StarWind Virtual SAN® Creating NVMe-oF

2024

TECHNICAL PAPERS

Trademarks

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.

Introduction

NVMe over Fabrics is the protocol that is set to be the standard for shared NVMe storage which is much faster and efficient than usual SAS. Moving things forward, StarWind Software designed NVMe over Fabrics [1] which can be connected to the host on demand and provide really fast NVMe storage there.

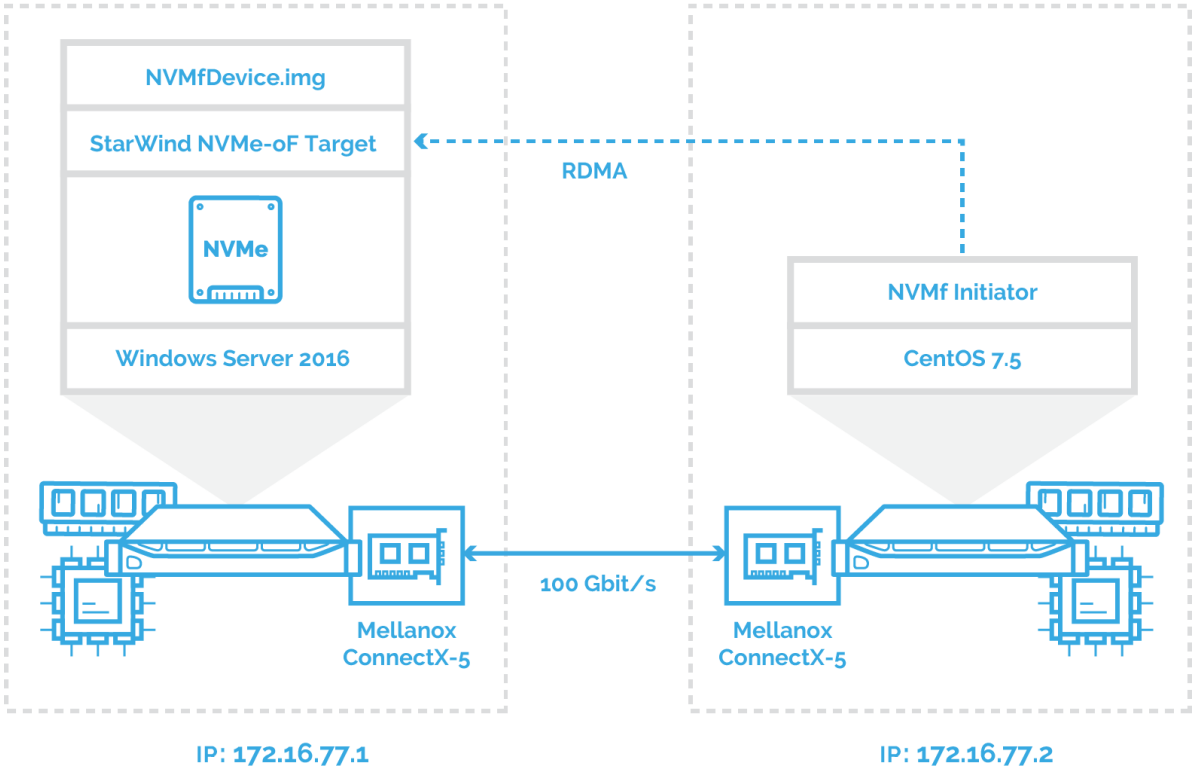
This guide is intended for IT professionals and enthusiasts who would like to configure the StarWind NVMe over Fabrics. It provides detailed instructions on how to configure the hosts and get shared NVMe-oF storage connected on the CentOS host.

A full set of up-to-date technical documentation can always be found here, or by pressing the Help button in StarWind Management Console.

For any technical inquiries please visit our online community, Frequently Asked Questions page or use the support form to contact our technical support department.

Preconfiguring The Servers

Here is a network diagram of the configuration described in this guide.



In this document, the host with Mellanox ConnectX-5 adapter and NVMe drive is running Windows Server 2016. The second host is running CentOS 7.5 and has Mellanox ConnectX-5 adapter installed accordingly. The NVMe-oF is created on Windows Server 2016 using StarWind NVMe over Fabrics service and discovered and connected on CentOS 7.5 via 172.16.77.x subnet.

Windows Server 2016 should have the latest Mellanox driver installed which is available here: http://www.mellanox.com/page/products_dyn?product_family=32&mtag=windows_sw_drivers

For proper RDMA functioning with Mellanox, the CentOS kernel version should be higher than 4.9. The guide on how to install or upgrade the kernel in CentOS 7 can be found at the following link: https://www.tecmint.com/install-upgrade-kernel-version-in-centos-7/

Please install important packages when connected to the CentOS host via terminal console (e.g. PuTTY) as it is specified in the Mellanox Red Hat Enterprise Linux (RHEL) 7.5 Driver User Manual (http://www.mellanox.com/pdf/prod_software/Red_Hat_Enterprise_Linux_(RHEL)_7.5_Driver_User_Manual.pdf):

```
# yum install libibverbs librdmacm libibcm libibmad libibumad
libmlx4 libmlx5 opensm ibutils infiniband-diags srptools
perftest mstflint rdmacm-utils ibverbs-utils librdmacm-utils –y
```

Also, install the libibverbs-utils package to get the ability to check that Infiniband is active:

```
# yum install libibverbs-utils
```

To check that Infiniband port state, run the following command. The port state must be Active.

```
# ibv_devinfo
```

```
[root@localhost ~]# ibv_devinfo
hca_id: mlx5_0
        transport:                      InfiniBand (0)
        fw_ver:                         14.20.1010
        node_guid:                      000c:29ff:fe01:73d0
        sys_image_guid:                 ec0d:9a03:00d9:979e
        vendor_id:                      0x02c9
        vendor_part_id:                 4118
        hw_ver:                         0x0
        board_id:                       MT_2420110004
        phys_port_cnt:                  1
                port:   1
                        state:                  PORT_ACTIVE (4)
                        max_mtu:                4096 (5)
                        active_mtu:             4096 (5)
                        sm_lid:                 0
                        port_lid:               0
                        port_lmc:               0x00
                        link_layer:             Ethernet
```

The Mellanox modules for ConnectX®-4/ConnectX®-4 Lx/ConnectX®-5/ConnectX®-5 Ex are: mlx5_core and mlx5_ib. To check that kernel modules mlx5_core, mlx5_ib and RDMA are loaded, run the following command:

```
# lsmod | grep mlx5
```

```
# lsmod | grep rdma
```

```
[root@localhost ~]# lsmod | grep mlx5
mlx5_ib               249856  0
ib_uverbs              98304  2 rdma_ucm,mlx5_ib
ib_core               233472  14 rdma_cm,ib_ipoib,rpcrdma,ib_srpt,ib_srp,nvme_rdma,iw_cm,ib_iser,ib_umad,ib_isert,
rdma_ucm,ib_uverbs,mlx5_ib,ib_cm
mlx5_core             528384  1 mlx5_ib
mlxfw                  20480  1 mlx5_core
ptp                    20480  1 mlx5_core
```

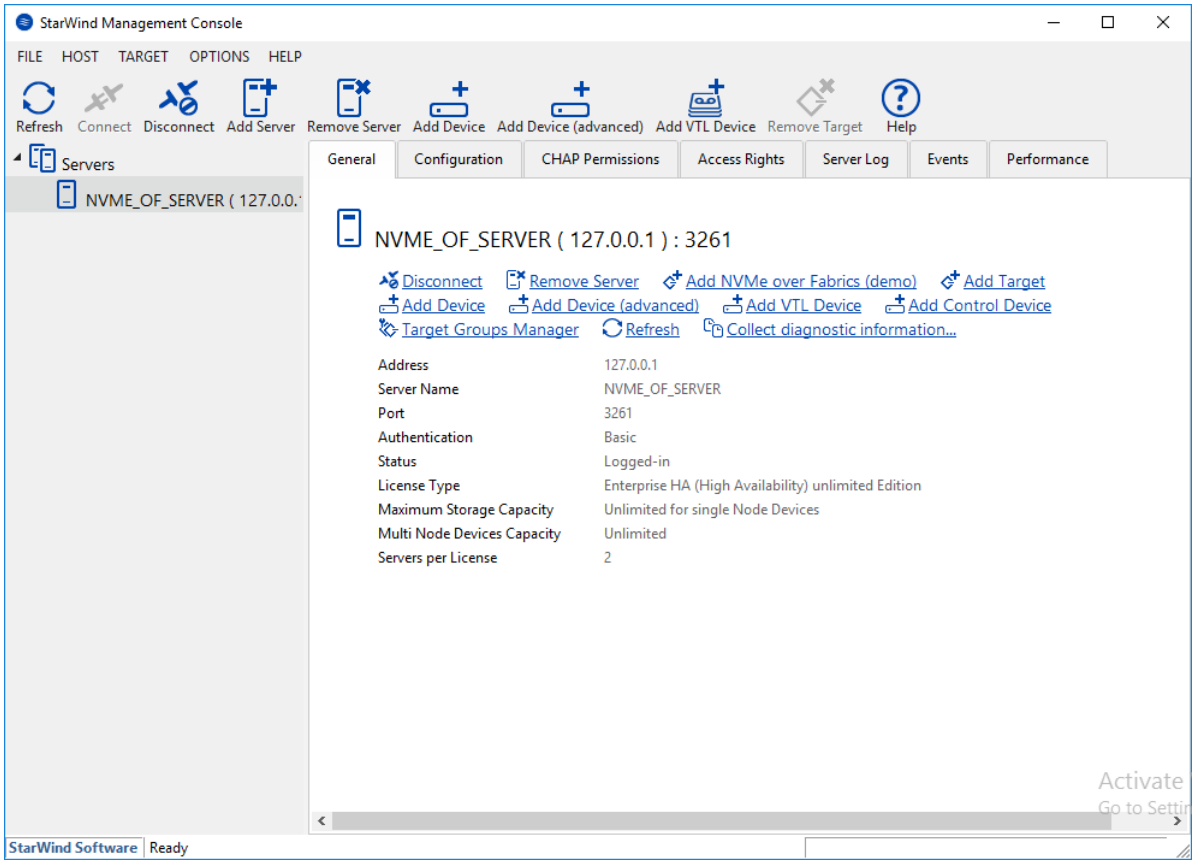
```
[root@localhost ~]# lsmod | grep rdma
rpcrdma               212992  0
sunrpc                356352  1 rpcrdma
rdma_ucm               28672  0
ib_uverbs              98304  2 rdma_ucm,mlx5_ib
nvme_rdma              28672  0
rdma_cm                57344  7 rpcrdma,ib_srpt,ib_srp,nvme_rdma,ib_iser,ib_isert,rdma_ucm
iw_cm                  45056  1 rdma_cm
ib_cm                  53248  4 rdma_cm,ib_ipoib,ib_srpt,ib_srp
ib_core               233472  14 rdma_cm,ib_ipoib,rpcrdma,ib_srpt,ib_srp,nvme_rdma,iw_cm,ib_iser,ib_umad,ib_isert
rdma_ucm,ib_uverbs,mlx5_ib,ib_cm
nvme_fabrics           20480  1 nvme_rdma
```

To check the bandwidth between Windows Server 2016 and CentOS 7.5 hosts, use a StarWind rPerf utility, which can be downloaded here:
https://www.starwindsoftware.com/starwind-rperf

Creating Nvme-Of

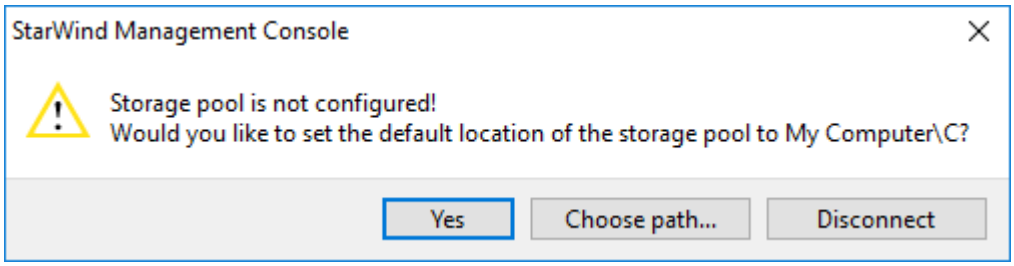
1. Login to Windows host and install StarWind VSAN by following the guide here: https://www.starwindsoftware.com/resource-library/starwind-virtual-san-quick-start-guide-installation

2. Launch StarWind Management Console from the StarWind icon in the tray.



3. If StarWind Service and Management Console are installed on the same server, the Management Console will automatically add the local StarWind instance to the Console tree after the first launch. Then, the Management Console automatically connects to it using default credentials. To add a StarWind server to the console, use the Add Server button on the control panel.

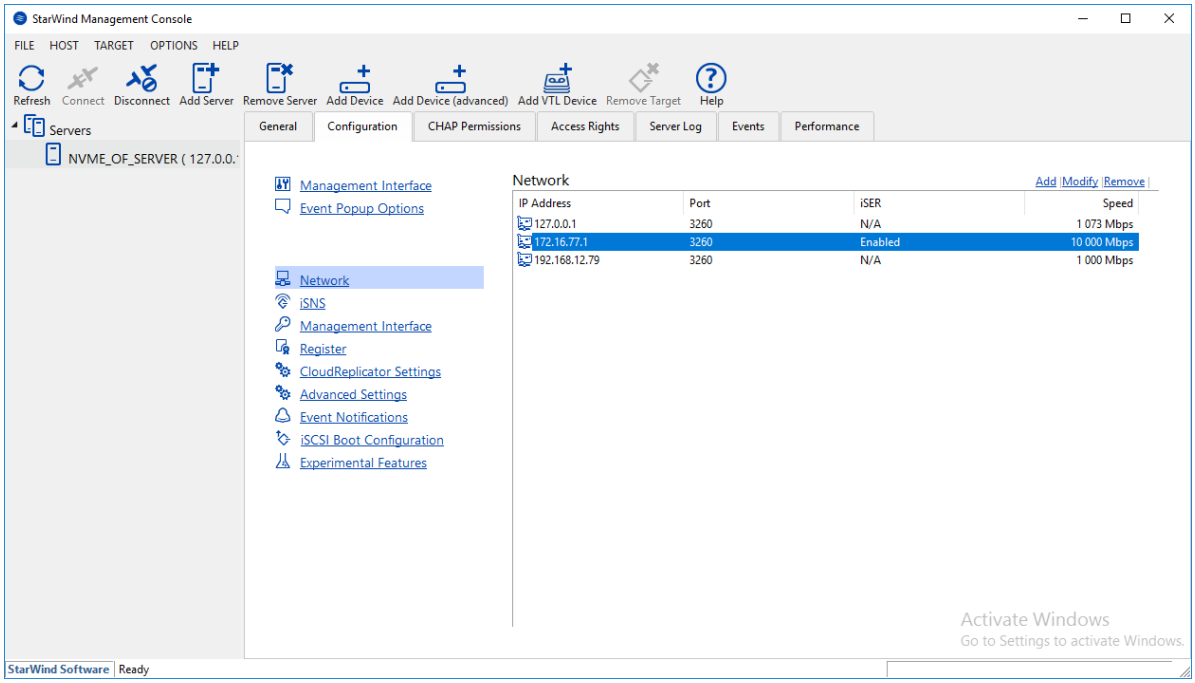
4. StarWind Management Console will ask to specify the default storage pool on the server which it connects to for the first time. Configure the default storage pool to use one of the volumes that have been prepared previously. Press the Yes button to configure the storage pool. If the storage pool destination needs to be changed, click Choose path... and point the browser to the necessary disk.



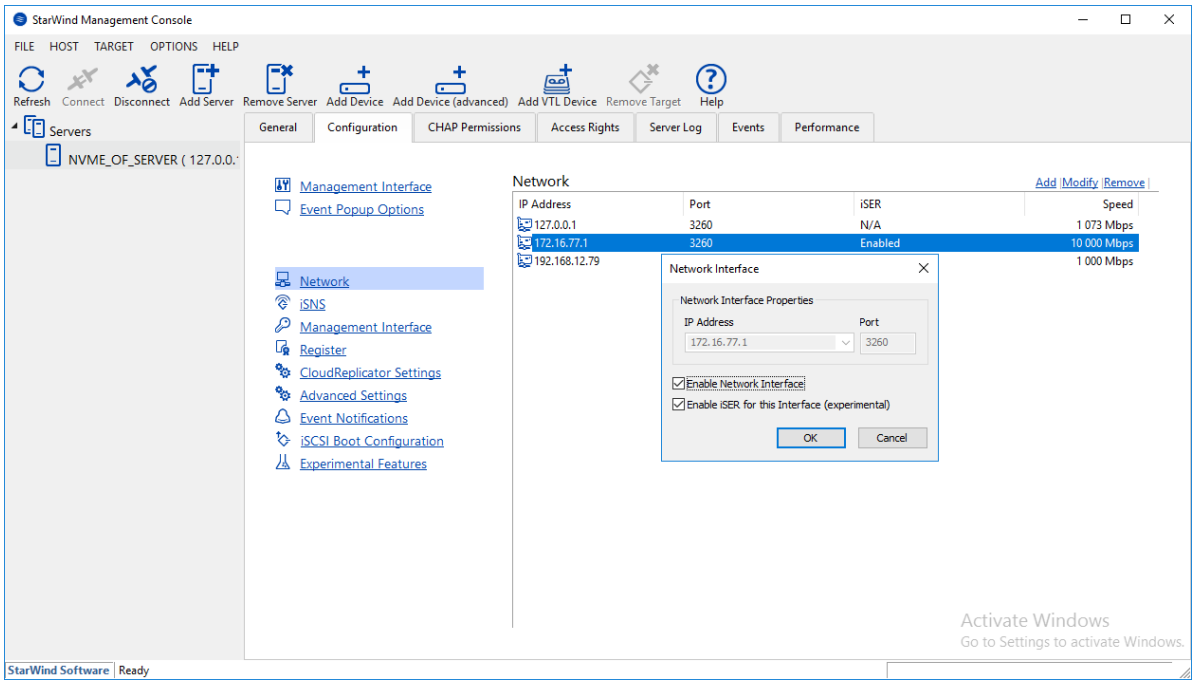
NOTE: Each of the arrays which will be used by StarWind Virtual SAN to store virtual disk images, should meet the following requirements:

- Be initialized as GPT,
- Have a single NTFS-formatted partition,
- Have a drive letter assigned.

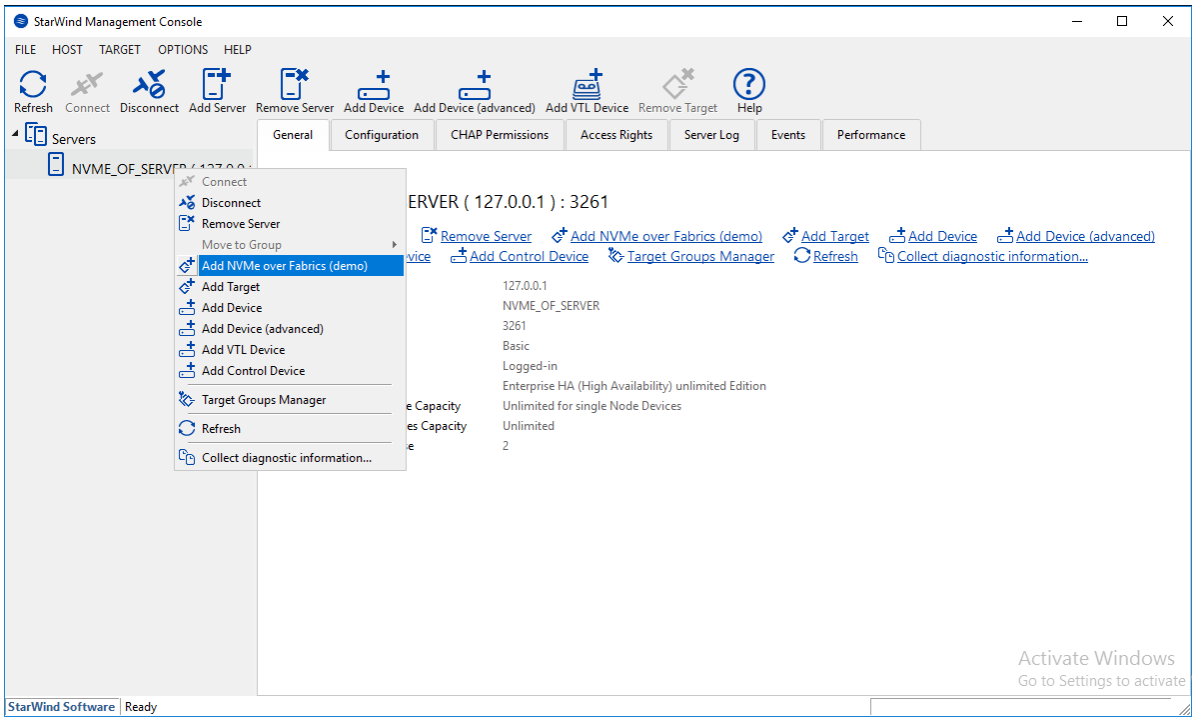
5. In StarWind Management Console, open the server Configuration tab and click the Network button. Check that ISER support is enabled.



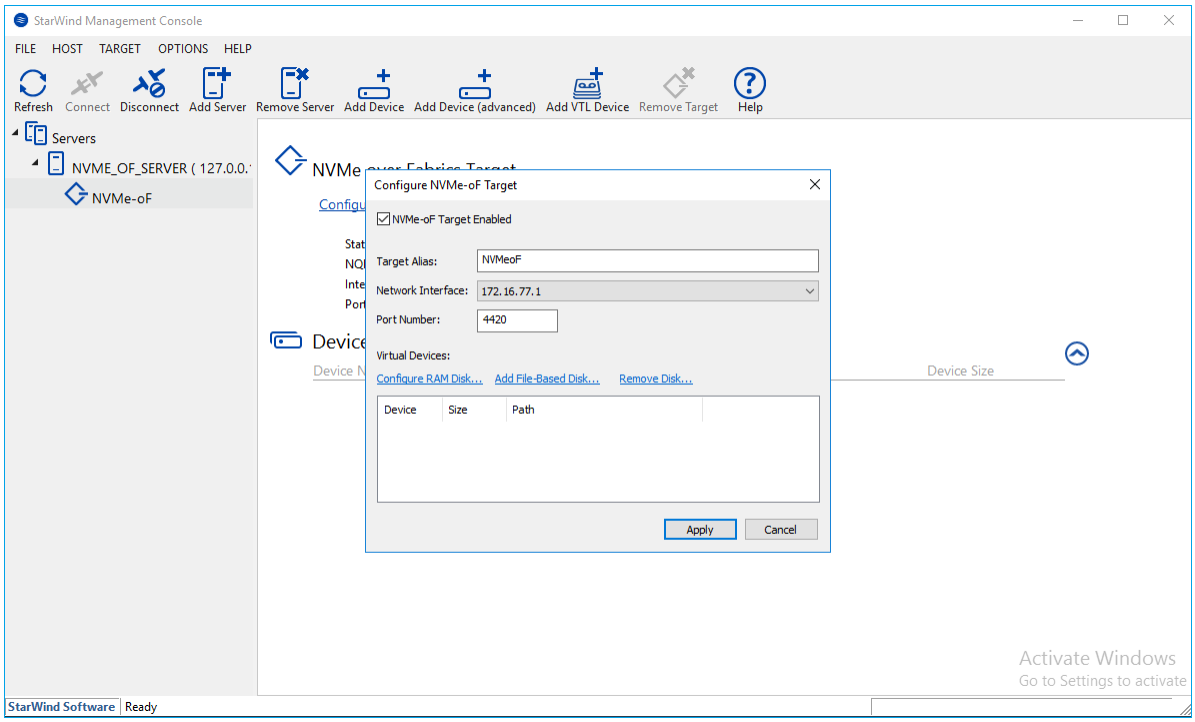
6. To enable iSER support, click on the network (here 172.16.77.1), press the Modify button and enable the checkbox Enable iSER for this Interface... in the appeared window. Click OK to confirm.



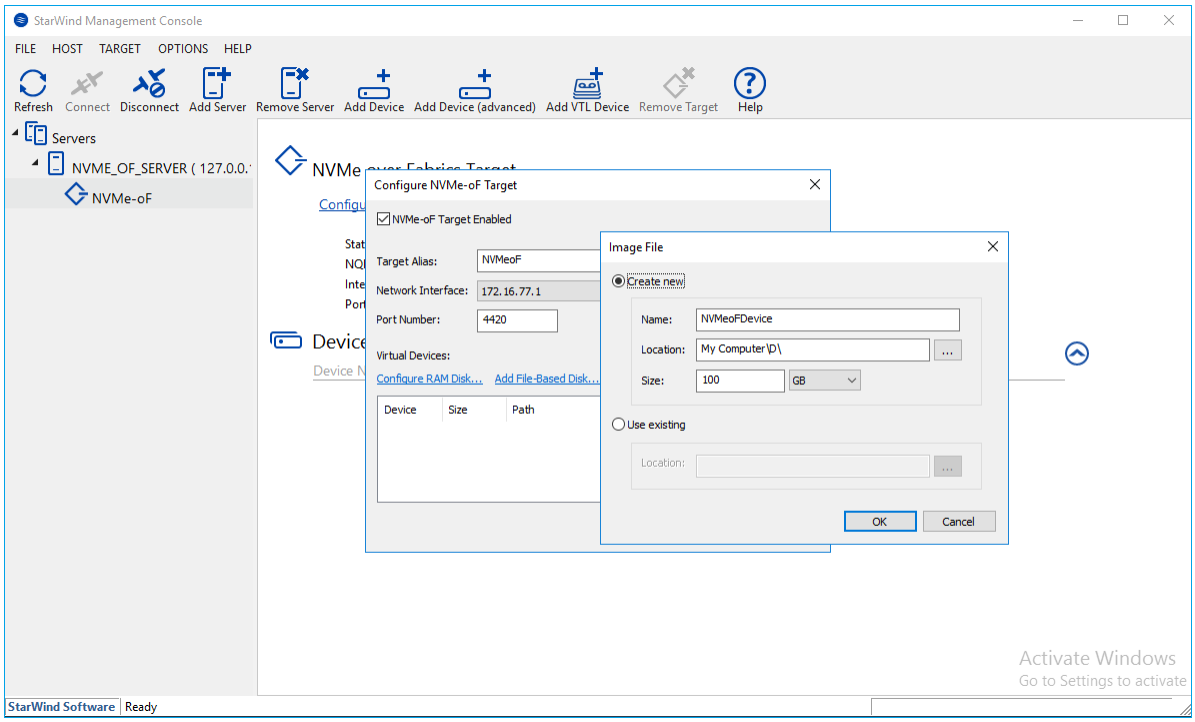
7. Right click on the server and choose Add NVMe-oF... option in the appeared list.



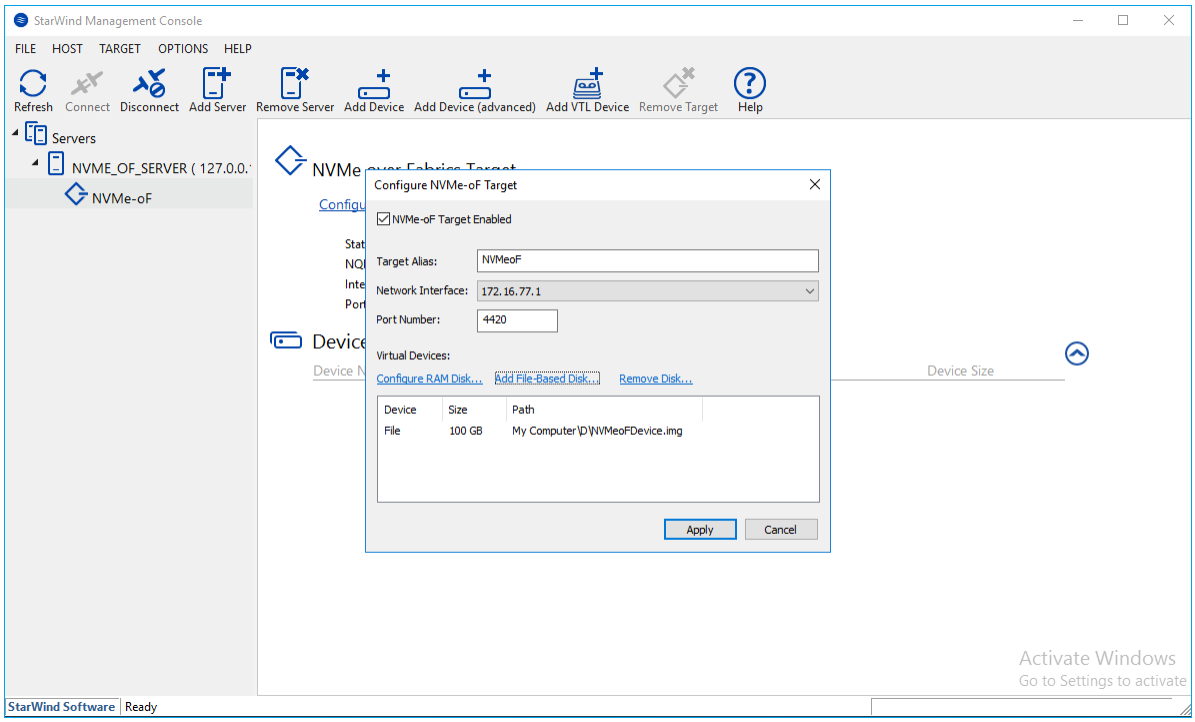
8. lick on the newly created target and press the Config... button. Enter Target Alias and port number (default is 4420) and click the Add File-Based Disk... button.



9. In the newly appeared window, specify virtual disk name, its location, and size. Click OK to save the settings.



10. Click Apply to confirm the NVMe-oF configuration.



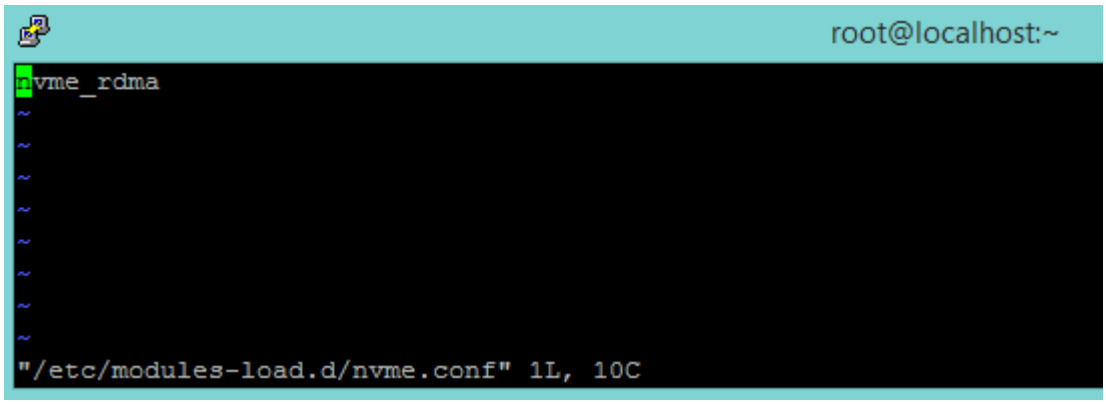
Discovering And Connecting Nvme-Of On Centos Host

1. The Linux kernel NVMe over Fabrics RDMA host support is provided by the nvme-rdma driver:

```
modprobe nvme-rdma
```

2. To make sure that nvme-rdma is loading during OS boot, edit nvme.conf using text editor and add nvme_rdma line there.

```
# vi /etc/modules-load.d/nvme.conf
```



3. The nvme-cli tool may be used to interface with the Linux kernel NVMe over Fabrics host. To install nvme-cli, run the following command:

```
# yum install nvme-cli
```

```
[root@localhost ~]# yum install nvme-cli
Loaded plugins: fastestmirror
Loading mirror speeds from cached hostfile
 * base: ftp.agh.edu.pl
 * elrepo: ftp.colocall.net
 * epel: epel.mirror.srv.magticom.ge
 * extras: ftp.vectranet.pl
 * updates: ftp.agh.edu.pl
Resolving Dependencies
--> Running transaction check
---> Package nvme-cli.x86_64 0:1.6-1.el7 will be installed
--> Finished Dependency Resolution

Dependencies Resolved

================================================================================
 Package            Arch              Version                Repository      Size
================================================================================
Installing:
 nvme-cli           x86_64            1.6-1.el7              base           250 k

Transaction Summary
================================================================================
Install  1 Package

Total download size: 250 k
Installed size: 448 k
Is this ok [y/d/N]: y
Downloading packages:
nvme-cli-1.6-1.el7.x86_64.rpm                               | 250 kB   00:00
Running transaction check
Running transaction test
Transaction test succeeded
Running transaction
  Installing : nvme-cli-1.6-1.el7.x86_64                                   1/1
  Verifying  : nvme-cli-1.6-1.el7.x86_64                                   1/1

Installed:
  nvme-cli.x86_64 0:1.6-1.el7

Complete!
[root@localhost ~]#
```

4. To discover the target, run nvme discover command by specifying the protocol, target host IP address, and port number:

```
# nvme discover -t rdma -a 172.16.77.1 -s 4420
```

```
[root@localhost ~]# nvme discover -t rdma -a 172.16.77.1 -s 4420

Discovery Log Number of Records 1, Generation counter 0
=====Discovery Log Entry 0======
trtype:  rdma
adrfam:  ipv4
subtype: nvme subsystem
treq:    not specified
portid:  0
trsvcid: 4420
subnqn:  nqn.2008-08.com.starwindsoftware:NVMfTarget
traddr:  172.16.77.1
rdma_prtype: not specified
rdma_qptype: connected
rdma_cms:    rdma-cm
rdma_pkey: 0x0000
[root@localhost ~]#
```

5. To connect the target, run nvme connect command by specifying target name (can be copied from the previous command execution result), target host IP address, and port number:

```
# nvme connect -t rdma -n
"nqn.2008-08.com.starwindsoftware:NVMfTarget" -a 172.16.77.1 -s
4420
```

6. To check the connected device, run nvme list command:

```
[root@localhost ~]# nvme connect -t rdma -n "nqn.2008-08.com.starwindsoftware:NVMfTarget" -a 172.16.77.1 -s 4420
[root@localhost ~]# nvme list
Node             SN                   Model                                    Namespace Usage
  Format           FW Rev
---------------- -------------------- ---------------------------------------- --------- -------------------------
- ---------------- --------
/dev/nvme0n1     2959370297           SPDK Virtual Controller                  1         107.37  GB / 107.37  GB
  512   B +  0 B   FFFFFFFF
[root@localhost ~]#
```

7. The device can be found in the list of devices, provided by the lsblk command:

```
[root@localhost ~]# lsblk
NAME    MAJ:MIN RM   SIZE RO TYPE MOUNTPOINT
nvme0n1 259:0    0   100G  0 disk
```

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