
☐ Use an Existing Virtual Disk

Location:

☐ Read-Only Mode

Next > Thick provisioned > N/A > No Flash Cache > New target > Create

Via replication manager stellen we replication in voor sw-vsan-01 :



Add replica > Synchronous Replication > ip adres van 2de node

Add Partner Node

Specify Partner Host Name or IP Address where Replication Node would be created

Host Name or IP Address

Port Number

Omdat we met 3 nodes werken kiezen we best voor Node Majority Failover

Failover Strategy



Heartbeat

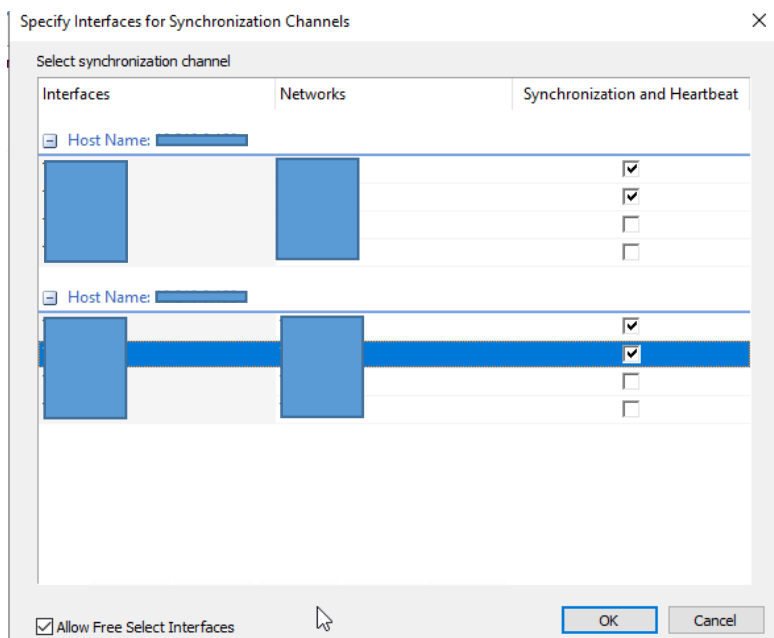
Process node and communication failures using additional communication channel (heartbeat). At least one synchronization or heartbeat channel must be functional for proper failover processing. Loss of all communication channels may lead to split brain issue, so it's recommended to use client iSCSI connection interfaces as heartbeat channel.



Node Majority

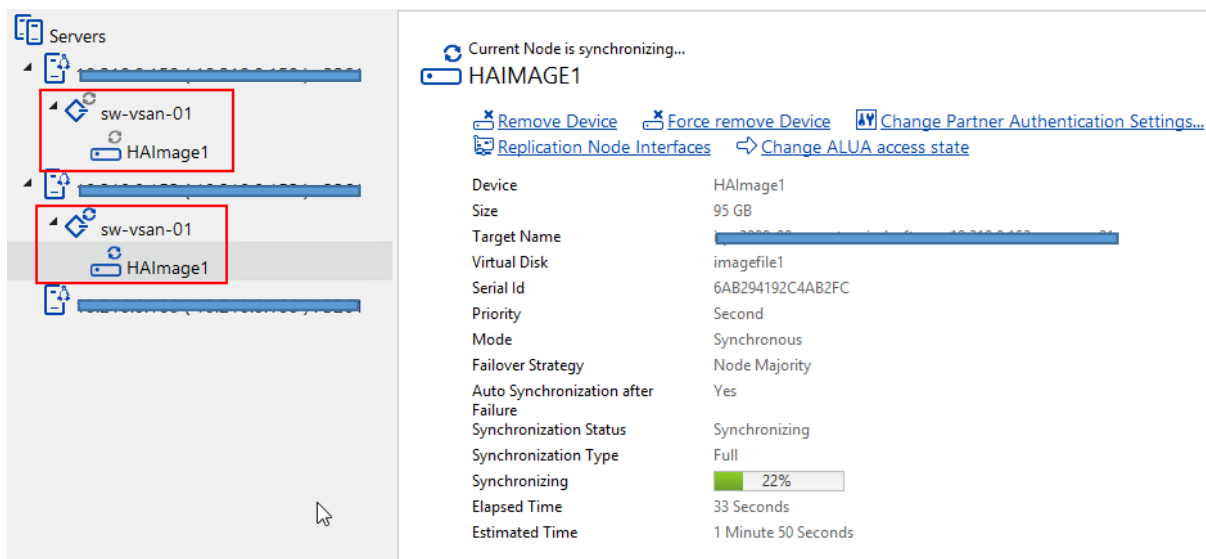
Process node and communication failures using majority policy: node stays active while it sees more than half of nodes including itself. In case of 2 storage nodes, requires configuring additional witness node. Does not require additional heartbeat channel.

Create new partner device > Change network settings > Selecteer de juiste NIC's



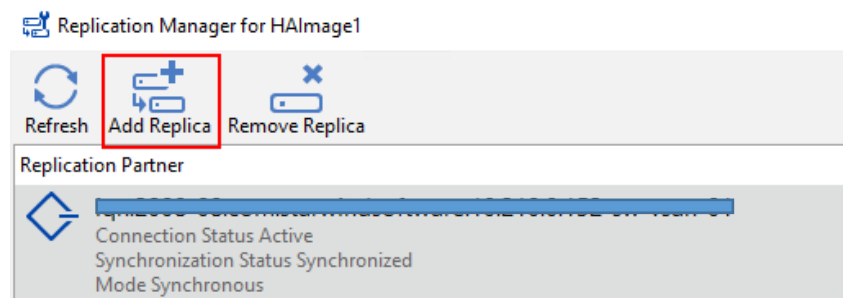
Create Replica

sw-vsan-01 wordt gerepliceerd naar de 2de node



Omdat we voor Node Majority failover hebben gekozen moeten we een witness node instellen. Onze 3^{de} node zal dienen als Witness

Rechterklik op sw-vsan-01 > Replication manager > Add replica



We voegen m.a.w. een extra replica partner toe die enkel als witness zal dienen

We kiezen dus voor Witness Node:

Replication Mode

- ☐ **Synchronous "Two-Way" Replication**
Replication Partner must be connected to Client as Source Device as well, MPIIO on Client must be enabled, needs dedicated high Performance Network Connection for Synchronization.
- ☐ **Asynchronous "One-Way" Replication**
Replica is used to store replicated Data, Data is stored as Snapshots, Client cannot connect to Replication Partner, mount Snapshot from Replica to get Access to replicated Data.
- ☒ **Witness Node**
Witness node doesn't contain user data. In case when Node Majority policy is set for Synchronous replication device and there are two storage nodes, Witness Node must be added to cluster to make number of nodes odd number and enable proper functioning of Node Majority policy.

En geven het ip adres van de 3^{de} node op

Add Partner Node

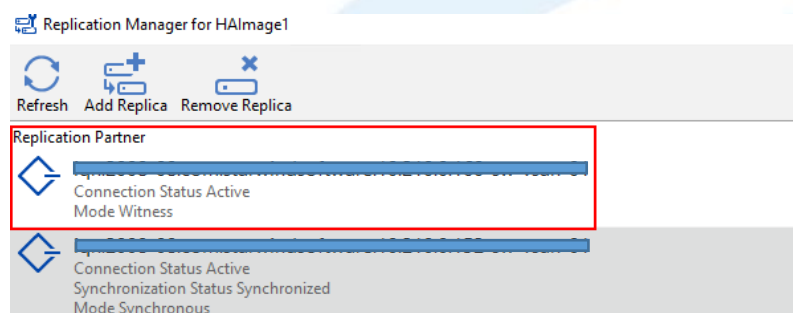
Specify Partner Host Name or IP Address where Replication Node would be created

Host Name or IP Address

Port Number

Next > Selecteer de juiste NIC's > Create replica

In de replication manager zien we nu dat er een extra partner is als Witness:



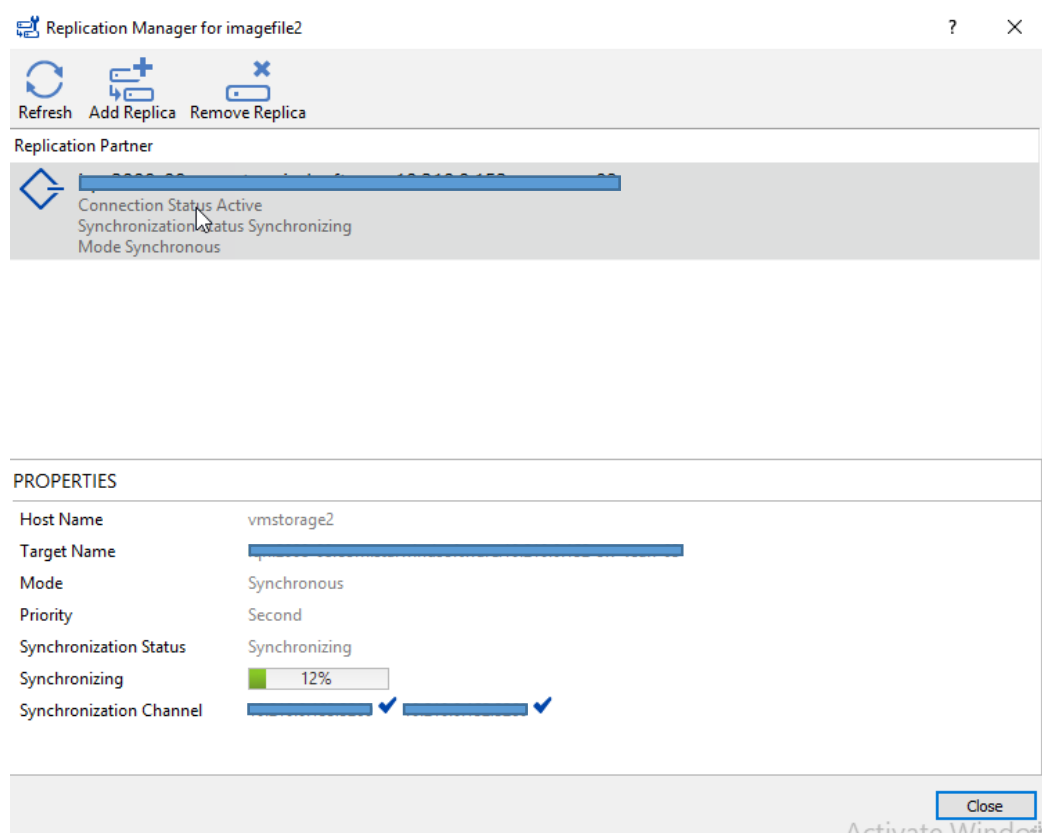
Om sw-vsan-02 aan te maken volgen we exact dezelfde stappen.
Het enige verschil is dat je deze dient in te stellen tussen node 2 en 3. Hier kan je best de 1^{ste} node selecteren als Witness node

We gaan verder met sw-vsan-03. Deze zal shared storage creëren tussen alle 3 de nodes:

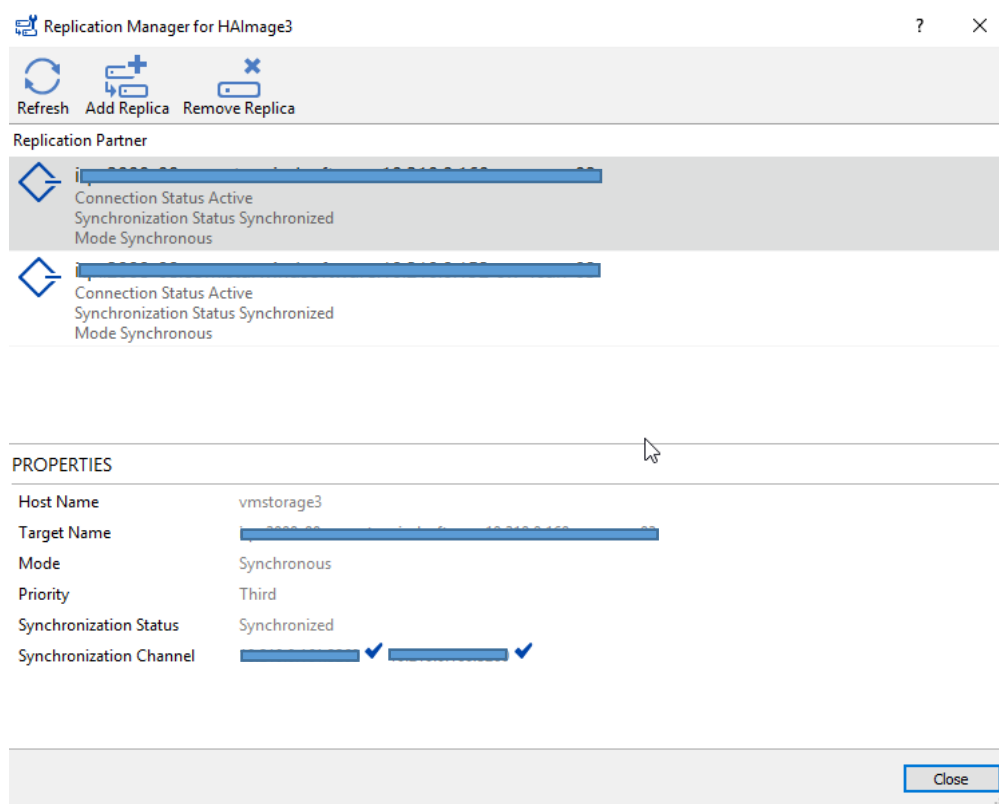
Node 1 > Add Device Advanced > Virtual disk > Thick provisioned > Create Device

Replication manager:

Voeg een synchrone replica toe naar node 2:

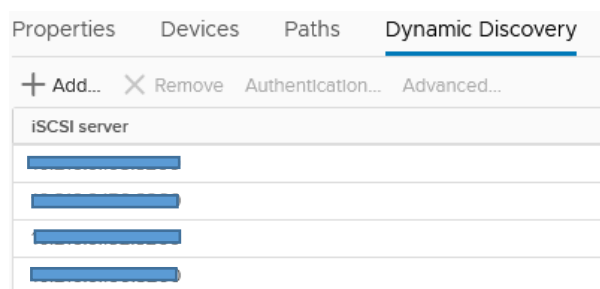


Voeg een extra replica toe naar node 3:






Vervolgens dienen we deze shared storage toe te voegen aan vSphere zoals we eerder deden bij de 2-node setup

Hiervoor moeten we het ip van de 3^{de} node nog toevoegen aan de iSCSI software adapter:



Create vervolgens de Datastores met gepaste naam:

-  sw-vsan-01
-  sw-vsan-02
-  sw-vsan-03

Testing 3-Node setup

Dankzij vorige stappen uit te voeren hebben we succesvol een 3-node setup gebouwd met virtual shared storage.

VM Migration (compute resource) tussen 3 nodes:

Succesvol. Compute resource kan snel gemigreerd worden tussen nodes.

Uitval van 1 VSA:

De opstelling kan de uitval van 1 node permitteren

Uitval van 2 VSA's:

Bij uitval van 2 VSA's zal de virtuele storage niet meer werken. Minstens de helft van de nodes moeten draaien om operationeel te blijven

Failover:

Failover wordt mee afgehandelt door de vCenter HA configuratie. Indien een ESXI server zal uitvallen waar de compute resource draait zal deze automatisch snel herstart worden op de andere beschikbare ESXI servers.