StarWind Virtual SAN® 2-node Compute and Storage Separated Scenario with VMware vSphere 6.5/6.7

2024

TECHNICAL PAPERS



Trademarks

“StarWind”, “StarWind Software” and the StarWind and the StarWind Software logos are registered trademarks of StarWind Software. “StarWind LSFS” is a trademark of StarWind Software which may be registered in some jurisdictions. All other trademarks are owned by their respective owners.

Changes

The material in this document is for information only and is subject to change without notice. While reasonable efforts have been made in the preparation of this document to assure its accuracy, StarWind Software assumes no liability resulting from errors or omissions in this document, or from the use of the information contained herein. StarWind Software reserves the right to make changes in the product design without reservation and without notification to its users.

Technical Support and Services

If you have questions about installing or using this software, check this and other documents first - you will find answers to most of your questions on the Technical Papers webpage or in StarWind Forum. If you need further assistance, please contact us .

About StarWind

StarWind is a pioneer in virtualization and a company that participated in the development of this technology from its earliest days. Now the company is among the leading vendors of software and hardware hyper-converged solutions. The company’s core product is the years-proven StarWind Virtual SAN, which allows SMB and ROBO to benefit from cost-efficient hyperconverged IT infrastructure. Having earned a reputation of reliability, StarWind created a hardware product line and is actively tapping into hyperconverged and storage appliances market. In 2016, Gartner named StarWind “Cool Vendor for Compute Platforms” following the success and popularity of StarWind HyperConverged Appliance. StarWind partners with world-known companies: Microsoft, VMware, Veeam, Intel, Dell, Mellanox, Citrix, Western Digital, etc.

Copyright ©2009-2018 StarWind Software Inc.

No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior written consent of StarWind Software.

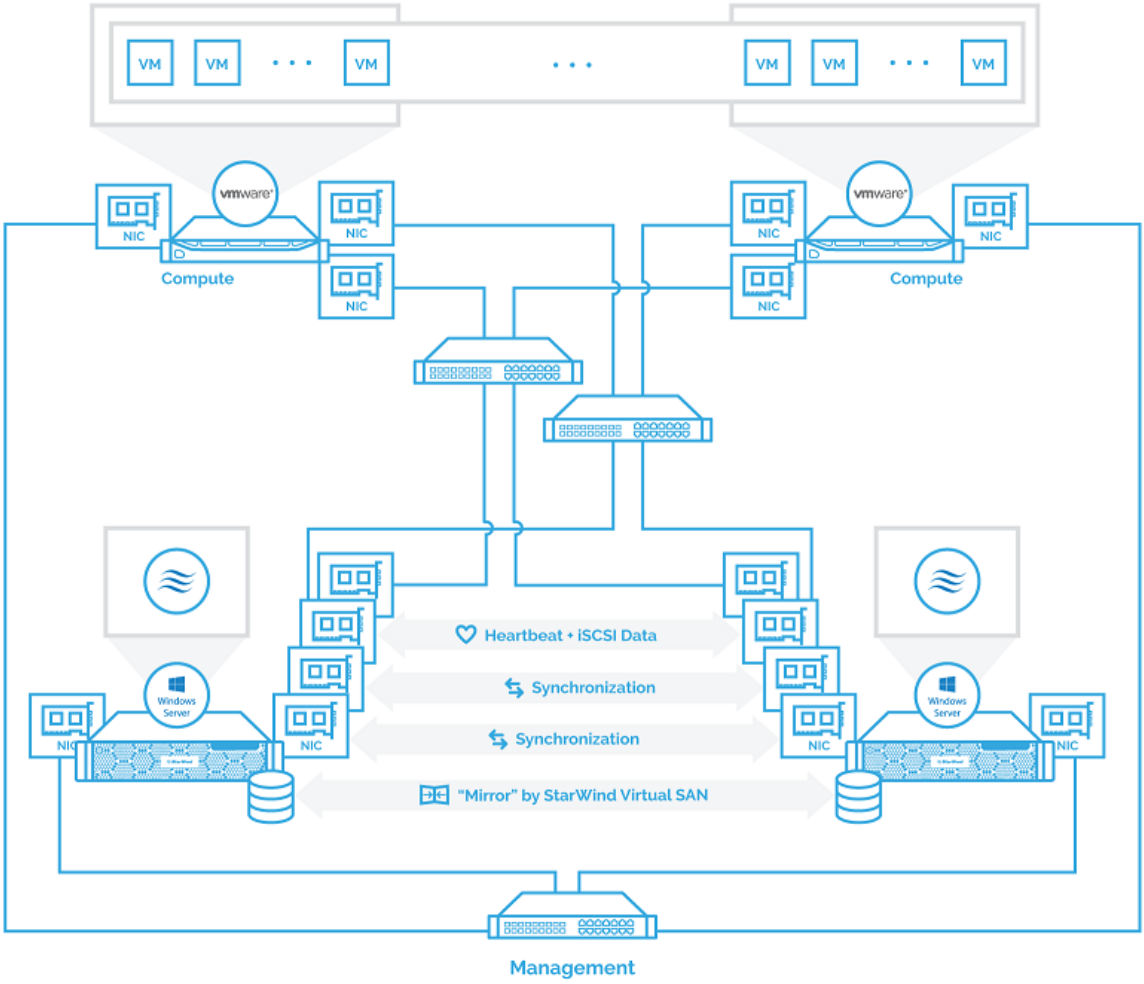
IMPORTANT NOTE: ESXi6.5 and ESXi 6.7 are no longer supported by VMware (see more details here: https://lifecycle.vmware.com/). This guide should be used for informational, test, and educational purposes only.

Introduction To Starwind Virtual San For Hyper-V

StarWind Virtual SAN is a native Windows hypervisor-centric hardware-less VM storage solution. It creates a fully fault-tolerant and high performing storage pool built for the virtualization workloads by mirroring the existing server's storage and RAM between the participating storage cluster nodes. The mirrored storage resources are then connected to all cluster nodes and treated just as a local storage by all hypervisors and clustered applications. High Availability (HA) is achieved by providing multipath access to all storage nodes. StarWind Virtual SAN delivers supreme performance compared to any dedicated SAN solution since it runs locally on the hypervisor and all I/O is processed by local RAM, SSD cache, and disks. This way it never gets bottlenecked by storage fabric.

Solution Diagram

The diagram below illustrates the network and storage configuration of the resulting solution described in the guide.



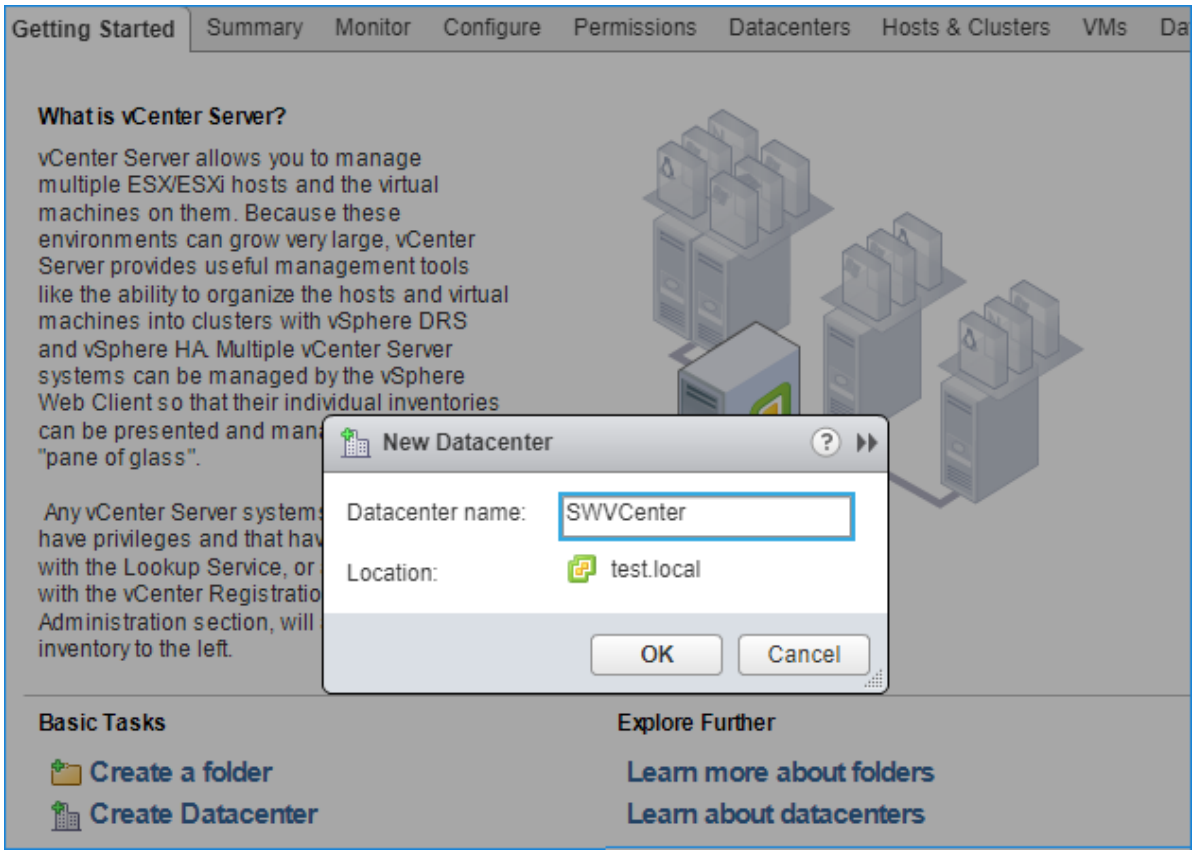
Preparing Hypervisor

Creating Datacenter

NOTE: Before creating Datacenter, the vCenter server should be deployed.

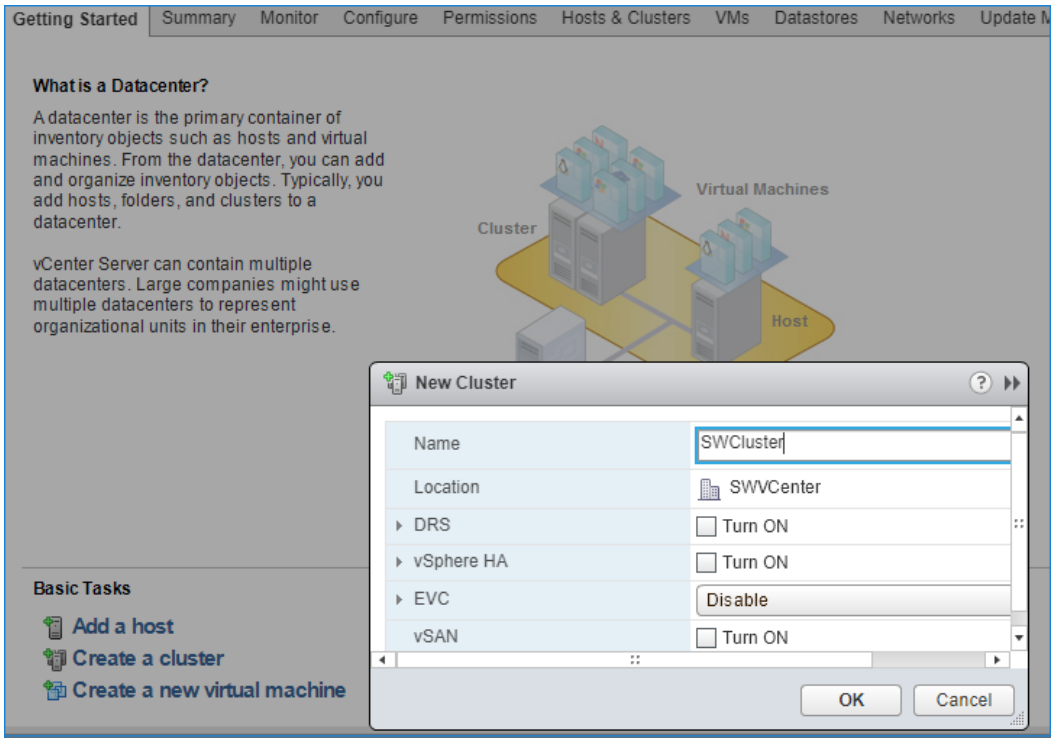
1. Connect to vCenter, select the Getting Started tab.

2. Click on Create Datacenter and enter Datacenter name.



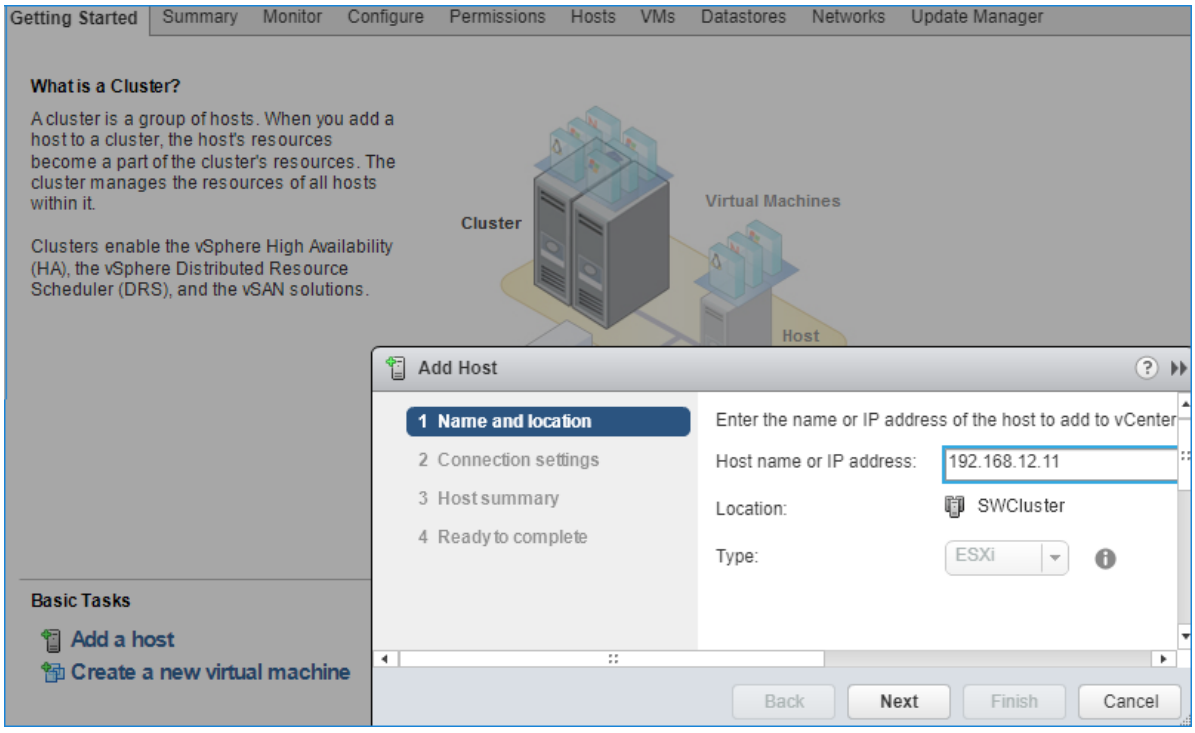
Creating Cluster

1. Click the Datacenter Getting Started tab and click Create a cluster. Enter cluster Name and click OK.



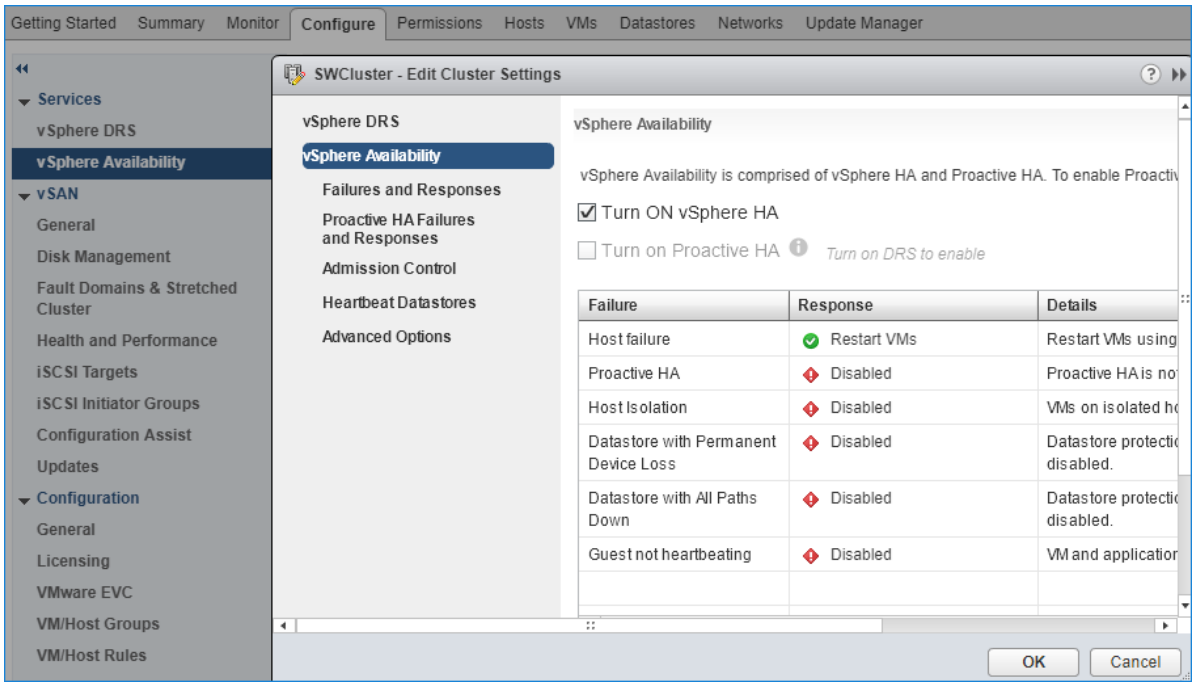
2. Open the Cluster pane and click Add host.

3. Enter the ESXi Host name or IP address and enter the administrative account.



4. Click Cluster -> Configure -> Edit. In vSphere Availability area, select the Turn on

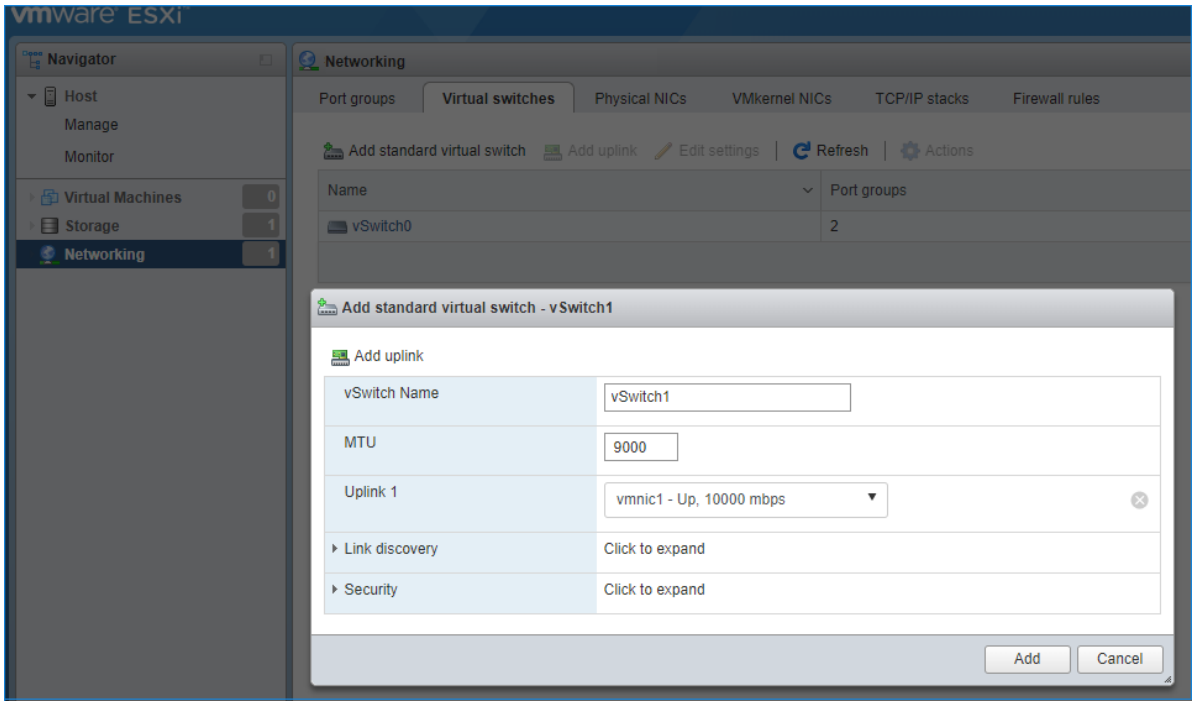
vSphere HA checkbox.



Configuring Networks

NOTE: Configure network interfaces on each node to make sure that the iSCSI interface is in a different subnet and connected physically according to the network diagram above. In this document, two links with 172.16.30.x and 172.16.40.x subnets per ESXi host are used for iSCSI traffic. All actions below should be taken for each ESXi server.

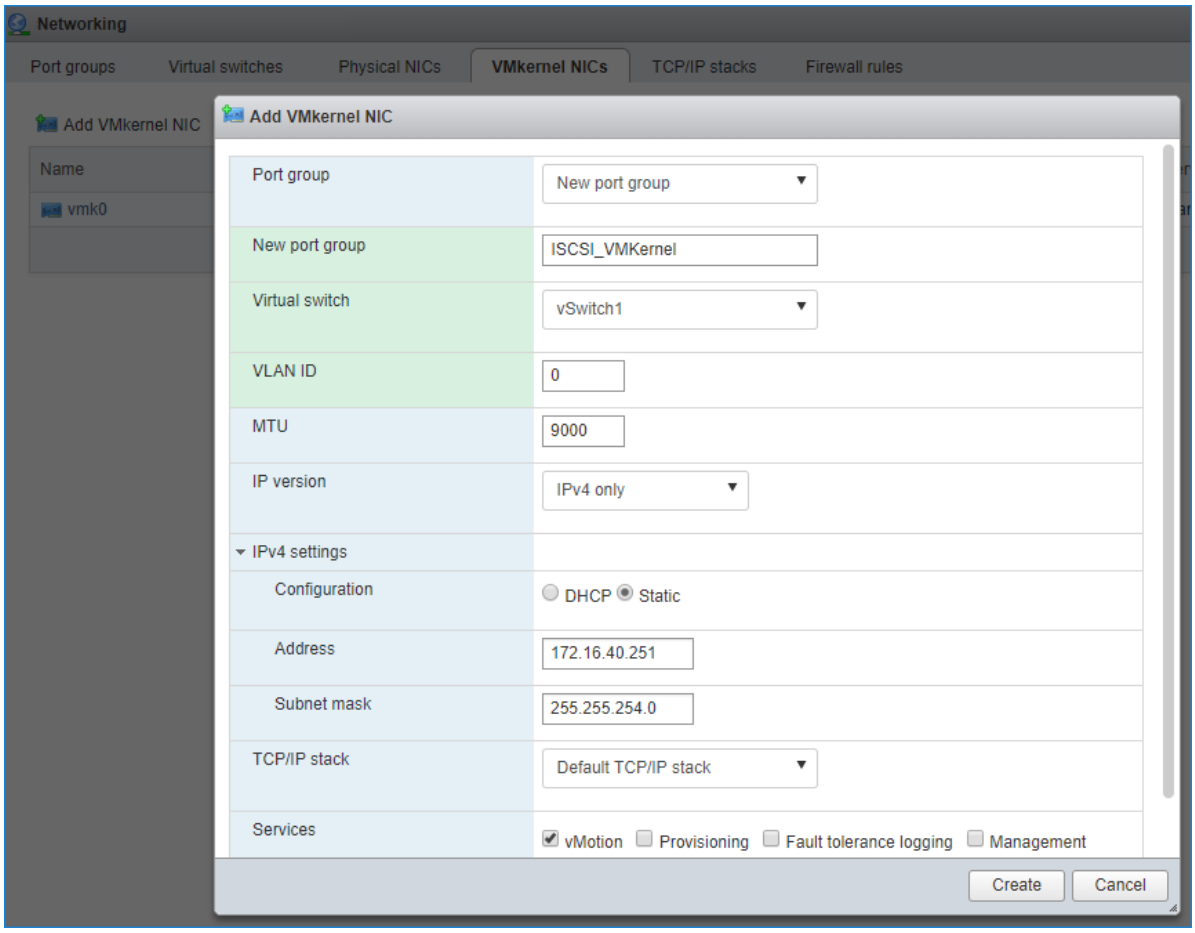
1. Create vSwitch for the iSCSI traffic via the first Ethernet Switch.



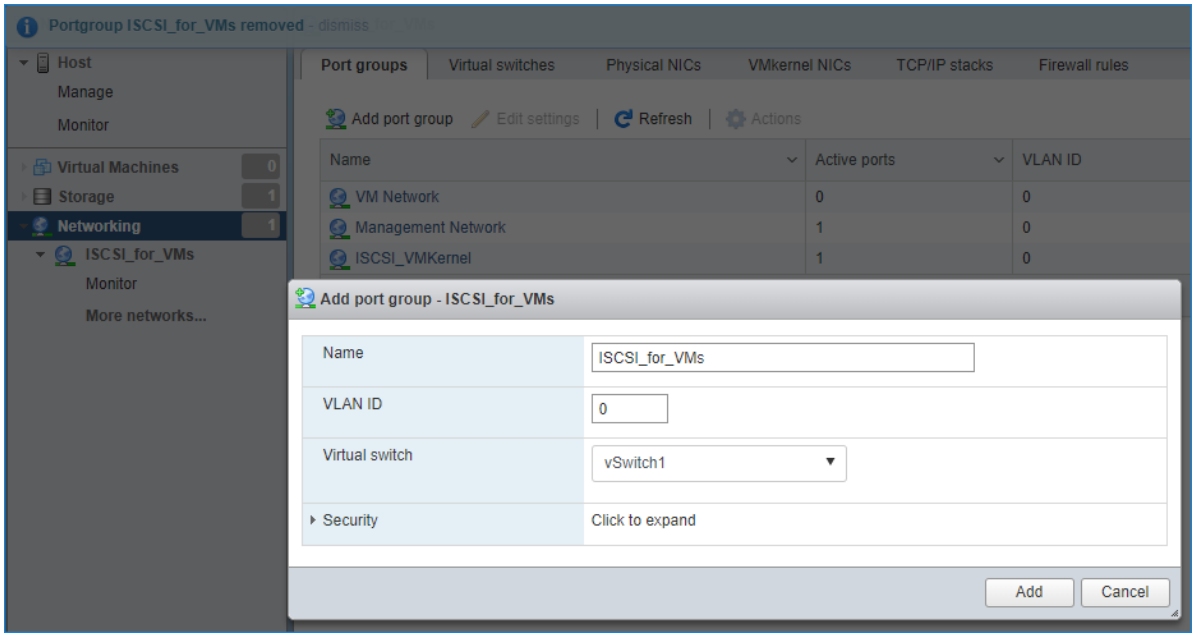
2. Create vSwitch for the iSCSI traffic via the second Ethernet Switch.

NOTE: iSCSI vSwitch requires creating Virtual Machine Port Group and VMKernel. VMKernel port requires assigning static IP address. It is recommended to set jumbo frames to 9000 on vSwitches and VMKernel ports for iSCSI traffic. Additionally, users can enable vMotion on VMKernel port.

3. Create VMKernel ports for iSCSI channels.



4. In VMkernel adapters pane, click Add host networking and Add Virtual Machine Port Group on the vSwitches for iSCSI traffic.



5. In Ready to complete, click Finish.

Preparing Starwind Servers

1. Install Windows Server (see supported OS on system requirements page: https://www.starwindsoftware.com/system-requirements) and StarWind VSAN on each host.

2. Make sure that server hardware used for StarWind Virtual SAN deployment satisfies the requirements listed below.

Minimum requirements:

RAM: 4 GB (plus the size of the RAM cache if applicable);

CPUs: 1 socket with 2 GHz;

Hard disk 1: 100 GB for OS (recommended);

Hard disk 2: Depends on the storage volume to be used as shared storage.

Network adapter 1: Management

Network adapter 2: iSCSI1

Network adapter 3: iSCSI2

Network adapter 4: Sync1

Network adapter 5: Sync2

NOTE: It is recommended to set jumbo frames to 9000 for iSCSI and Synchronization traffic. StarWind host allows adding Active Directory Domain Services role if necessary, thus it can serve as a domain controller.

Installing Starwind Vsan For Hyper-V

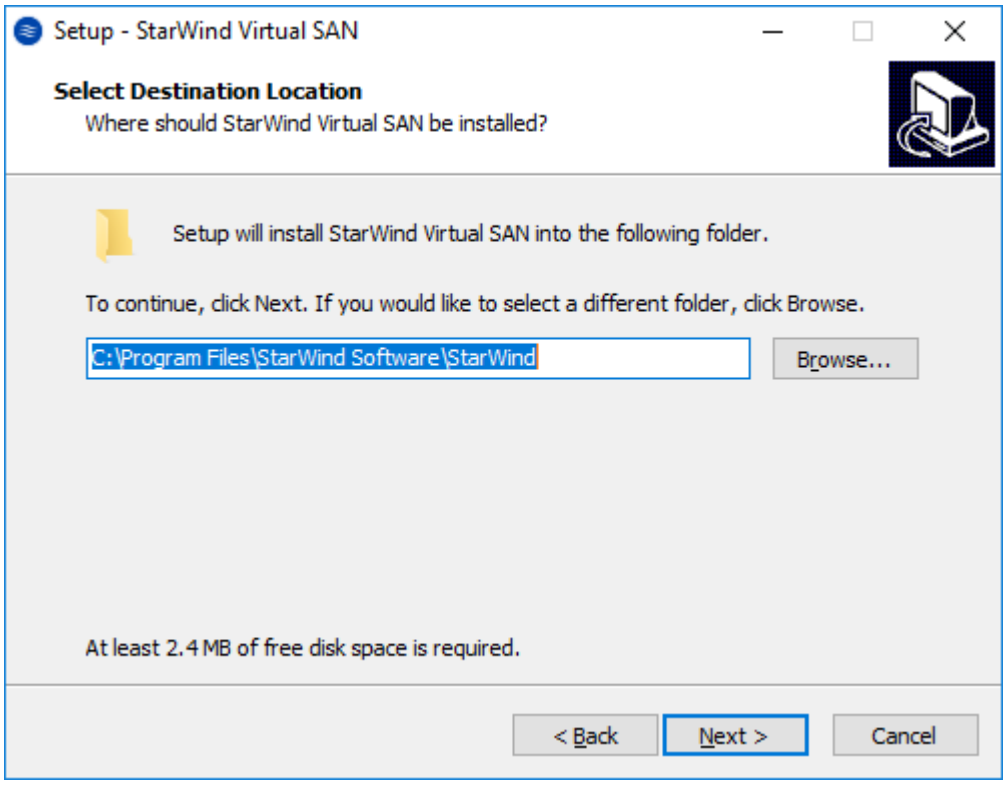
1. Download the StarWind setup executable file from the StarWind website: https://www.starwind.com/registration-starwind-virtual-san

2. Launch the downloaded setup file on the server to install StarWind Virtual SAN or one of its components. The Setup wizard will appear. Read and accept the License Agreement.



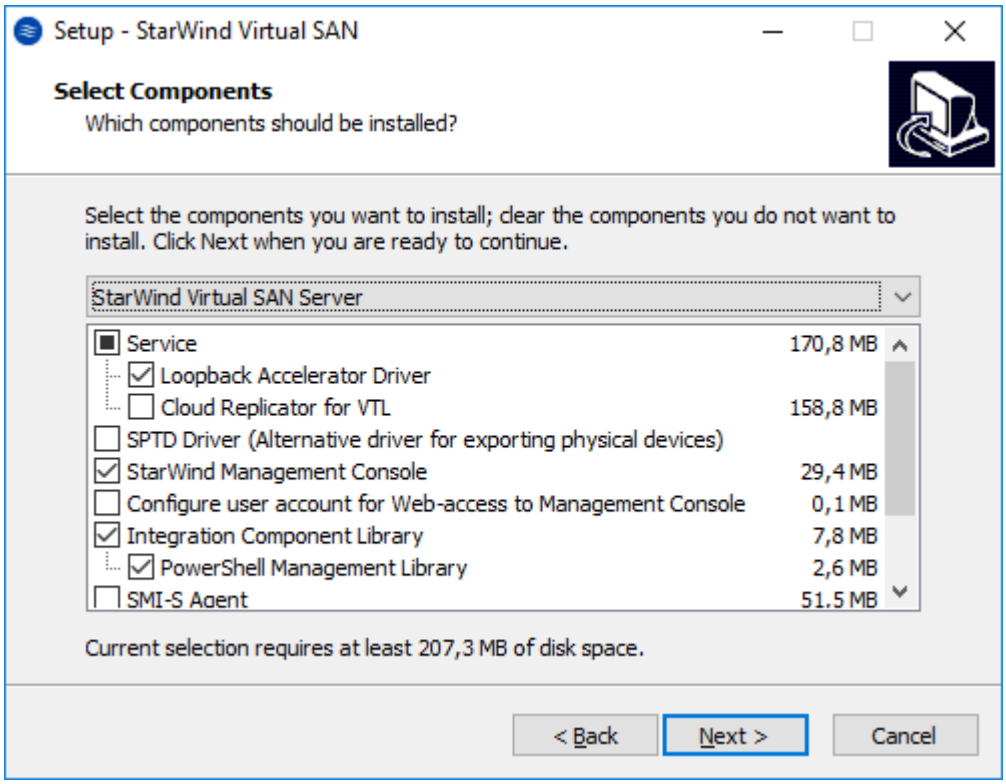
3. Carefully read the information about the new features and improvements. Red text indicates warnings for users that are updating the existing software installations.

4. Select Browse to modify the installation path if necessary. Click on Next to continue.

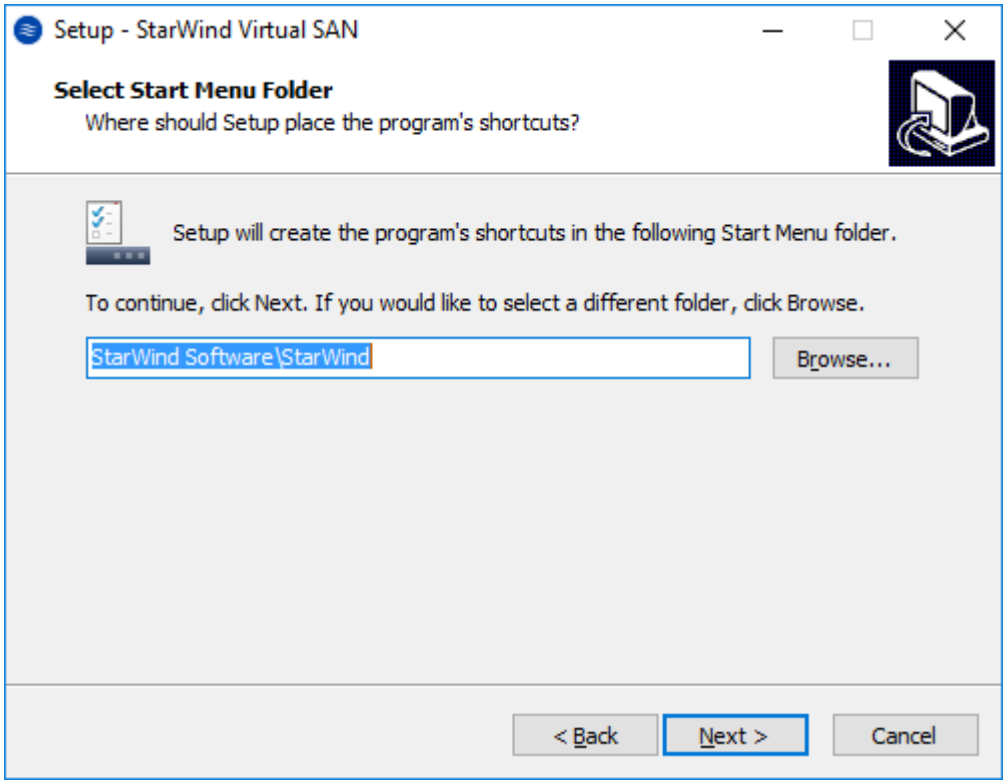


5. Select the following components for the minimum setup:

- StarWind Virtual SAN Service. The StarWind Virtual SAN service is the “core” of the software. It can create iSCSI targets as well as share virtual and physical devices. The service can be managed from StarWind Management Console on any Windows computer that is on the same network. Alternatively, the service can be managed from StarWind Web Console deployed separately.
- StarWind Management Console. Management Console is the Graphic User Interface (GUI) part of the software that controls and monitors all storage-related operations (e.g., allows users to create targets and devices on StarWind Virtual SAN servers connected to the network).
  NOTE: To manage StarWind Virtual SAN installed on a Windows Server Core edition with no GUI, StarWind Management Console should be installed on a different computer running the GUI-enabled Windows edition.



6. Specify Start Menu Folder.



7. Enable the checkbox if a desktop icon needs to be created. Click on Next to continue.

8. When the license key prompt appears, choose the appropriate option:

- request time-limited fully functional evaluation key.
- request FREE version key.
- relect the previously purchased commercial license key.

9. Click on the Browse button to locate the license file.

10. Review the licensing information.

11. Verify the installation settings. Click on Back to make any changes or Install to proceed with installation.

12. Enable the appropriate checkbox to launch StarWind Management Console right after the setup wizard is closed and click on Finish.

13. Repeat the installation steps on the partner node.

Provisioning Storage With Starwind Vsan

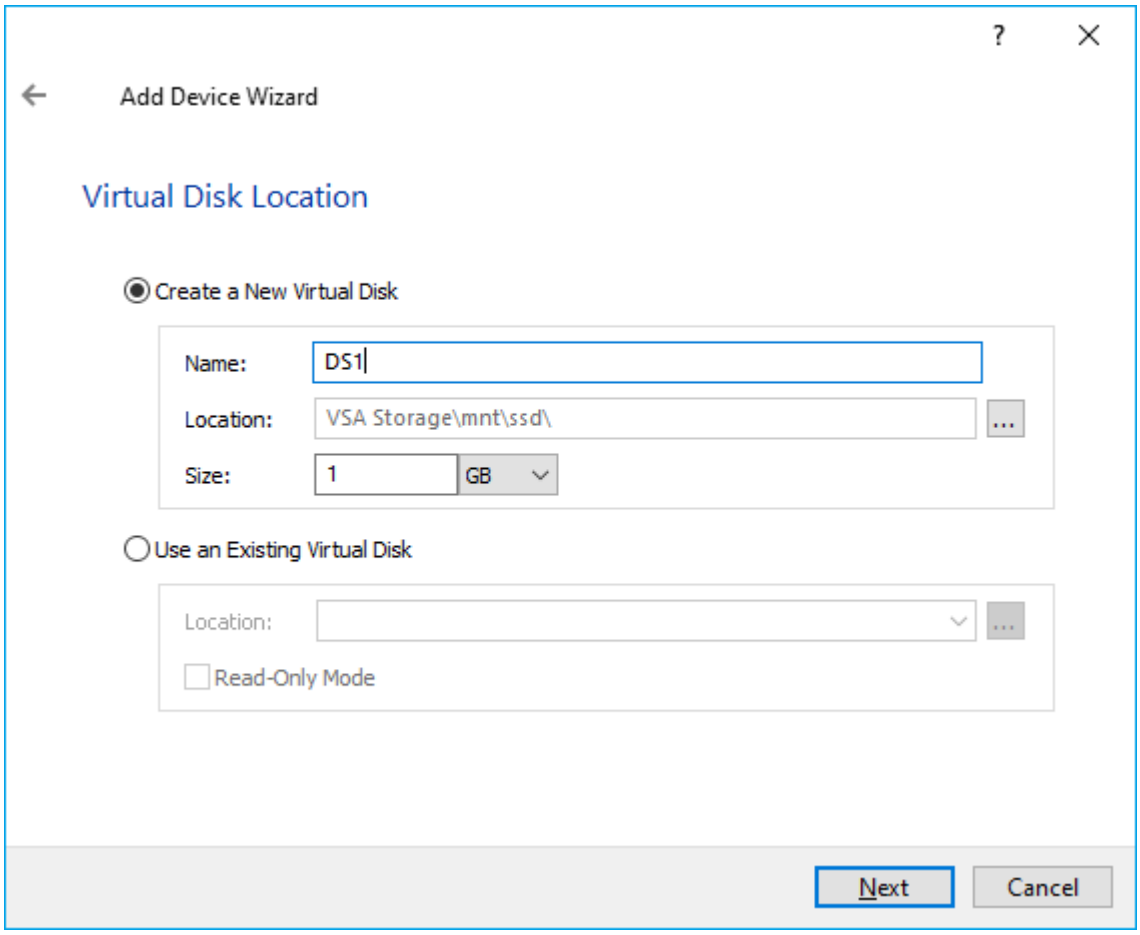
Creating StarWind HA devices

1. Open StarWind Management Console and click on the Add Device (advanced) button.

2. Select Hard disk device as the type of device to be created. Click Next to continue.

3. Select Virtual disk. Click Next to continue.

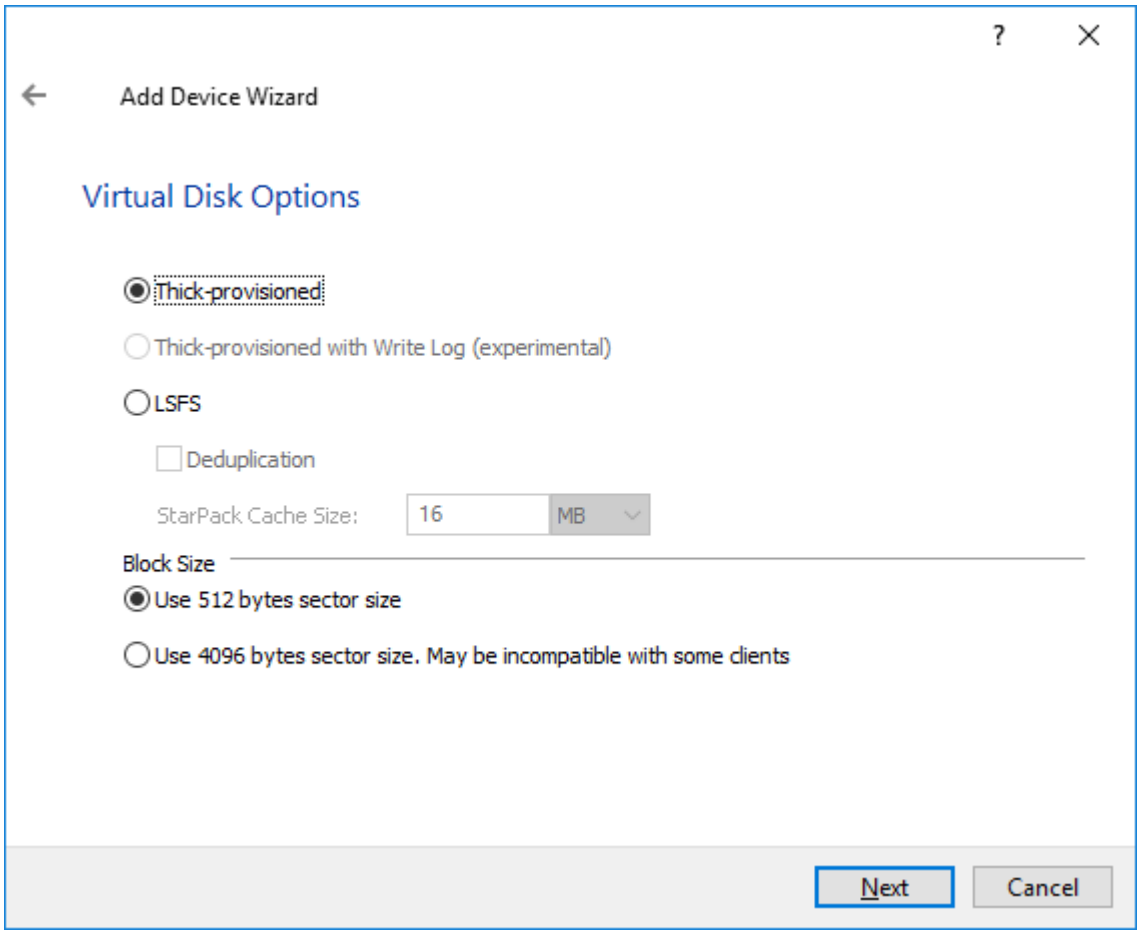
4. Specify virtual disk Name, Location, and Size.



NOTE: Image file device and HA device can be extended later according to the steps in KB article:

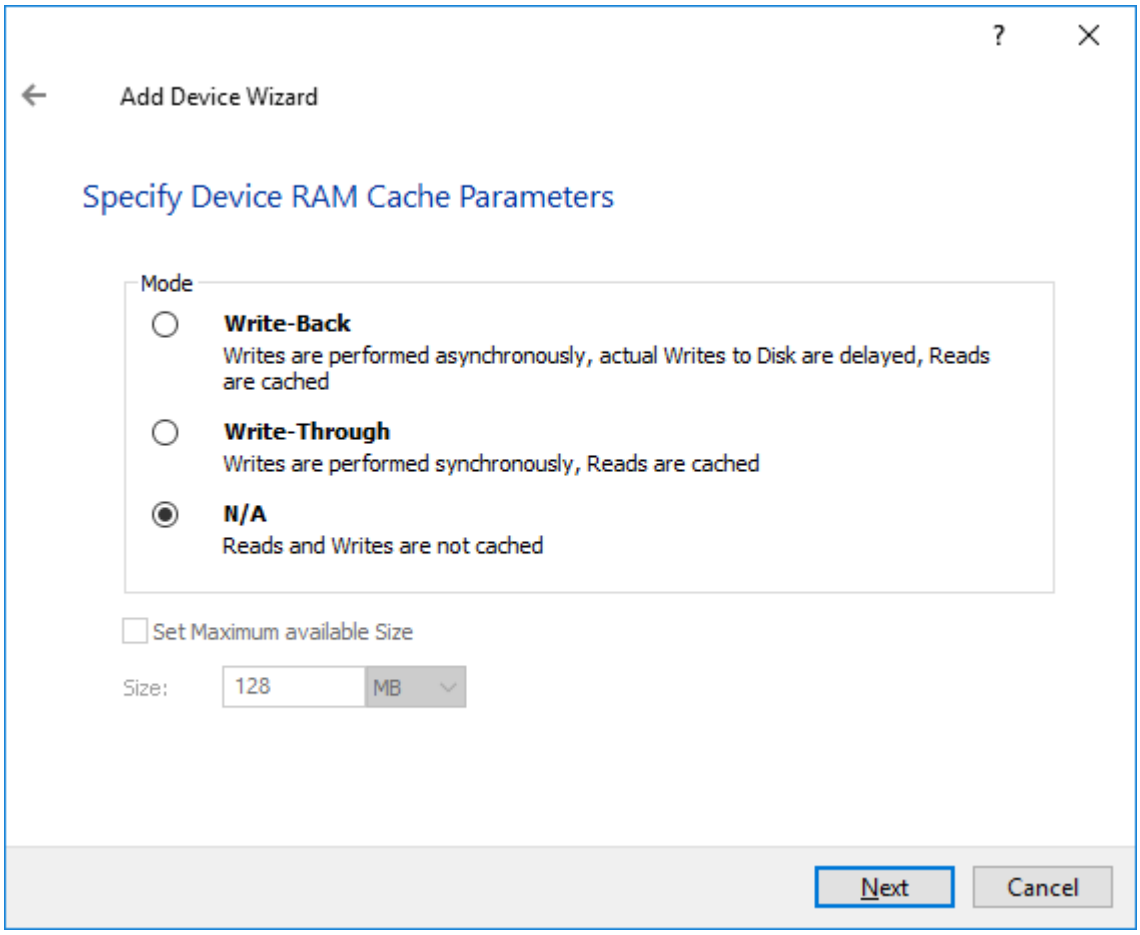
How to extend Image File or High Availability device

5. Specify virtual disk options.

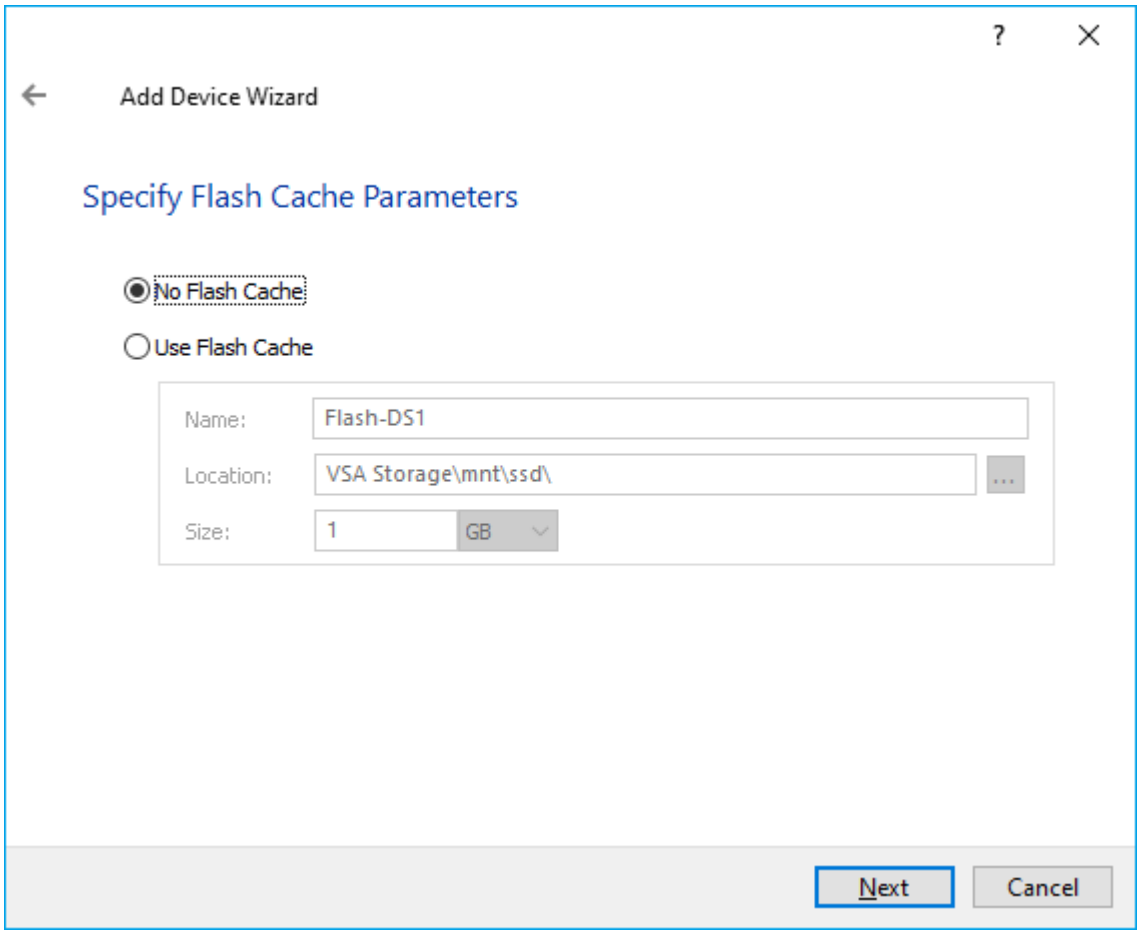


NOTE: Sector size should be 512 bytes in case of using ESXi.

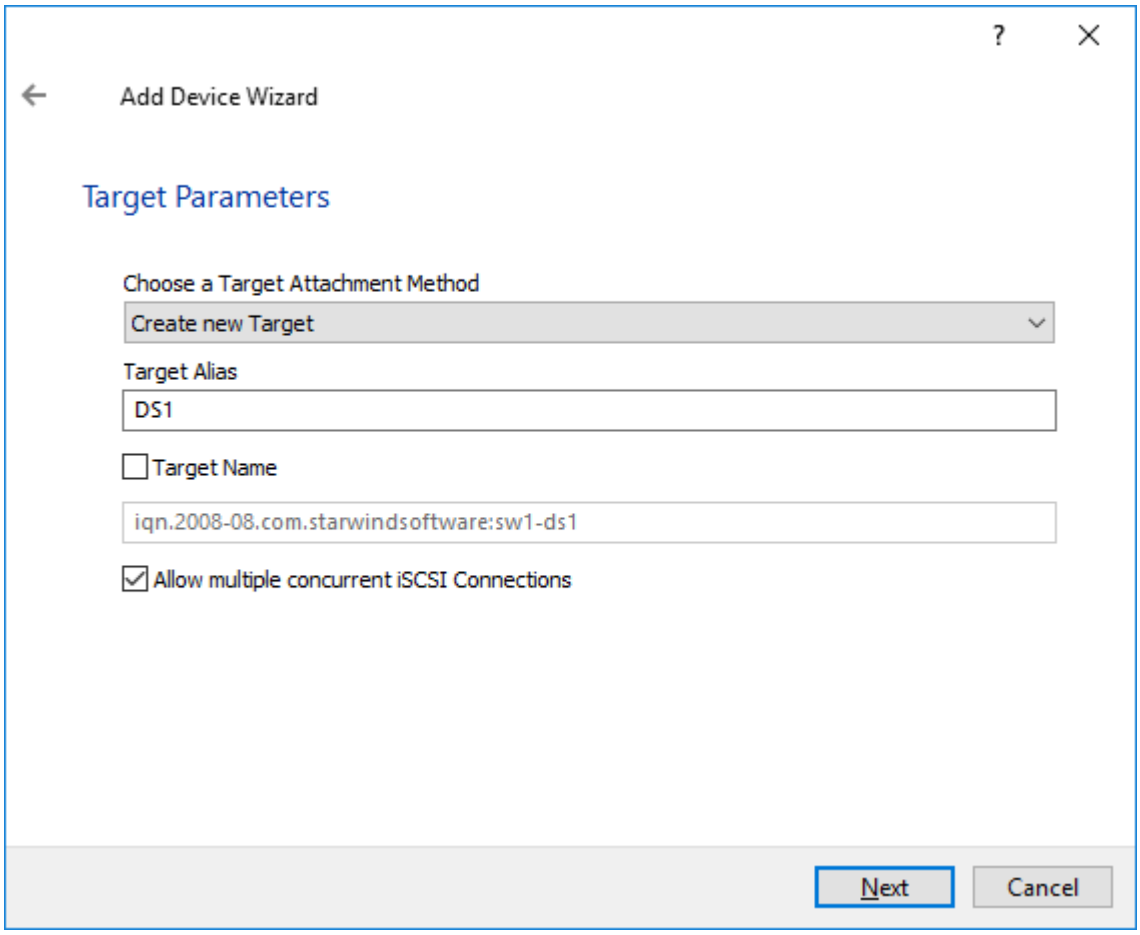
6. Define the RAM caching policy and specify the cache size (in the corresponding units) and click Next to continue.



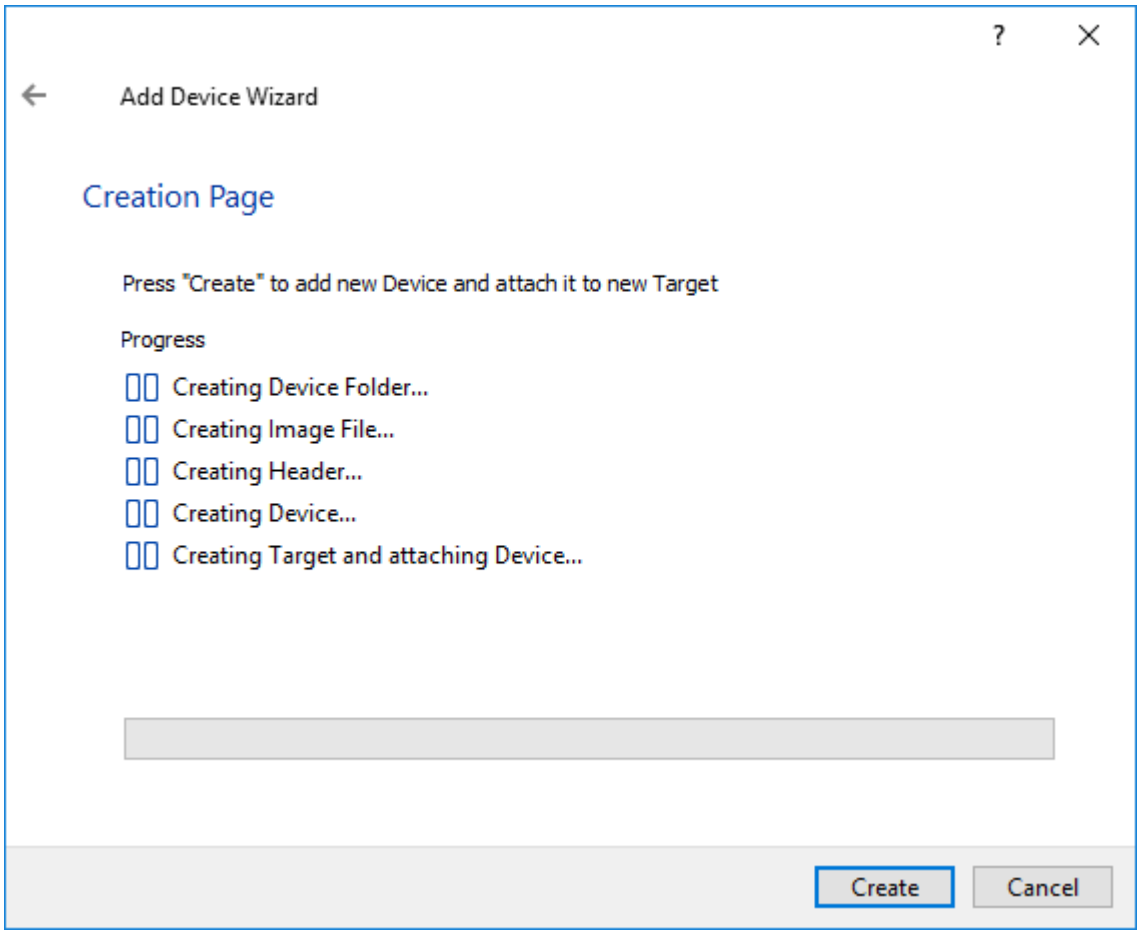
7. Define the Flash caching policy and the cache size. Click Next to continue.



8. Specify Target Parameters and select the Allow multiple concurrent iSCSI connections checkbox to enable several clients to connect simultaneously to the target.



9. Click Create to add a new device and attach it to the target.

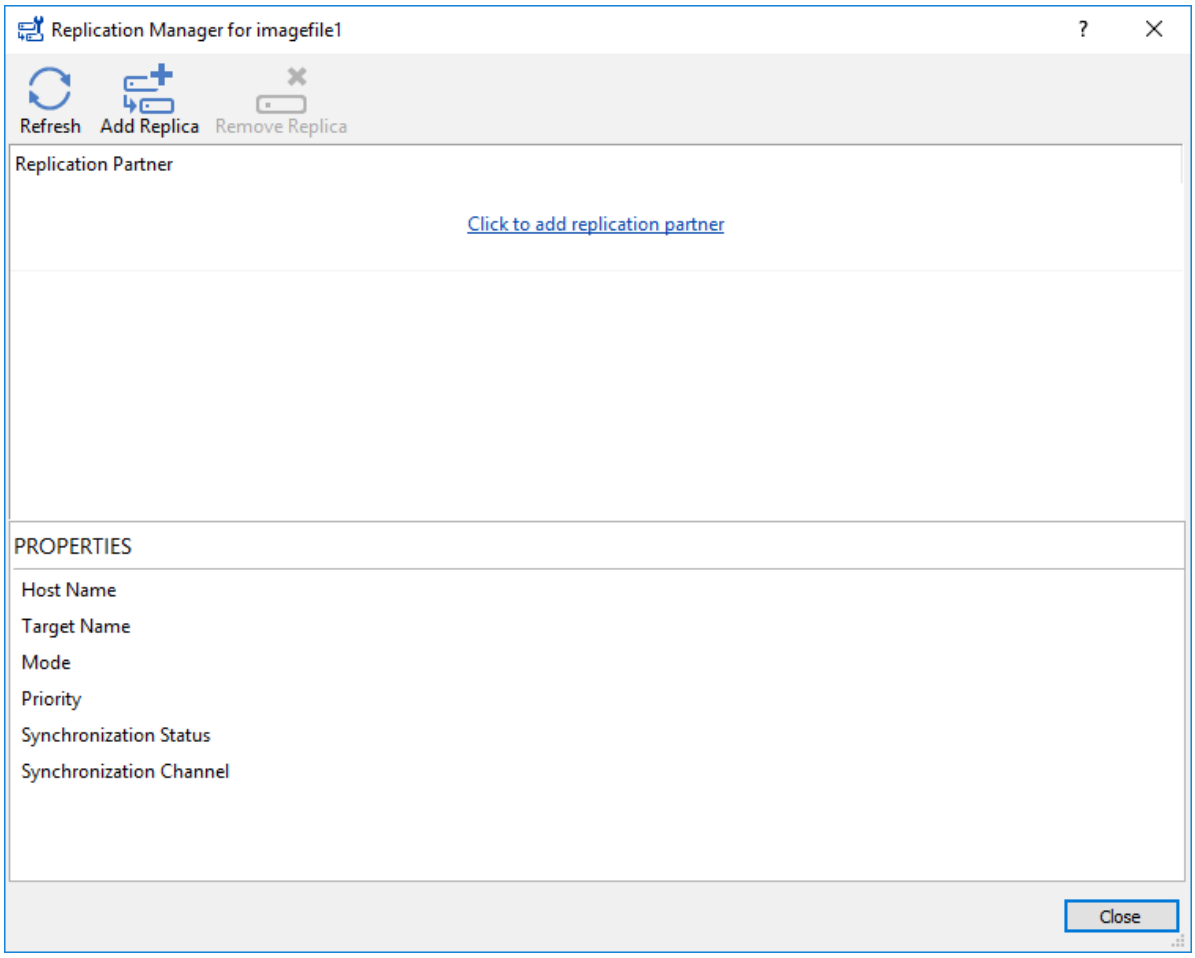


10. Click Create to add a new device and attach it to the target.

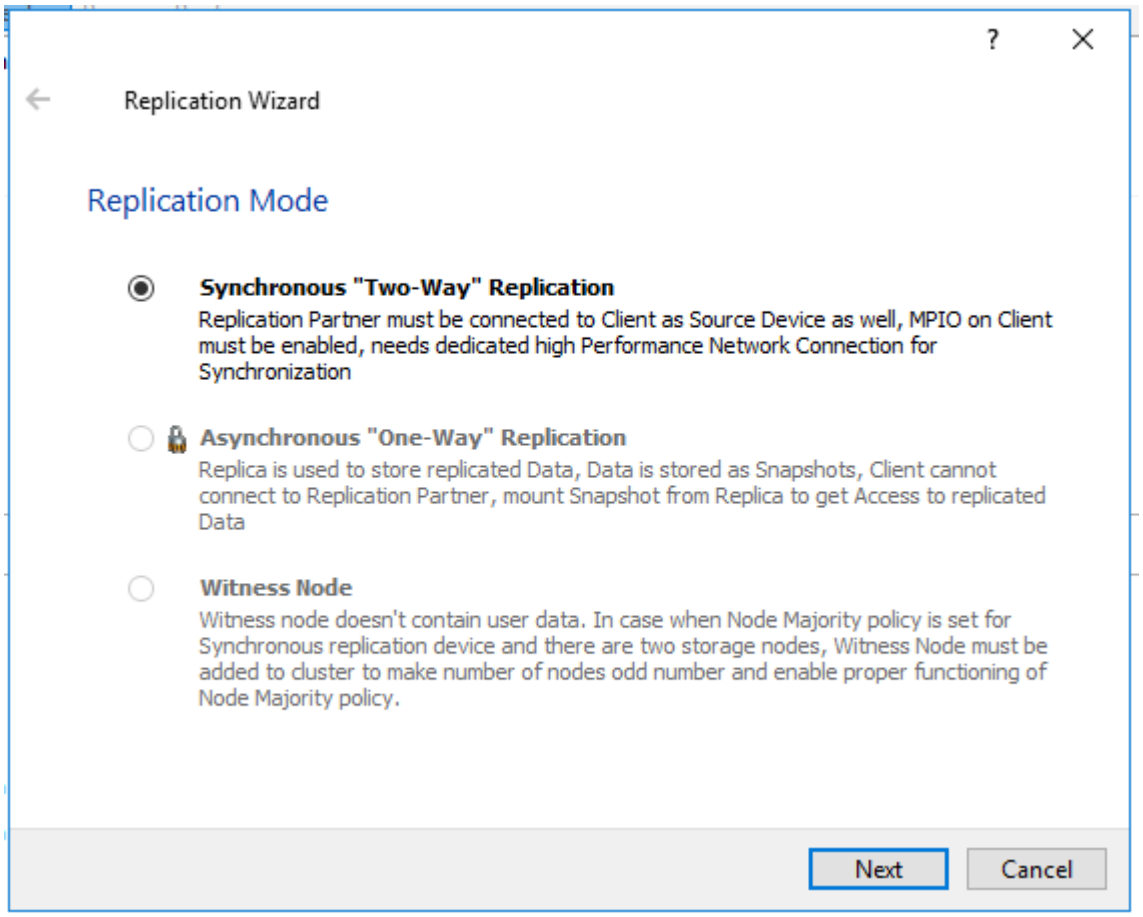
11. Click Finish to close the wizard.

12. Right-click on the recently created device and select Replication Manager from the shortcut menu.

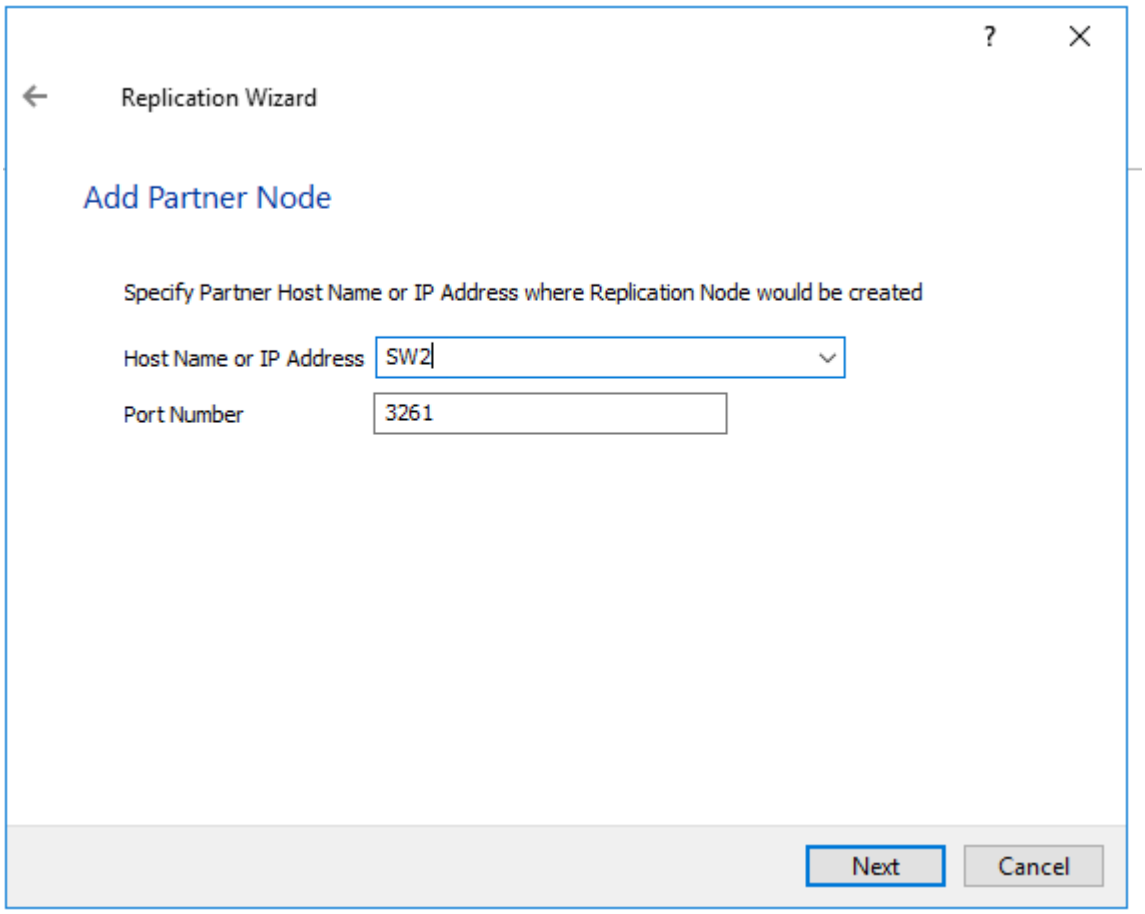
13. Click Add replica.



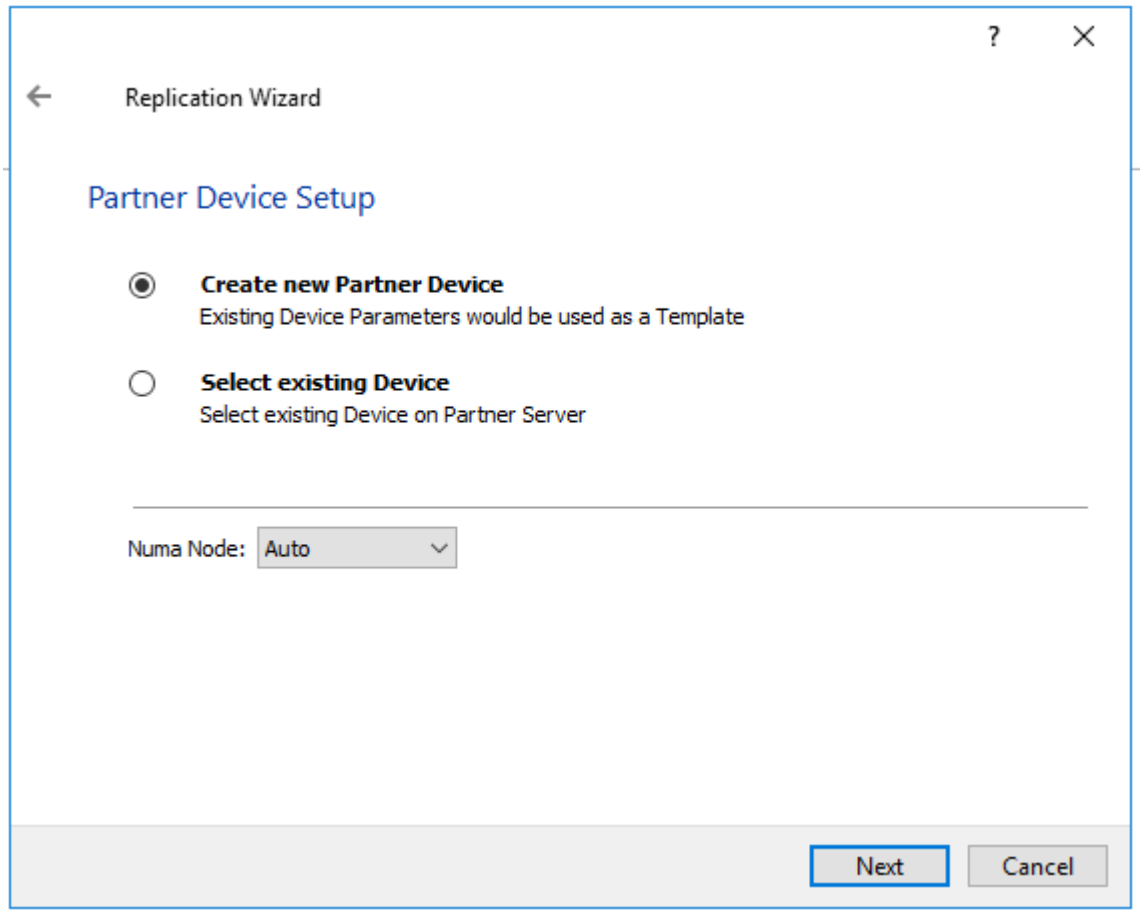
14. Select Synchronous “Two-Way” Replication as a replication mode.



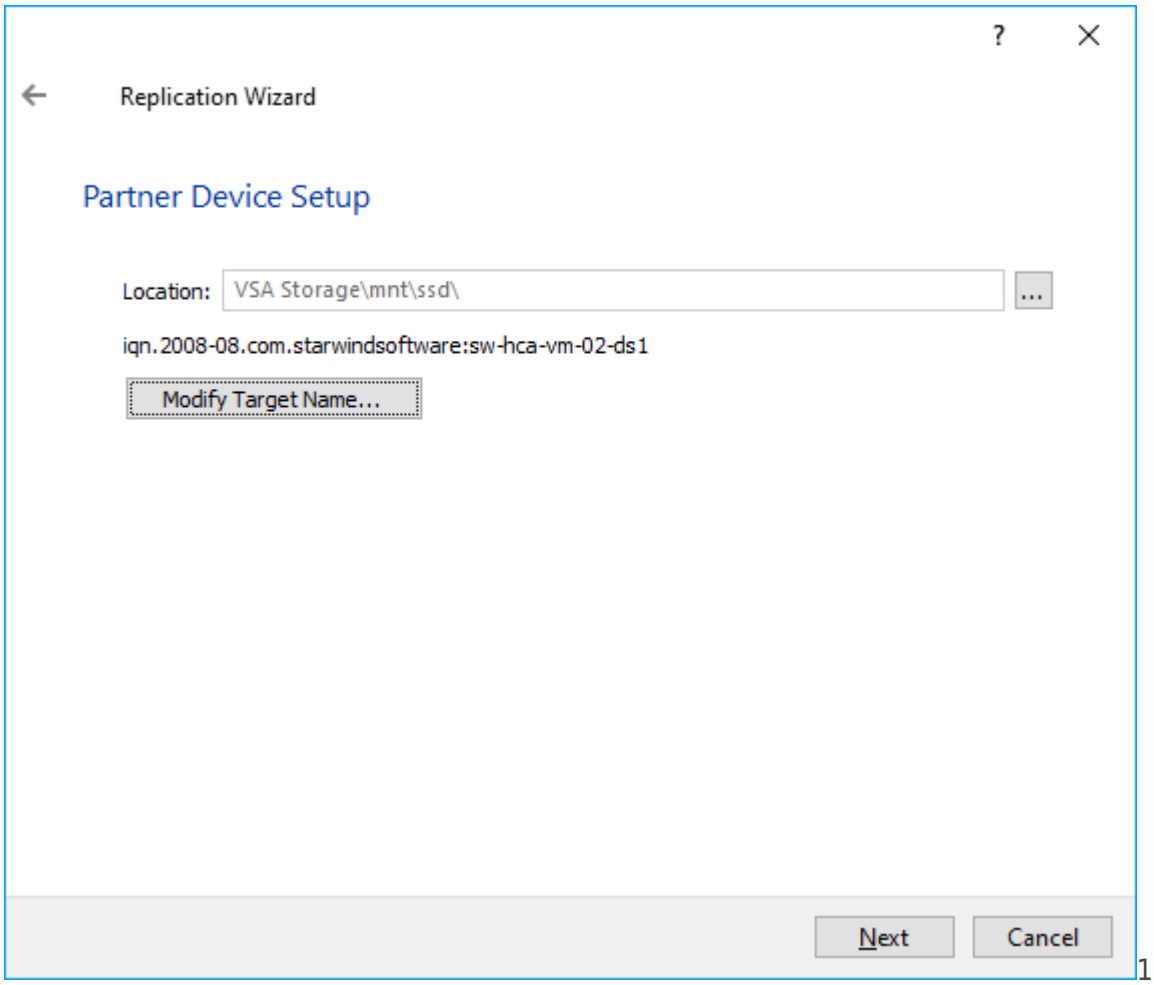
15. Specify Partner Hostname or IP Address and Port Number. Click Next.



16. Choose Create new Partner Device and click Next.



17. Choose the device Location and specify the target name if required. Otherwise, the name is generated automatically in accordance with the specified target alias.



18. Select Synchronization Journal Strategy and click Next.

NOTE: There are several options – RAM-based journal (default) and Disk-based journal with failure and continuous strategy, that allow to avoid full synchronization cases.

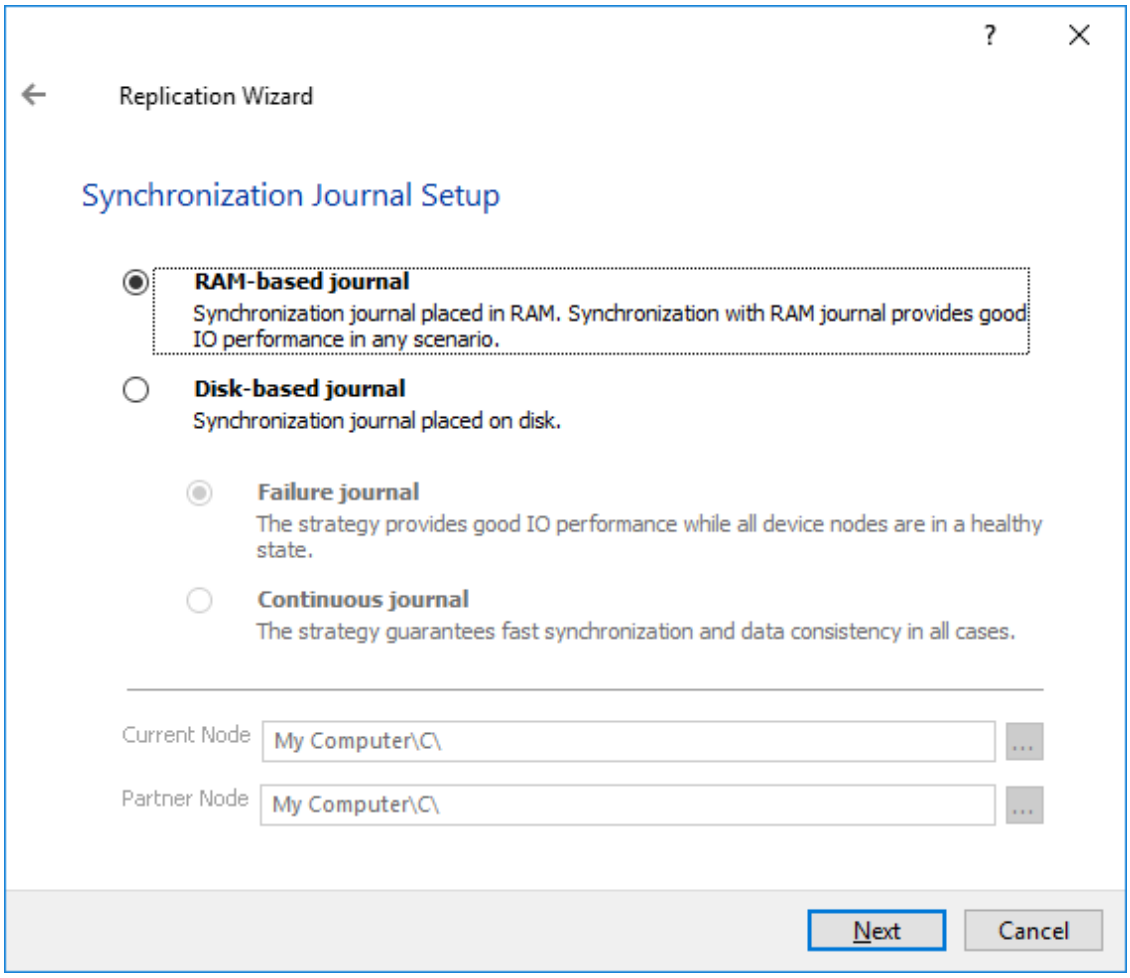
RAM-based (default) synchronization journal is placed in RAM. Synchronization with RAM journal provides good I/O performance in any scenario. Full synchronization could occur in the cases described in this KB:
https://knowledgebase.starwindsoftware.com/explanation/reasons-why-full-synchronization-may-start/

Disk-based journal placed on a separate disk from StarWind devices. It allows to avoid full synchronization for the devices where it's configured even when StarWind service is being stopped on all nodes.
Disk-based synchronization journal should be placed on a separate, preferably faster disk from StarWind devices. SSDs and NVMe disks are recommended as the device performance is defined by the disk speed, where the journal is located. For example, it can be placed on the OS boot volume.
It is required to allocate 2 MB of disk space for the synchronization journal per 1 TB of HA

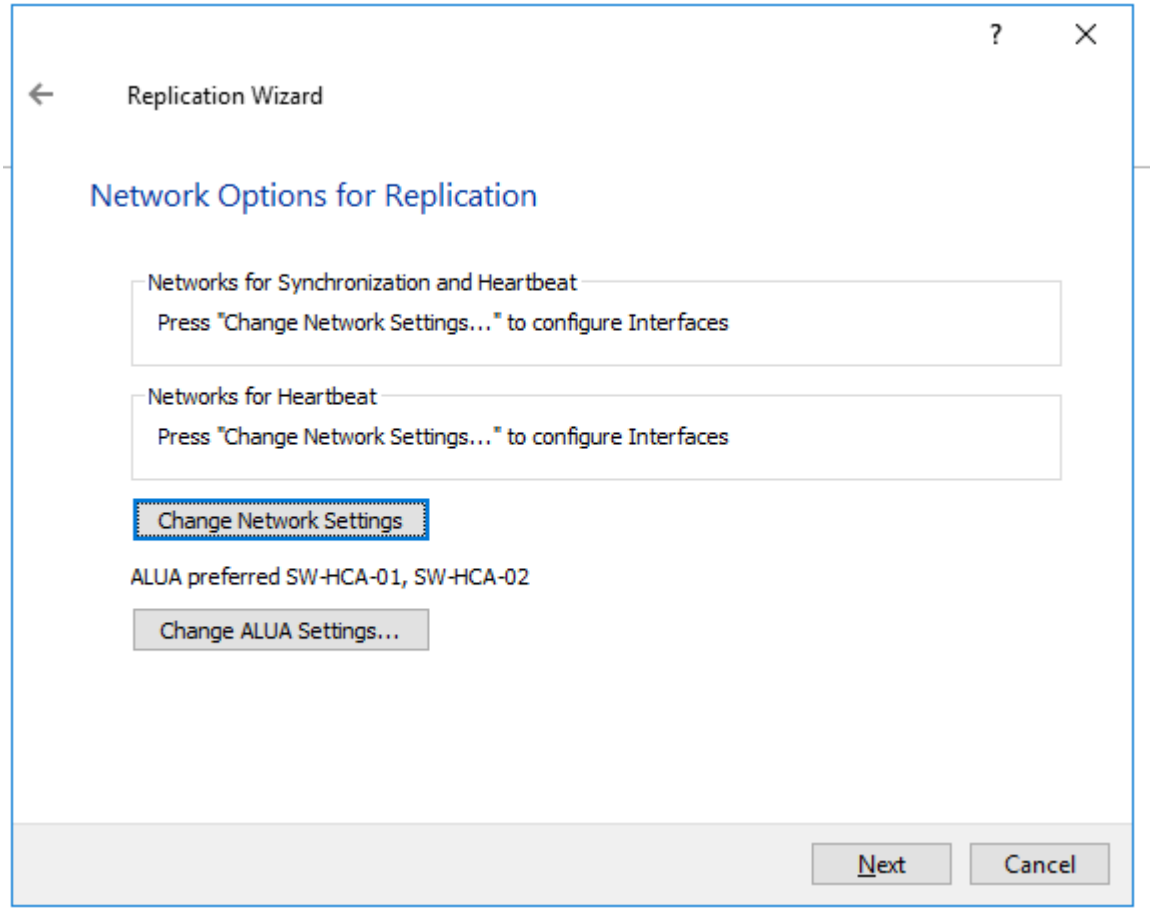
device size with a disk-based journal configured and 2-way replication and 4MB per 1 TB of HA device size for 3-way replication.

Failure journal – provides good I/O performance, as a RAM-based journal, while all device nodes are in a healthy synchronized state. If a device on one node went into a not synchronized state, the disk-based journal activates and a performance drop could occur as the device performance is defined by the disk speed, where the journal is located. Fast synchronization is not guaranteed in all cases. For example, if a simultaneous hard reset of all nodes occurs, full synchronization will occur.

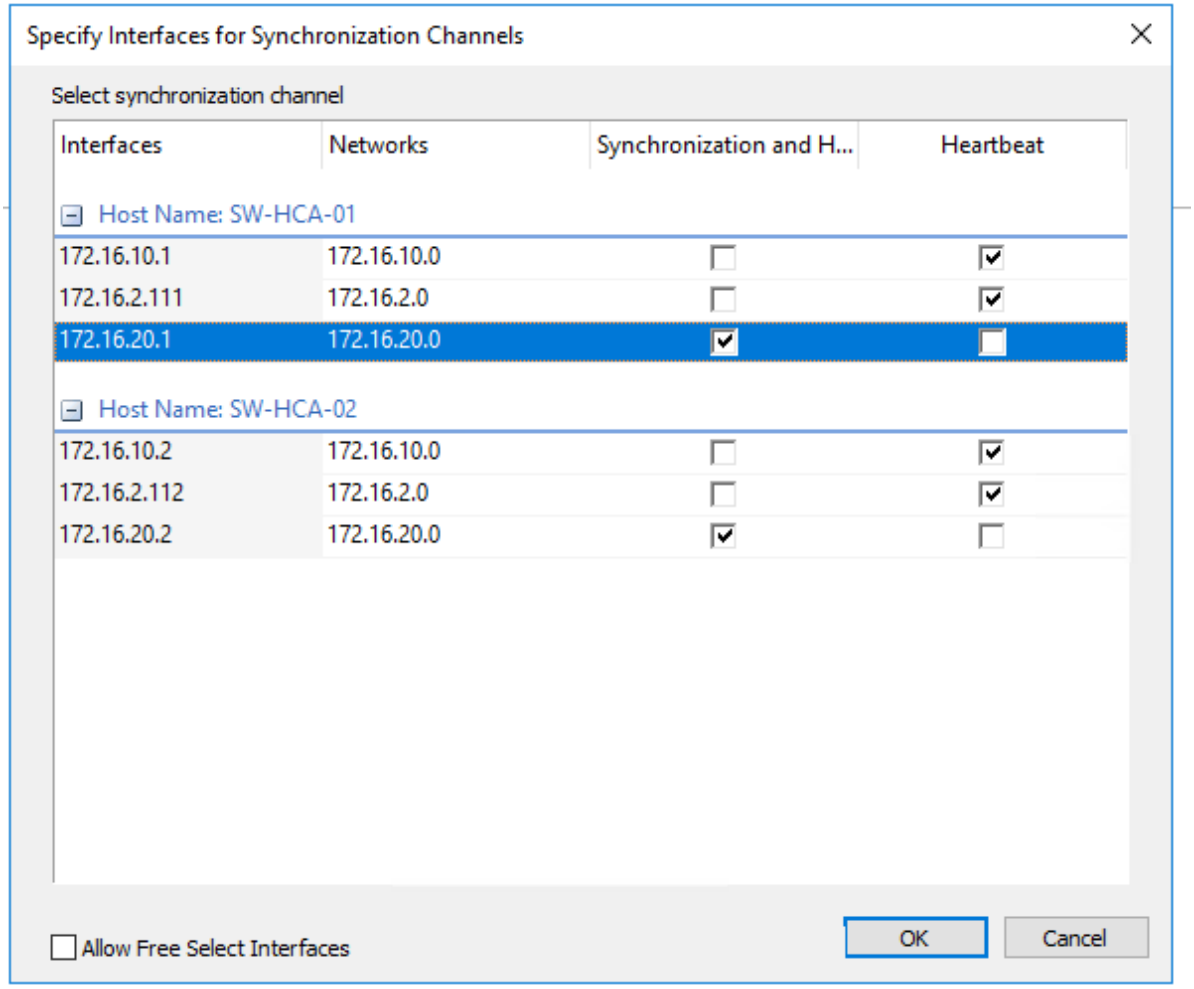
Continuous journal – guarantees fast synchronization and data consistency in all cases. Although, this strategy has the worst I/O performance, because of frequent write operations to the journal, located on the disk, where the journal is located.



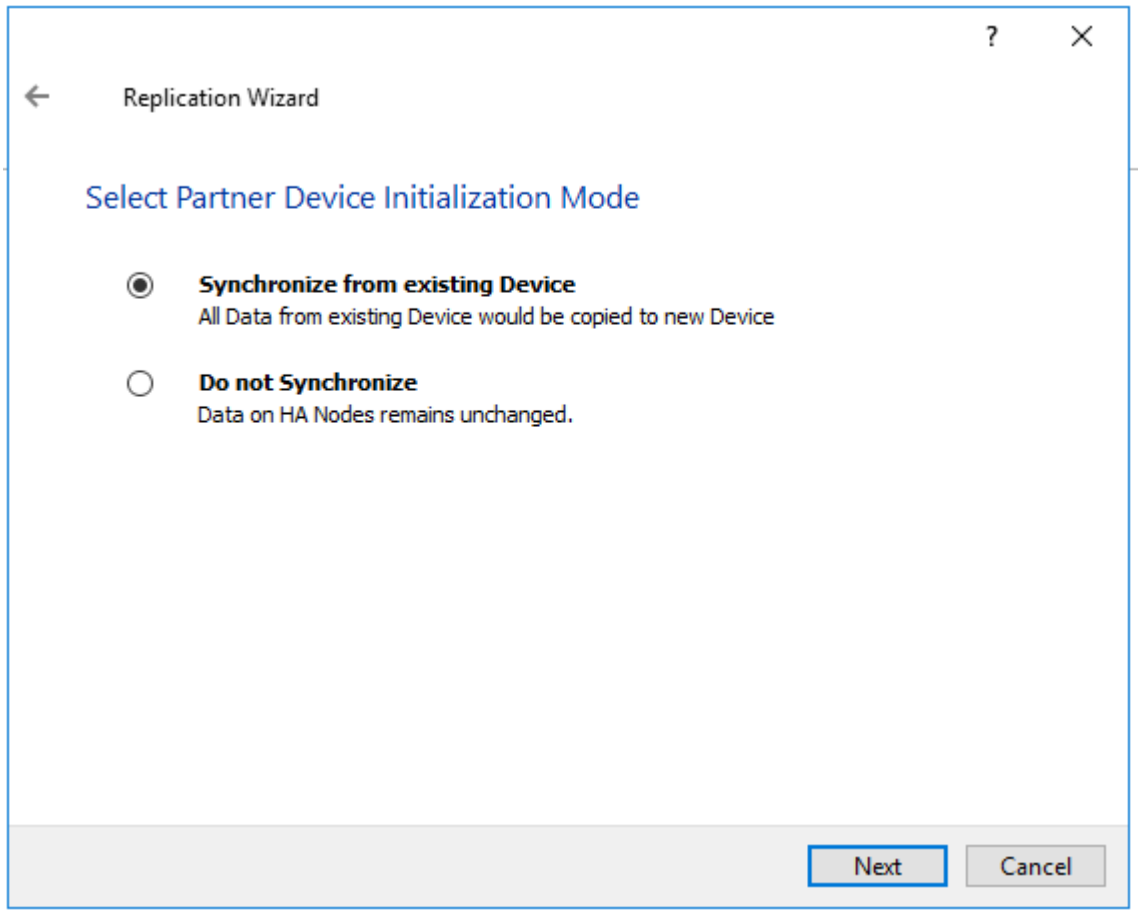
19. Click Change Network Settings.



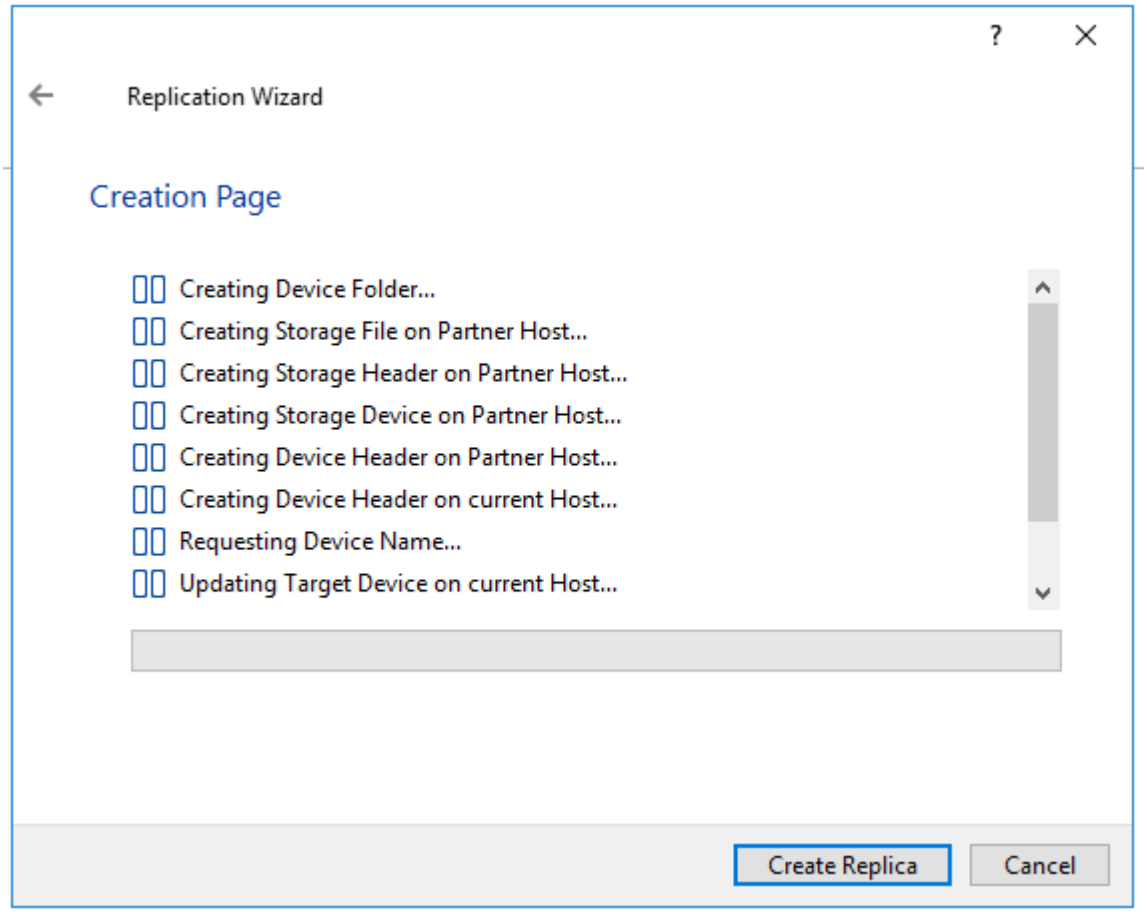
20. Specify the interfaces for Synchronization and Heartbeat channels. Click OK. Then click Next.



21. Choose Synchronize from existing Device

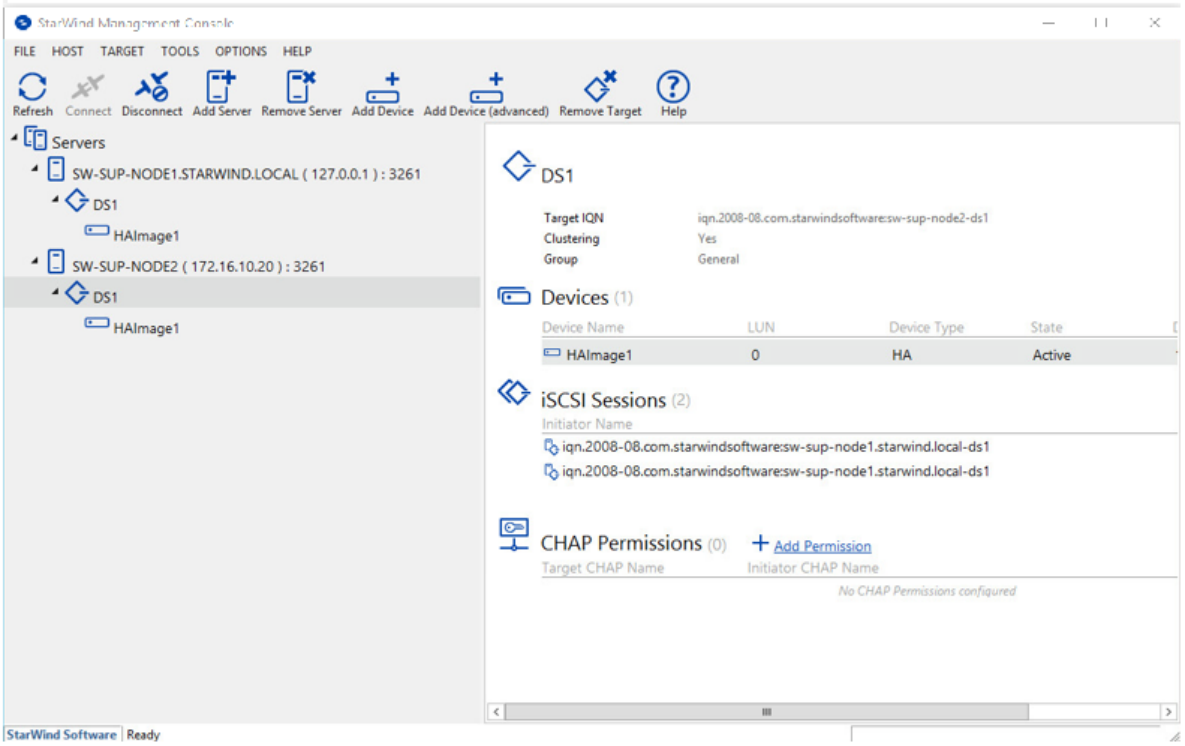


22. Click Create Replica.

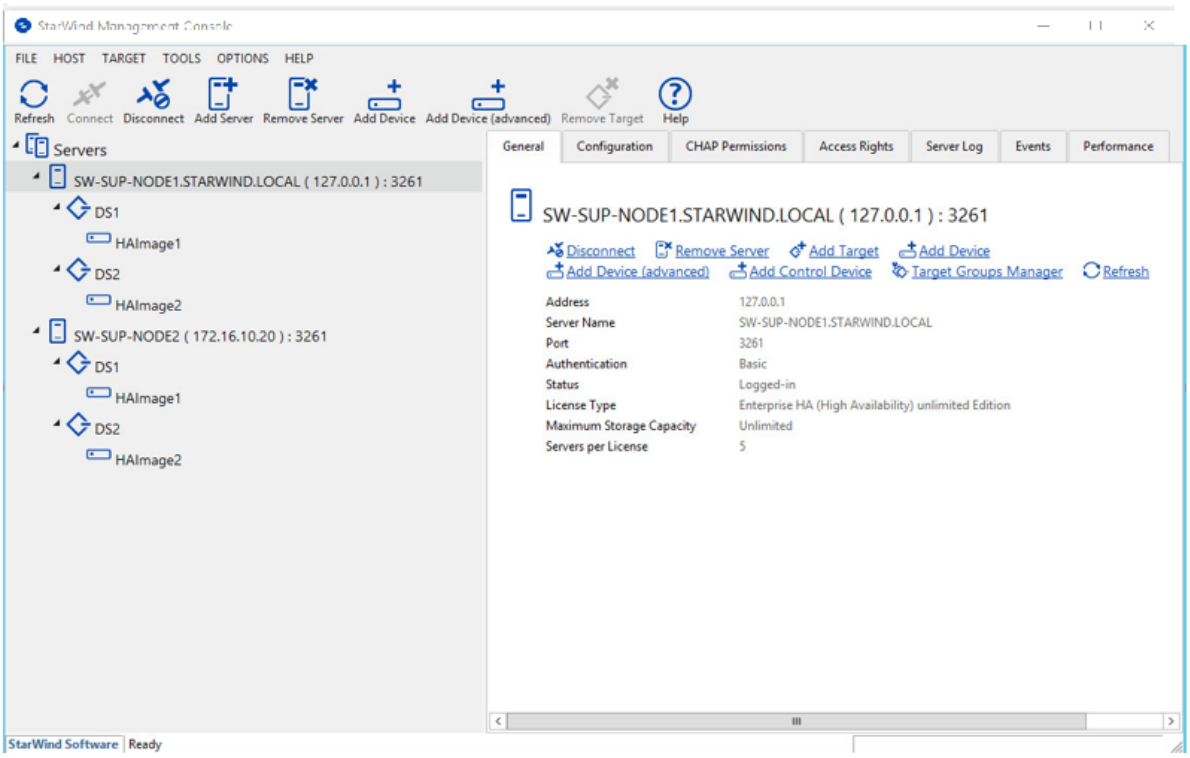


Click Finish to close the wizard.

23. The successfully added devices appear in the StarWind Management Console.



24. Follow the similar procedure to create other virtual disks that will be used as datastores.



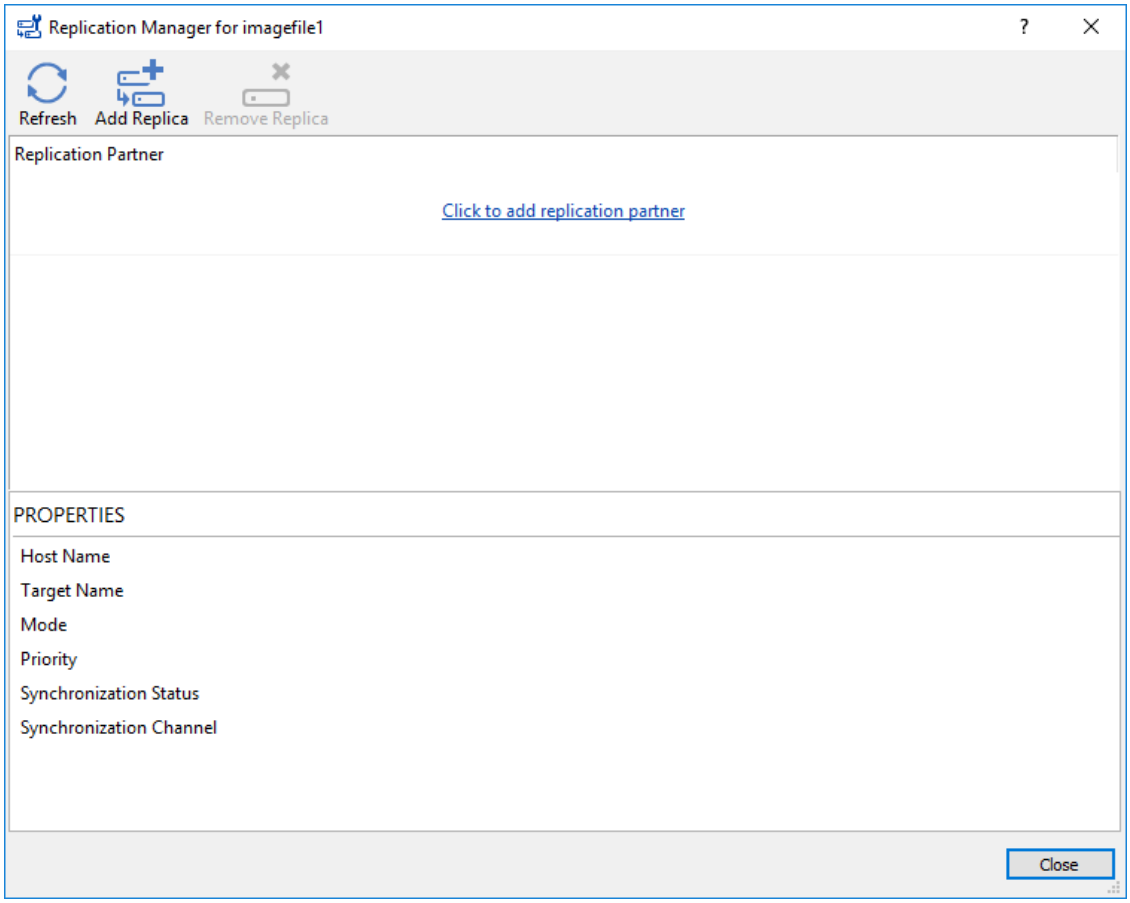
Select The Required Replication Mode

The replication can be configured using Synchronous “Two-Way” Replication mode: Synchronous or active-active replication ensures real-time synchronization and load balancing of data between two or three cluster nodes. Such a configuration tolerates the failure of two out of three storage nodes and enables the creation of an effective business continuity plan. With synchronous mirroring, each write operation requires control confirmation from both storage nodes. It guarantees the reliability of data transfers but is demanding in bandwidth since mirroring will not work on high-latency networks.

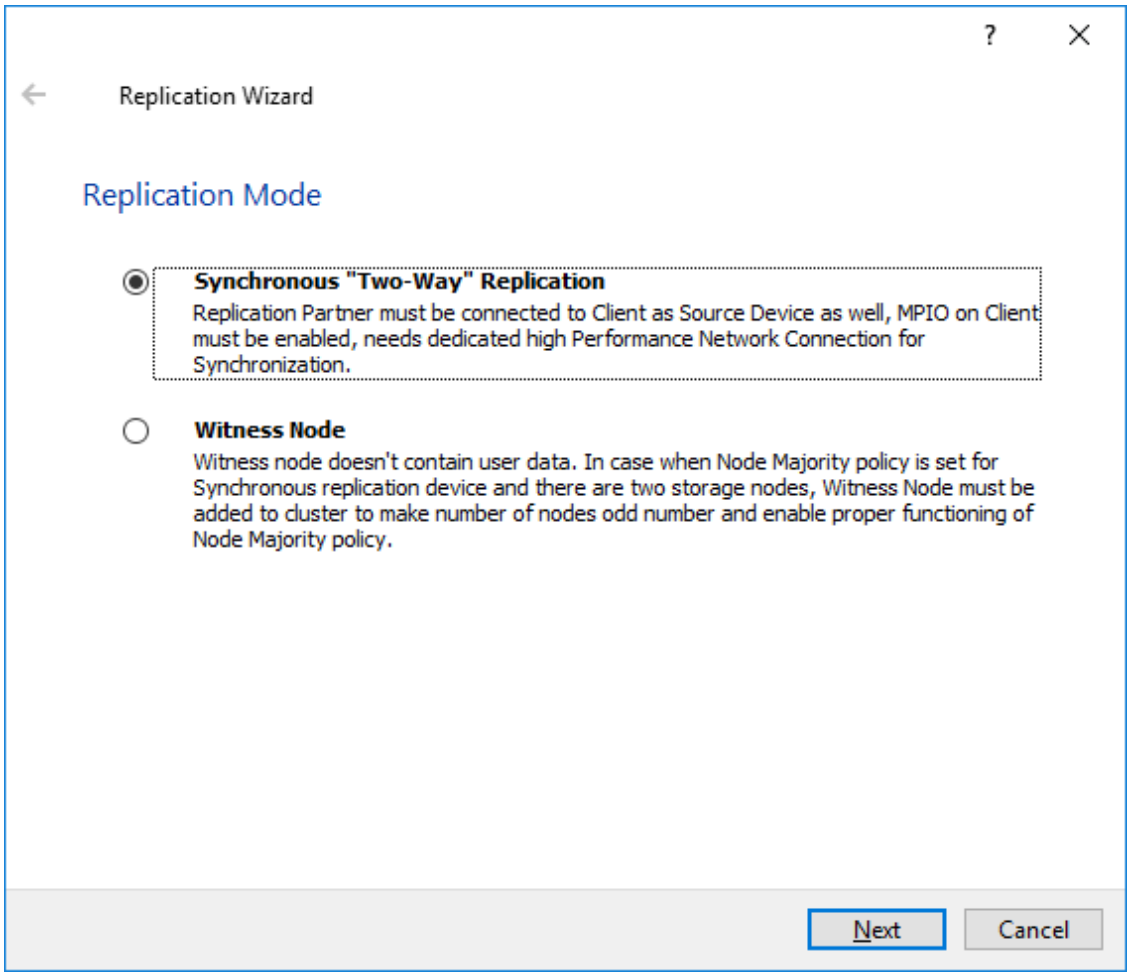
Synchronous “Two-Way” Replication

1. Right-click the recently created device and select Replication Manager from the shortcut menu.

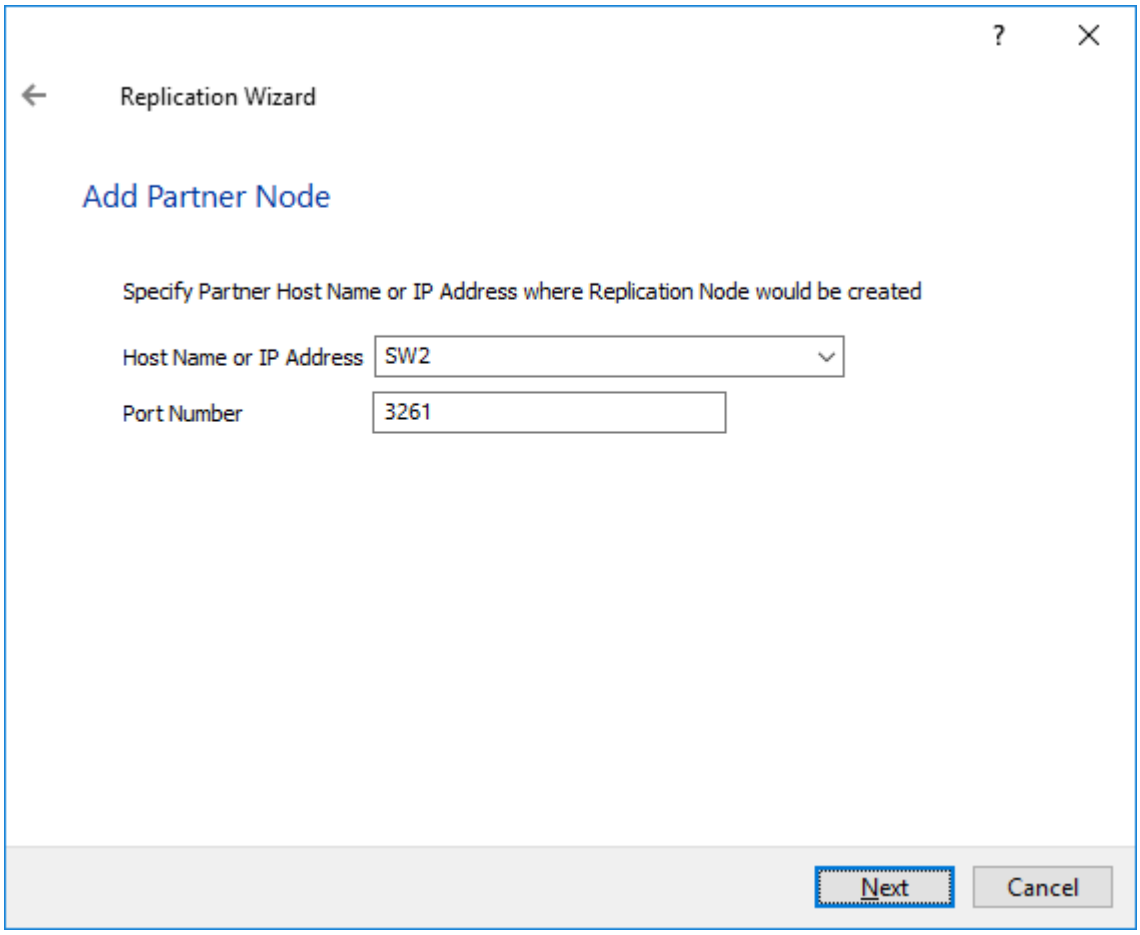
2. Select the Add Replica button in the top menu.



3. Select Synchronous “Two-Way” replication as a replication mode.



4. Specify a partner Host name or IP address and Port Number.



Selecting The Failover Strategy

StarWind provides 2 options for configuring a failover strategy:

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

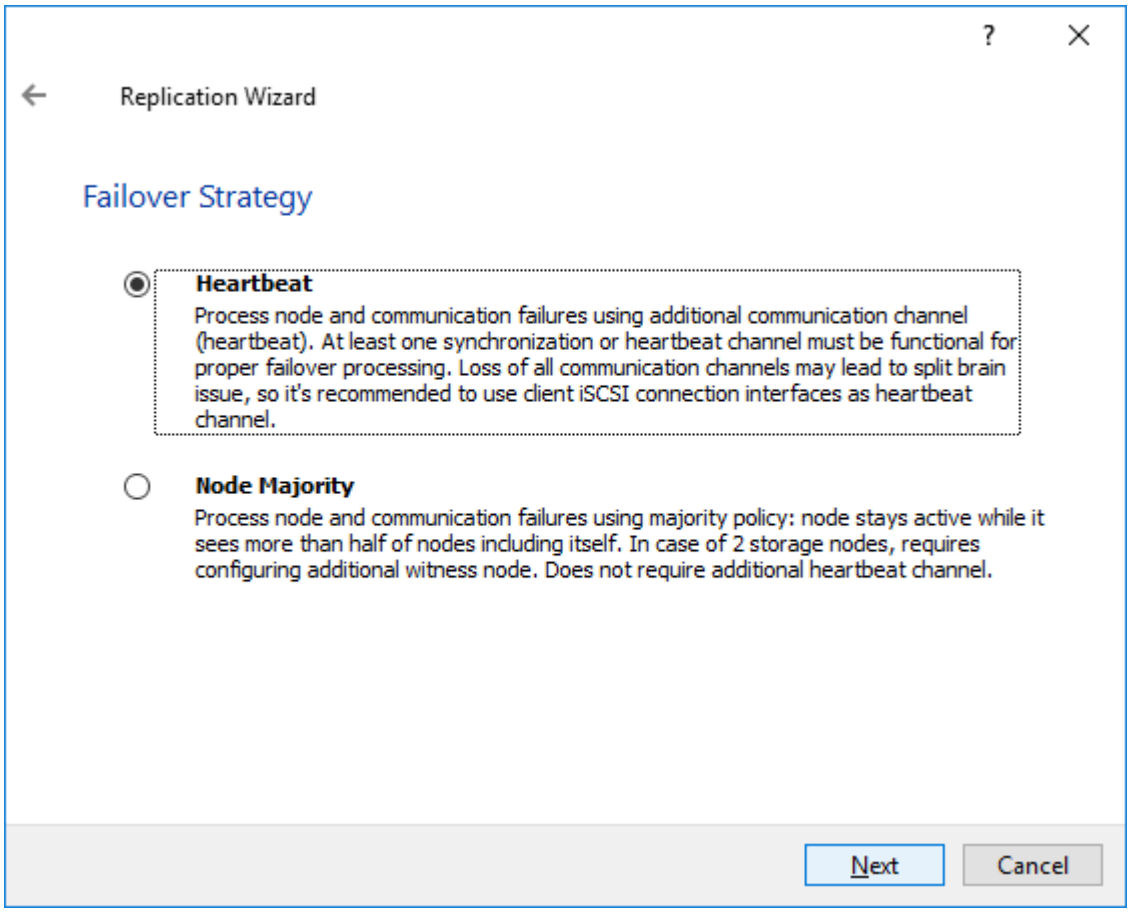
heartbeat channels during the replica creation to improve system stability and avoid the “split-brain” issue.
With the heartbeat failover strategy, the storage cluster will continue working with only one StarWind node available.

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:

Heartbeat

1. Select Failover Strategy.



2. Select Create new Partner Device and click Next.

3. Select a partner device Location.

4. Select Synchronization Journal Strategy and click Next.
NOTE: There are several options – RAM-based journal (default) and Disk-based journal with failure and continuous strategy, that allow to avoid full synchronization cases.

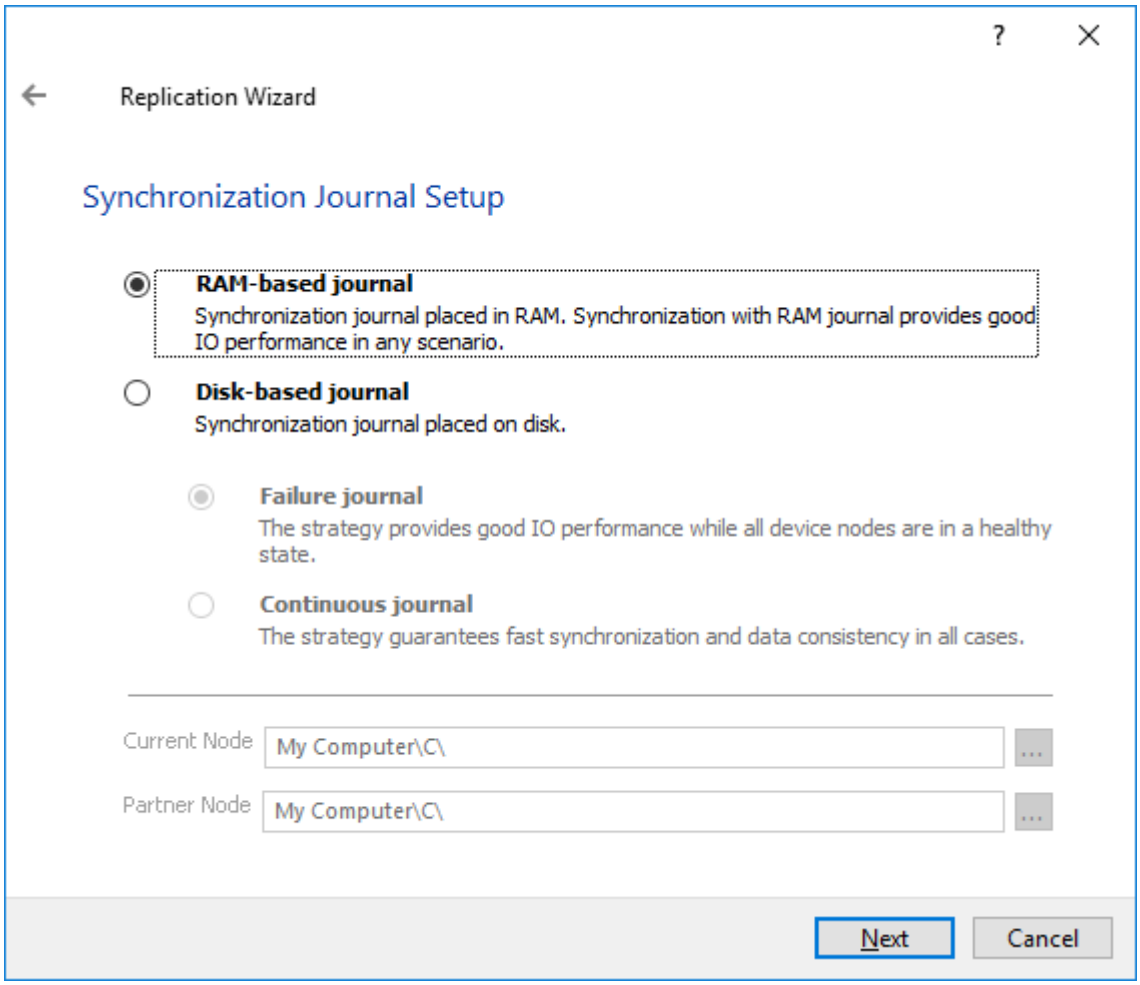
RAM-based (default) synchronization journal is placed in RAM. Synchronization with RAM journal provides good I/O performance in any scenario. Full synchronization could occur in the cases described in this KB:
https://knowledgebase.starwindsoftware.com/explanation/reasons-why-full-synchronization-may-start/

Disk-based journal placed on a separate disk from StarWind devices. It allows to avoid full synchronization for the devices where it's configured even when StarWind service is being stopped on all nodes. Disk-based synchronization journal should be placed on a separate, preferably faster disk from StarWind devices. SSDs and NVMe disks are recommended as the device performance is defined by the disk speed, where the journal is located. For example, it can be placed on the OS boot volume.

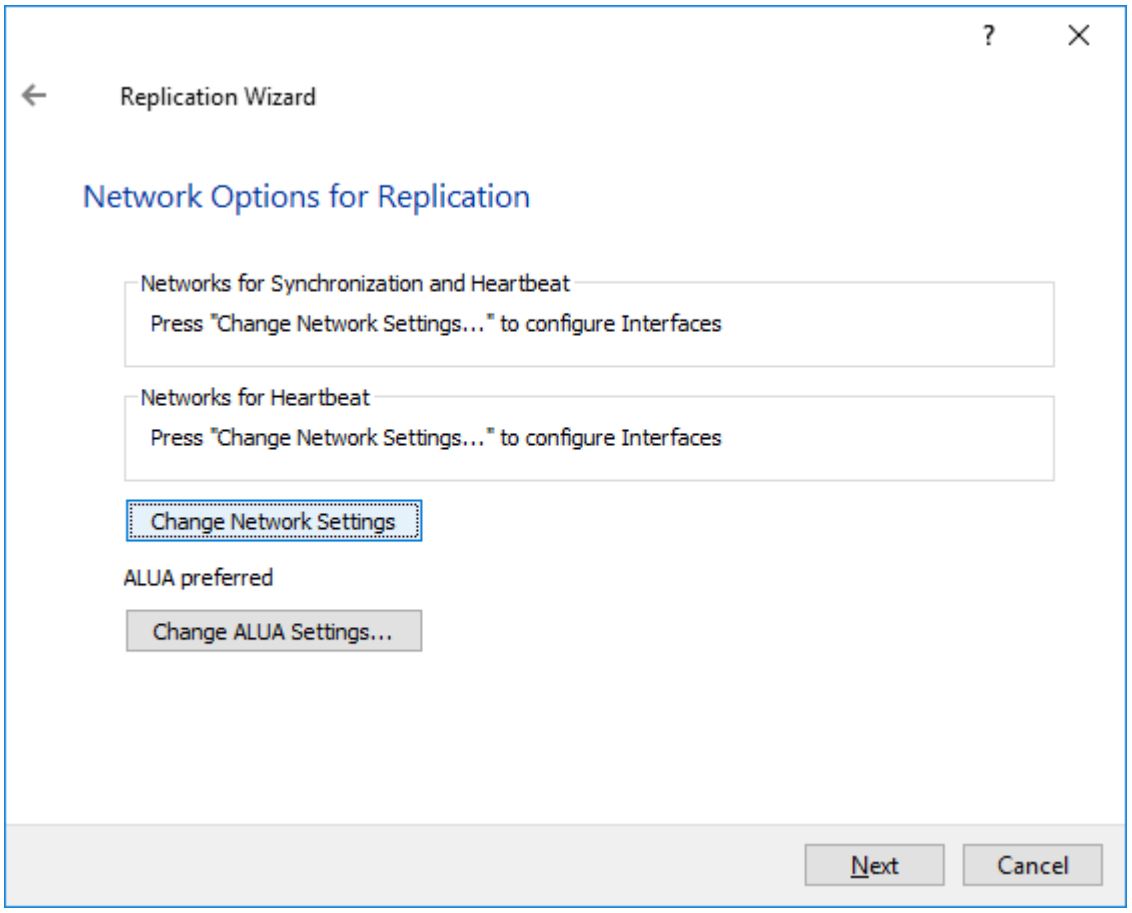
It is required to allocate 2 MB of disk space for the synchronization journal per 1 TB of HA device size with a disk-based journal configured with 2-way replication and 4MB per 1 TB of HA device size for 3-way replication.

Failure journal
The strategy provides good I/O performance, as a RAM-based journal, while all device nodes are in a healthy synchronized state. If a device on one node went into a not synchronized state, the disk-based journal activates and a performance drop could occur as the device performance is defined by the disk speed, where the journal is located. Fast synchronization is not guaranteed in all cases. For example, if a simultaneous hard reset of all nodes occurs, full synchronization will occur.

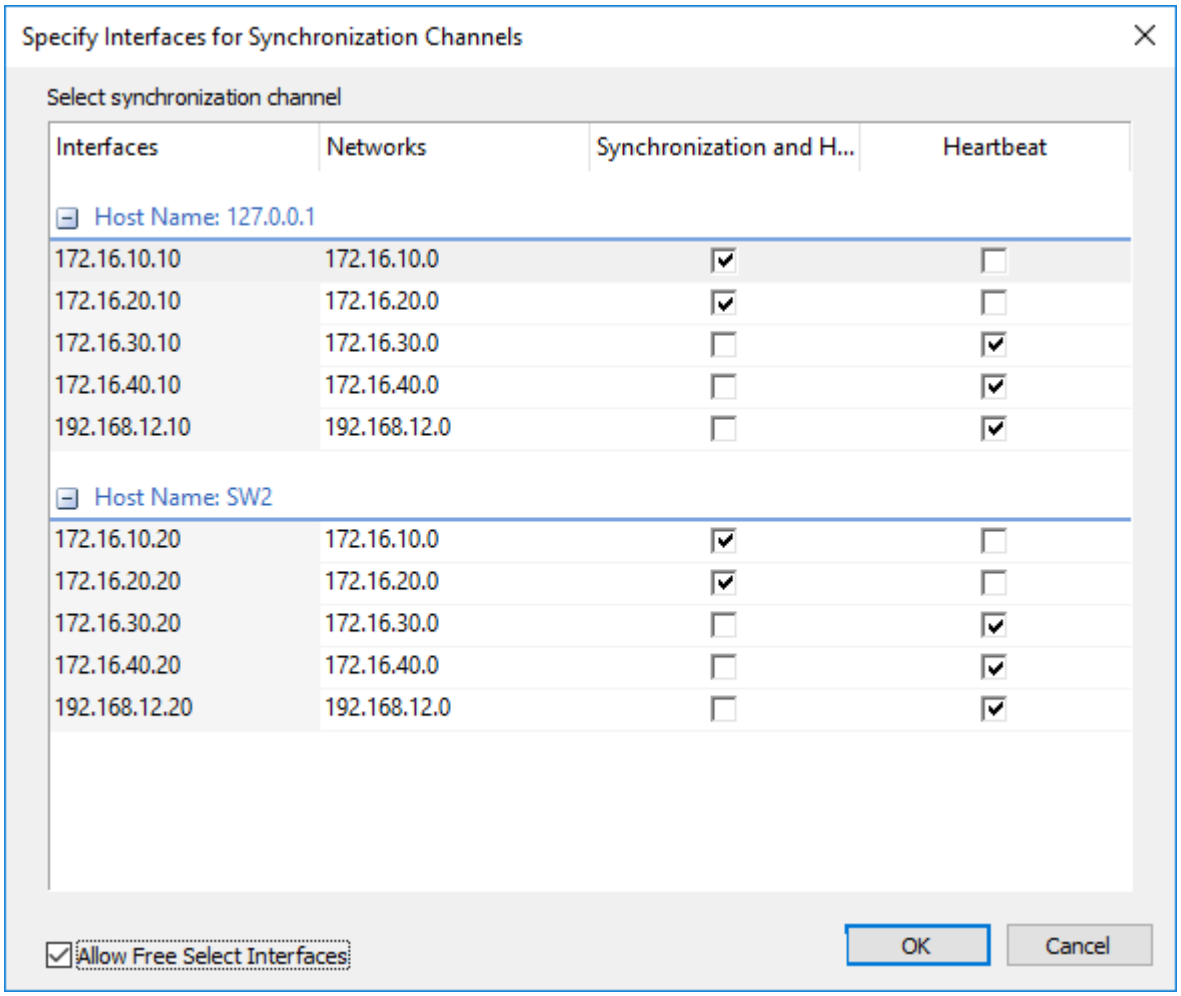
Continuous journal
The strategy guarantees fast synchronization and data consistency in all cases. Although, this strategy has the worst I/O performance, because of frequent write operations to the journal, located on the disk, where the journal is located.



5. Click Change Network Settings.



6. Specify the interfaces for Synchronization and Heartbeat Channels. Click OK and then click Next.



7. In Select Partner Device Initialization Mode, select Synchronize from existing Device and click Next.

8. Click Create Replica. Click Finish to close the wizard.
The successfully added device appears in StarWind Management Console.

9. Follow the similar procedure for the creation of other virtual disks that will be used as storage repositories.

Node Majority

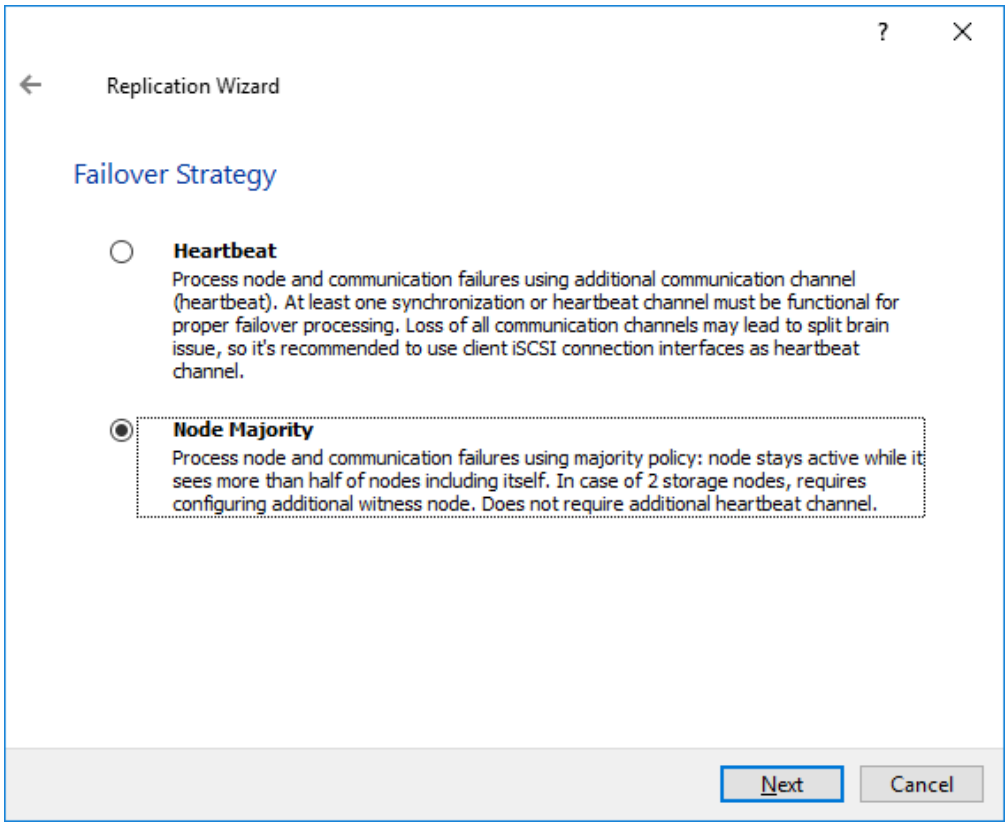
There are two ways to configure Witness for 2-nodes StarWind HA device, created with Node Majority Failover Strategy: File Share (SMB) as Witness and additional server as Witness Node.

– Creating HA device with File SHare(SMB) as Witness:

SMB Witness is a file, located on SMB share, which can be accessed by both nodes and help them to eliminate the split-brain issue in case of synchronization connection interruption between the nodes. To set up the SMB file share as a Witness for 2-nodes HA device with Node Majority Failover Strategy, perform the actions, described on this page:
https://www.starwindsoftware.com/help/ConfiguringFileShareSMBasWitness.html

– Creating HA device with Witness Node:

1. Select the Node Majority failover strategy and click Next.



2. Choose Create new Partner Device and click Next.

3. Specify the partner device Location and modify the target name if necessary. Click Next. Select Synchronization Journal strategy and location and click Next.

4. In Network Options for Replication, press the Change network settings button and select the synchronization channel for the HA device.

5. In Specify Interfaces for Synchronization Channels, select the checkboxes with the appropriate networks and click OK. Then click Next.

6. Select Synchronize from existing Device as the partner device initialization mode.

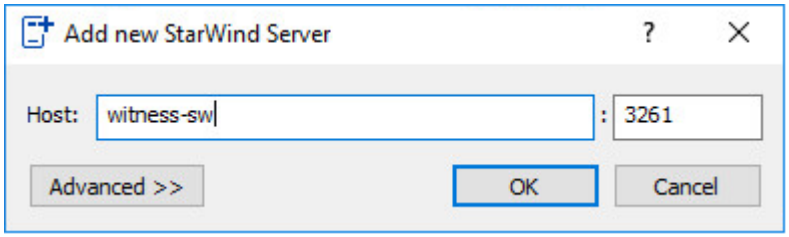
7. Press the Create Replica button and close the wizard.

8. The added devices will appear in StarWind Management Console.
Repeat the steps above to create other virtual disks if necessary.

Adding Witness Node

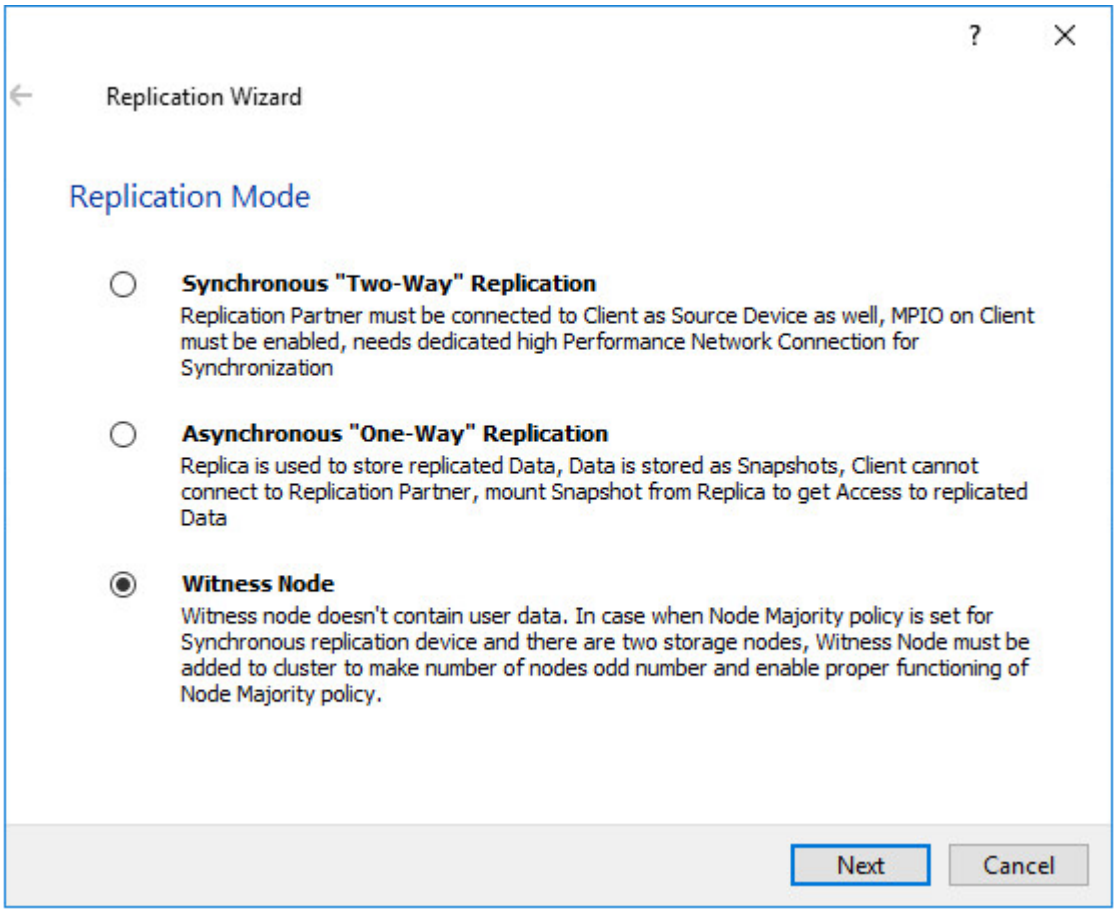
Witness node can be configured on a separate host or as a virtual machine in a cloud. It requires StarWind Virtual SAN service installed on it.
NOTE: Since the device created in this guide is replicated between 2 active nodes with the Node Majority failover strategy, a Witness node must be added to it.

1. Open StarWind Management Console, right-click on the Servers field and press the Add Server button. Add a new StarWind Server which will be used as the Witness node and click OK.

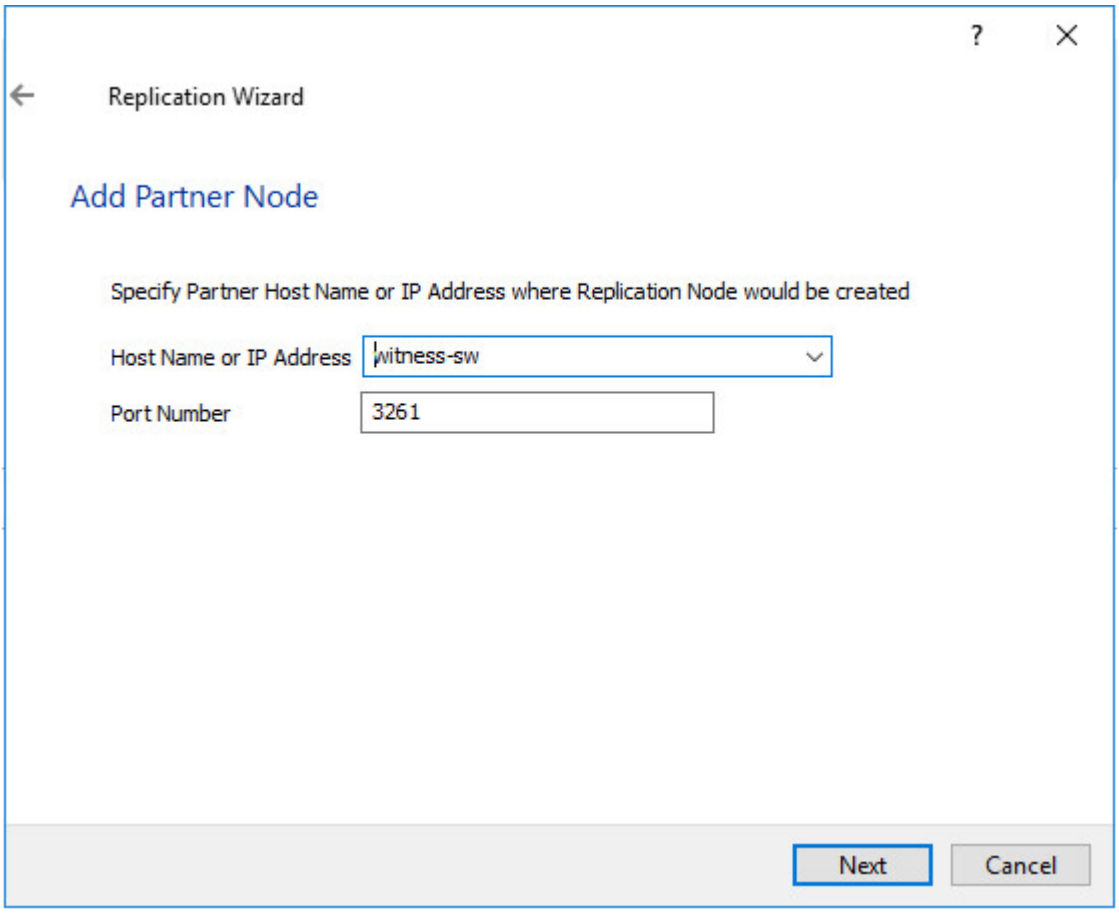


2. Right-click on the HA device with the configured Node Majority failover policy and select Replication Manager and press the Add Replica button.

3. Select Witness Node.



4. Specify the Witness node Host Name or IP address. The default Port Number is 3261.



5. In Partner Device Setup, specify the Witness device Location. Optionally, modify the target name by clicking the appropriate button.

6. In Network Options for Replication, select the synchronization channel with the Witness node by clicking the Change Network Settings button.

7. Specify the interface for Synchronization and Heartbeat and click OK.

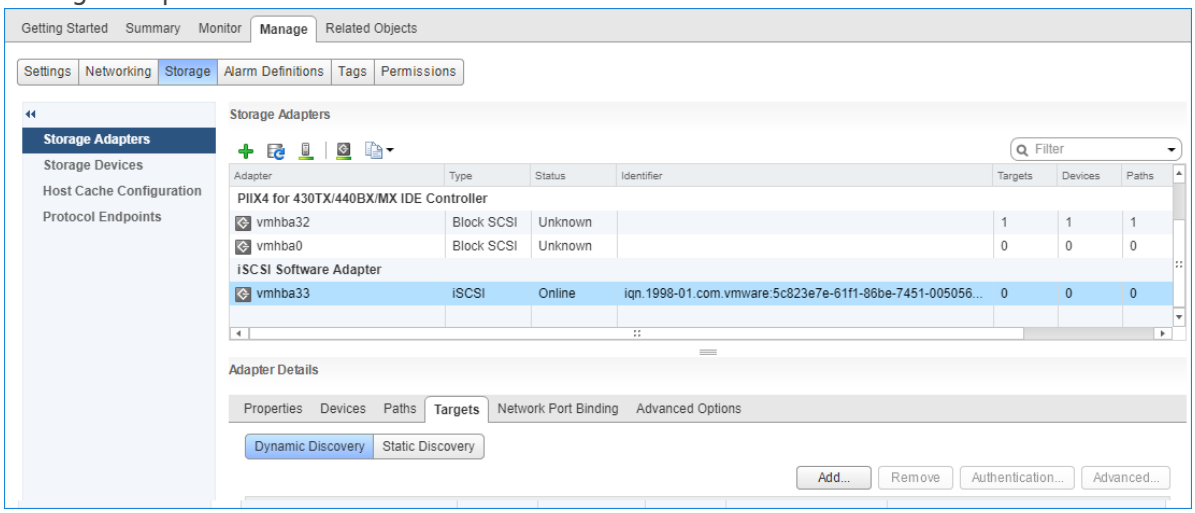
8. Click Create Replica and then close the wizard.

9. Repeat the steps above to create other virtual disks if necessary.
NOTE: To extend an Image File or a StarWind HA device to the required size, please check the article below:

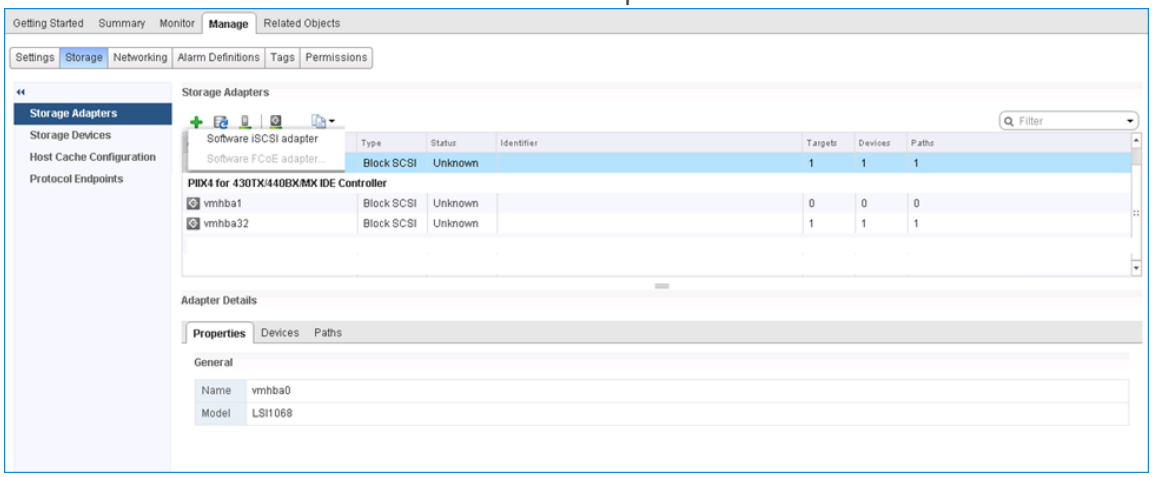
https://knowledgebase.starwindsoftware.com/maintenance/how-to-extend-image-file-or-high-availability-device/

Adding Discover Portals

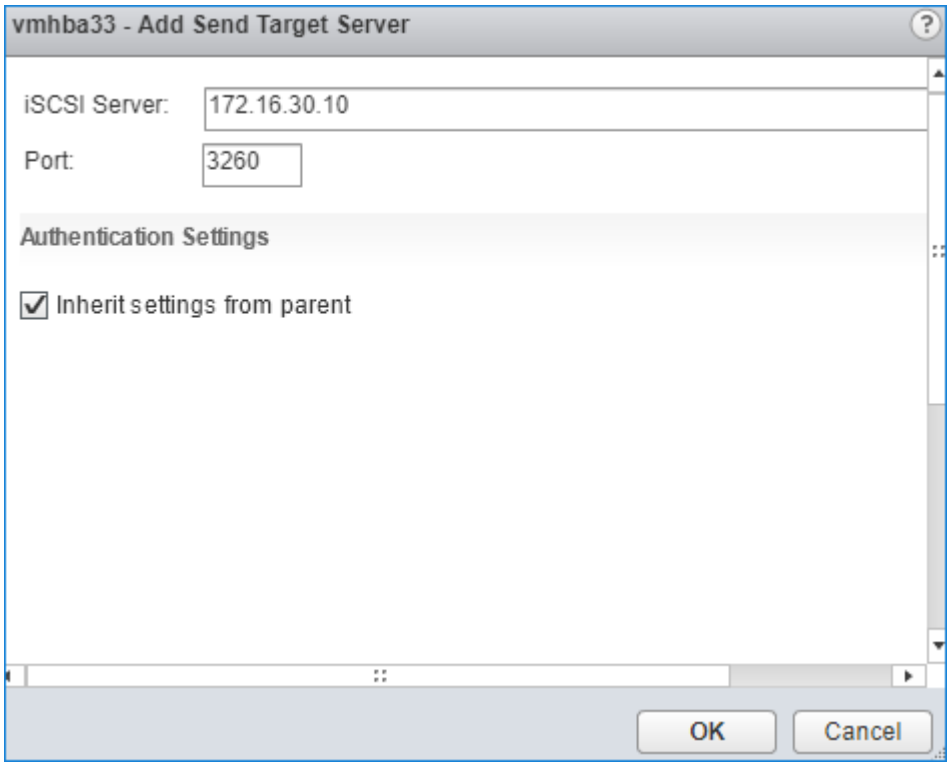
1. To connect the previously created devices to the ESXi host, click on the Storage -> Adapters -> Configure iSCSI and choose the Enabled option to enable Software iSCSI storage adapter.



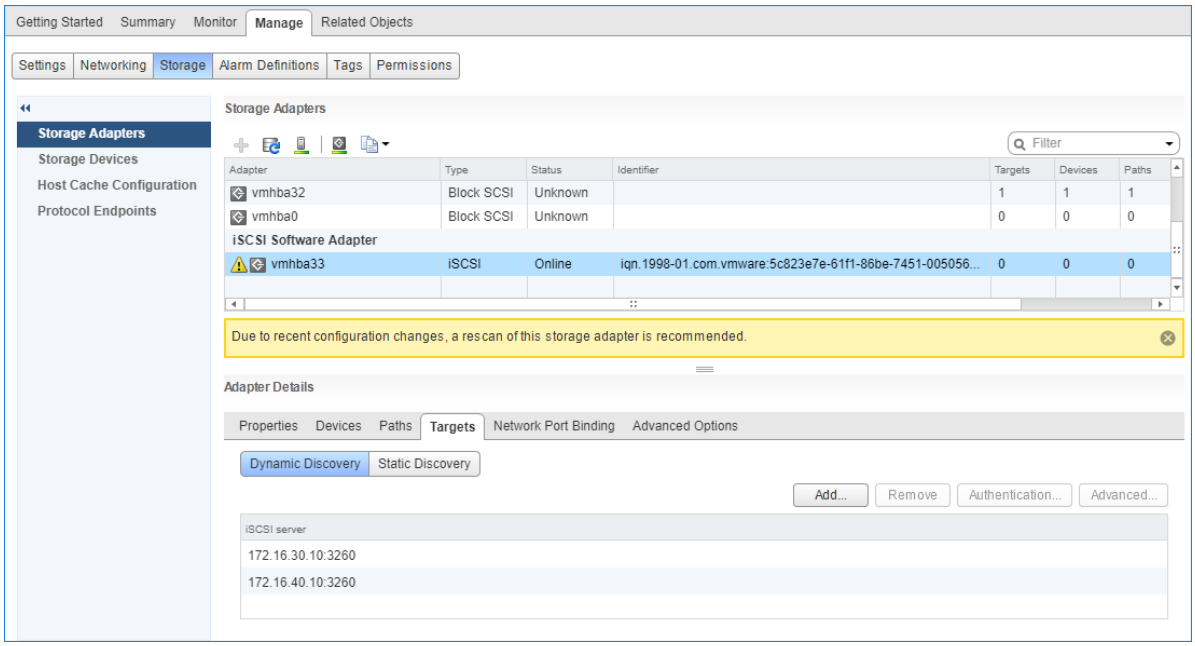
2. Click Add and select Add Software iSCSI adapter. Click OK.



3. In the Configure iSCSI window, under Dynamic Targets, click on the Add Dynamic Target button to specify iSCSI interfaces. Add dynamic targets of both storage hosts.



5. Perform the same procedure for each StarWind server by clicking Add and specifying the server IP address.



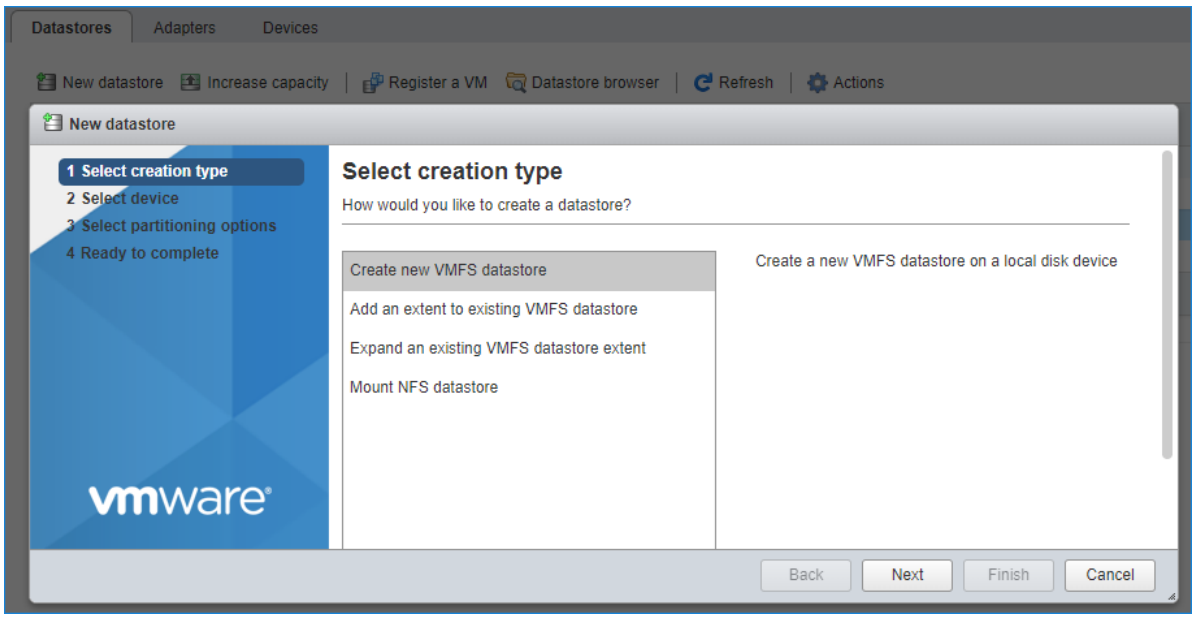
6. Select Scan for new Storage Devices and Scan for new VMFS Volumes. Then click OK.



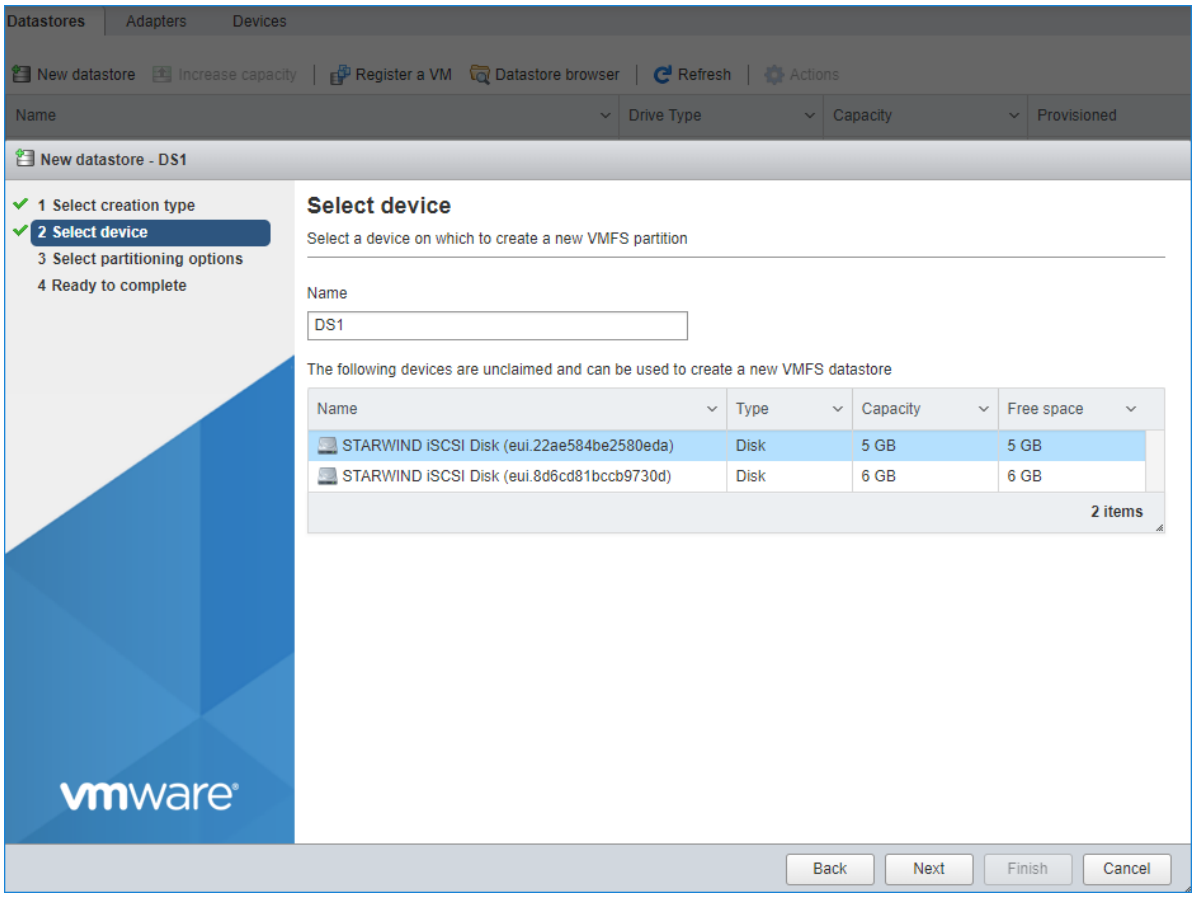
7. Repeat the steps 1-6 on the other ESXi node, specifying corresponding IP addresses for the iSCSI subnet.

Creating Datastores

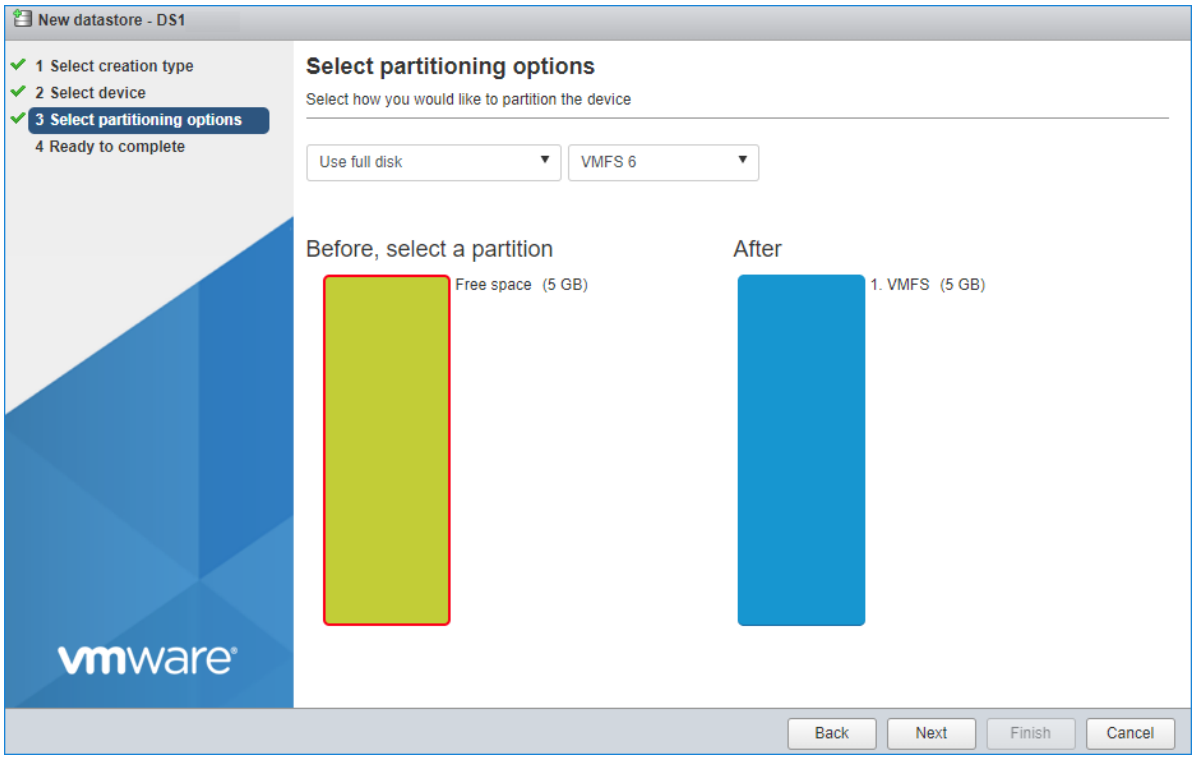
1. Right-click on the host, select New Datastore and choose VMFS.



2. Specify Datastore name and select the device to add to the datastore.

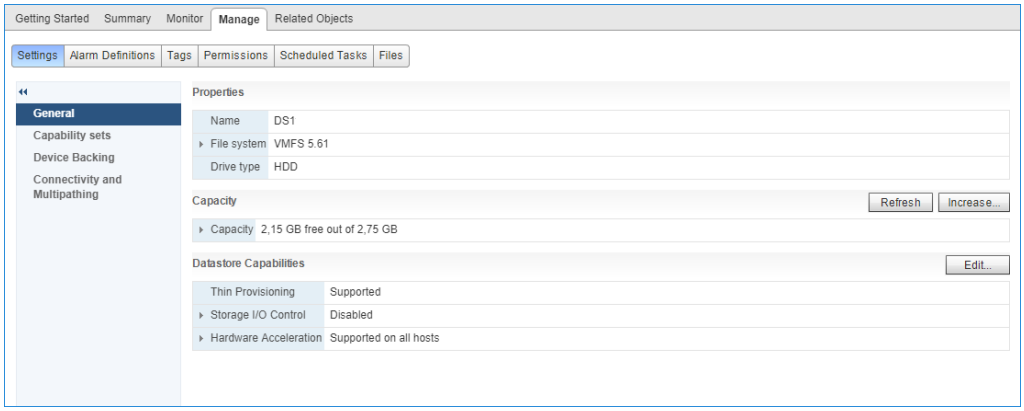


3. Enter Datastore Size and click Next.



4. In Ready to complete, verify the settings and click Finish.

5. Check another host for a new datastore.



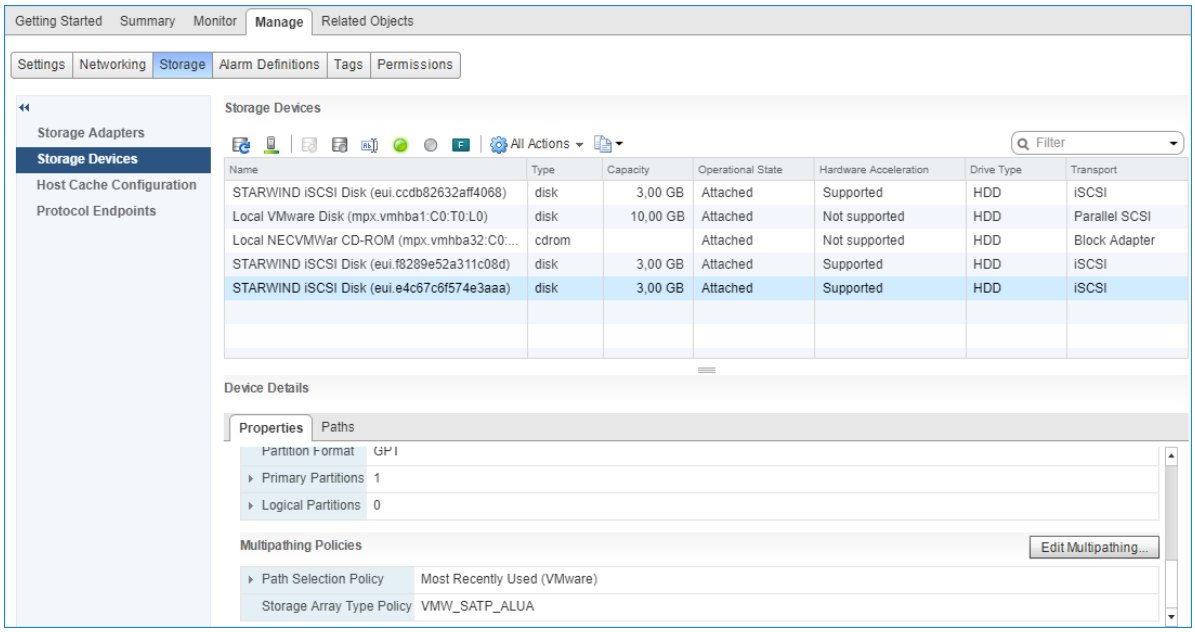
6. If the list of existing datastores does not include new datastore, click Rescan Storage...

7. Add another Datastore in the same way and select the second device for the second datastore.

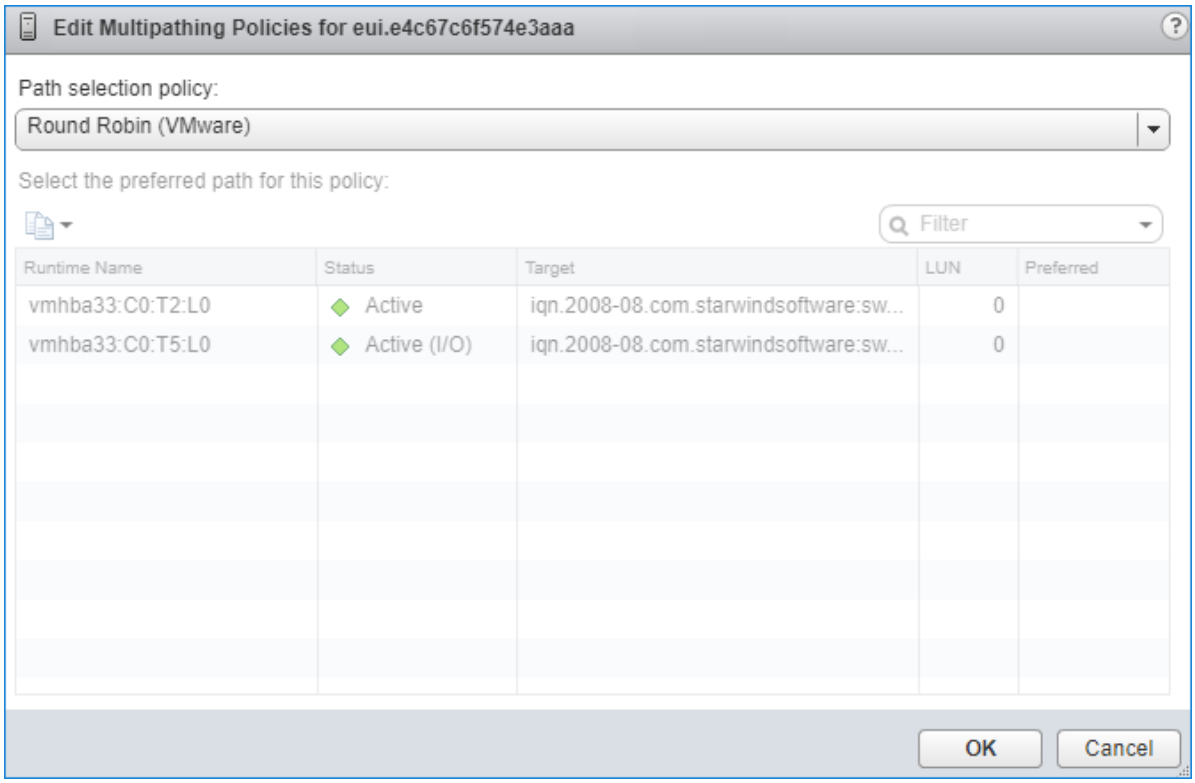
8. Verify that the storages are connected to both hosts. Otherwise, rescan the storage adapter.

9. Path selection policy for datastores changes from Most Recently Used to Round Robin automatically. To check it, click the Configure button, choose the Storage Devices tab,

select the device, and click the Edit Multipathing button.



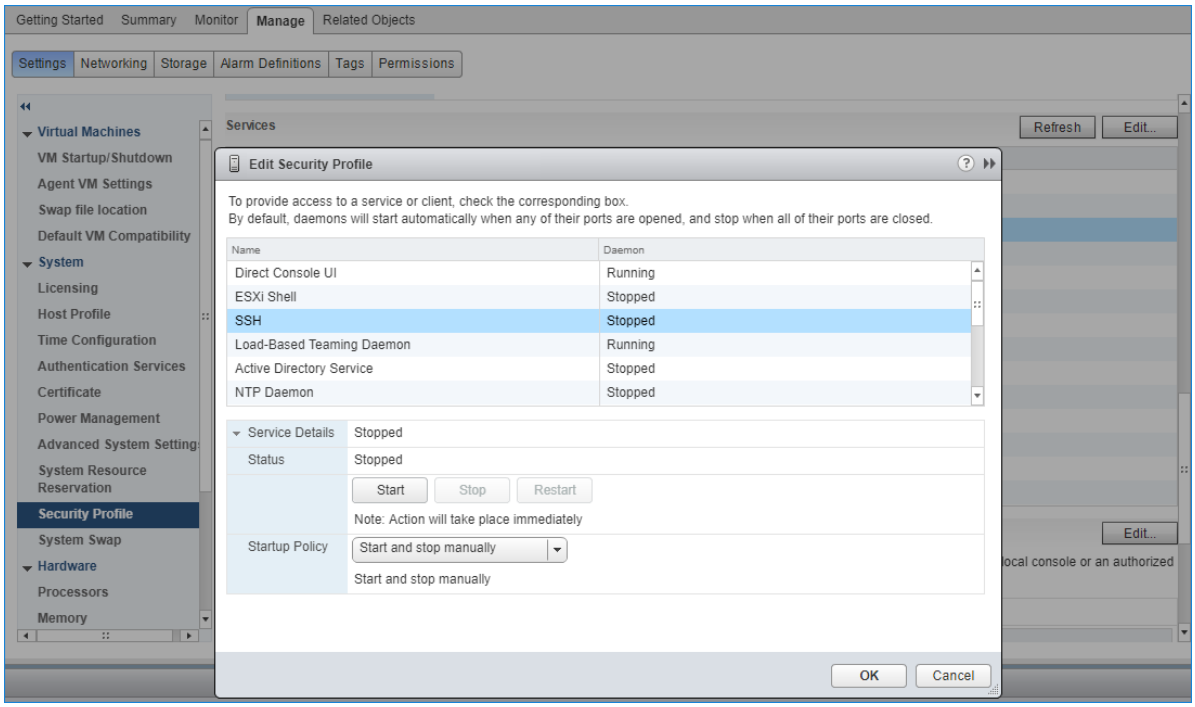
10. Select Round Robin MPIO policy.



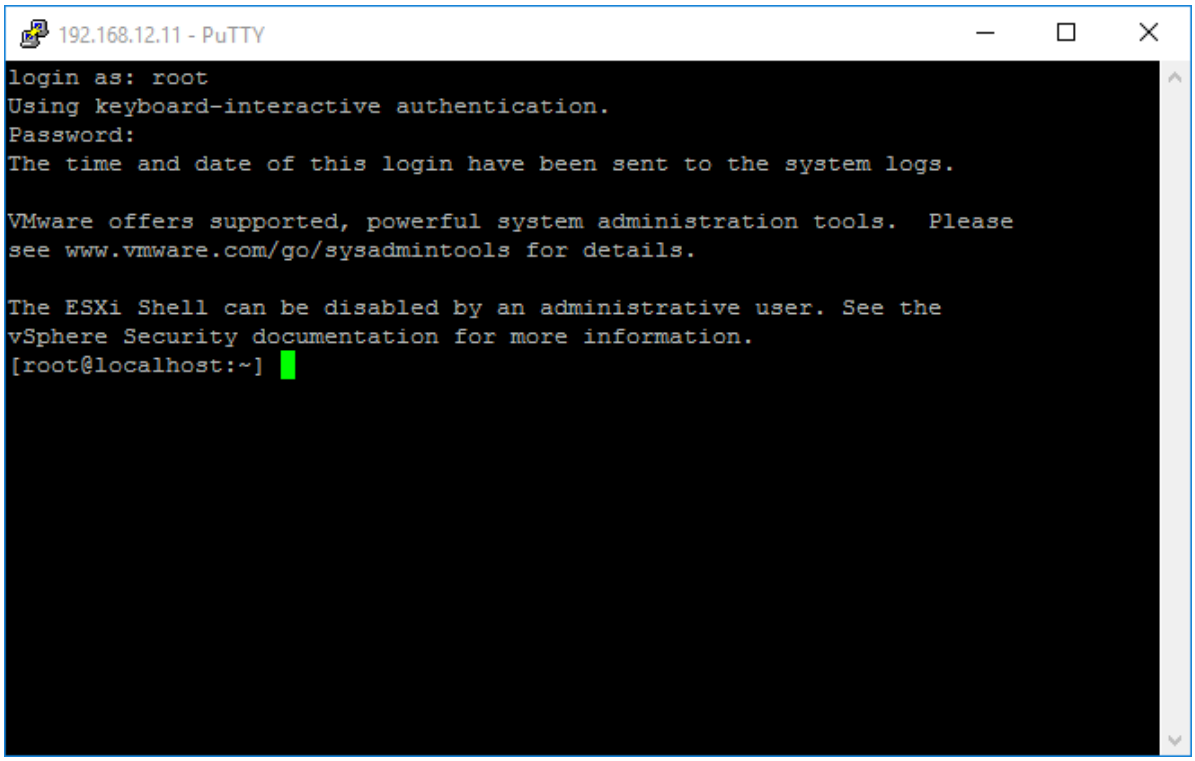
Additional Tweaks

1. Click the Manage -> Settings->Security Profile->Services->SSH->Edit.

2. Select Start and stop with host in the Startup Policy drop-down list.



3. Connect to the host using SSH client (e.g. Putty).



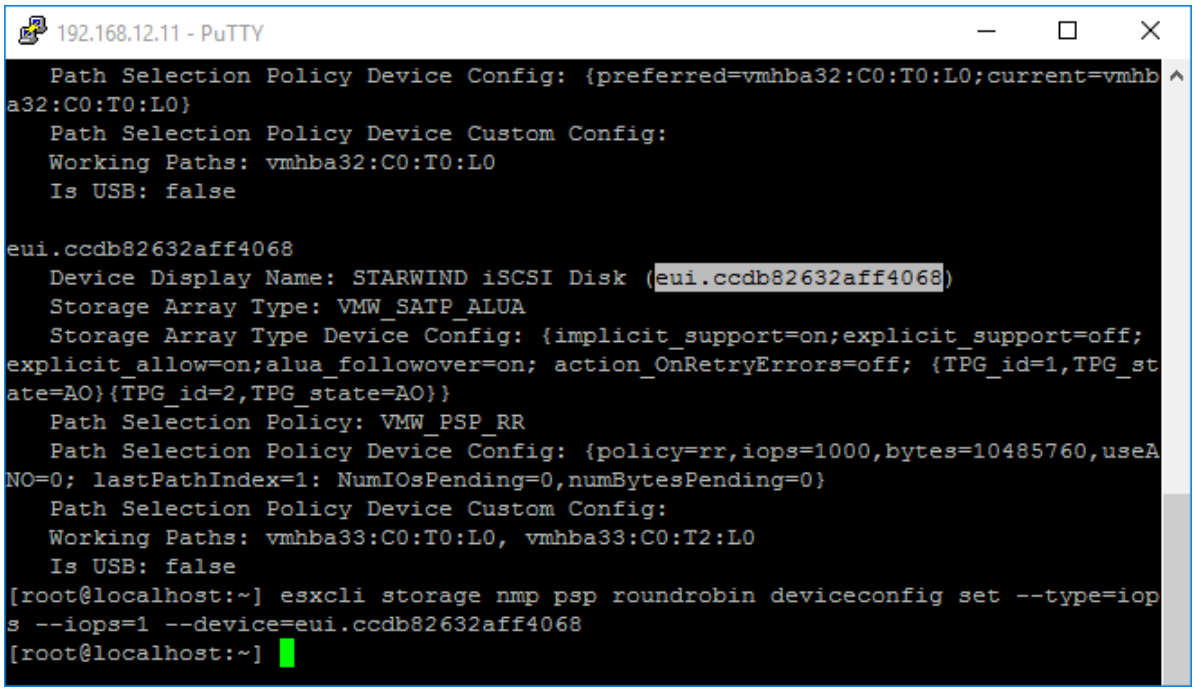
4. Check the device list using the following commands:

```
esxcli storage nmp device list
```

For devices, adjust Round Robin size from 1000 to 1 using the following command:

```
esxcli storage nmp psp roundrobin deviceconfig set --type=iops
--iops=1 --device=
```

NOTE. Paste StarWind device UID in the end of the command. Also, Rescan Script already contains this parameter and the system performs this action automatically.



5. Repeat the steps 1-4 on each host for each datastore.

Contacts

US Headquarters	EMEA and APAC
+1 617 829 44 95	+44 2037 691 857 (United Kingdom)
+1 617 507 58 45	+49 800 100 68 26 (Germany)
+1 866 790 26 46	+34 629 03 07 17 (Spain and Portugal)
	+33 788 60 30 06 (France)

Customer Support Portal: https://www.starwind.com/support

Support Forum: https://www.starwind.com/forums

Sales: sales@starwind.com

General Information: info@starwind.com

StarWind

StarWind Software, Inc. 100 Cummings Center Suite 224-C Beverly MA 01915, USA
www.starwind.com ©2024, StarWind Software Inc. All rights reserved.