

FlexSDS Build Scale-out Storage for VMware ESXi

Friday, February 10, 2023

FlexSDS Software Limited.

www.flexsds.com

Copyright © FlexSDS Software Limited 2016-2023. All right reserved.

Table of Contents

Overview.....	4
Topology.....	4
Single node mode.....	4
Dual nodes HA mode.....	5
3(+) Nodes Scale-out SDS mode.....	6
Install and Configure the FlexSDS Software.....	7
Setup in the ESXi.....	7
Hardware Configuration.....	8
Setup RoCE network in ESXi.....	8
Install network HBA driver.....	8
Setup Network in VMware ESXi.....	9
Enable NVMe over Fabrics adapter.....	10
Check if nvme rdma module is loaded.....	11
Enable the nvme rdma adapter.....	11
Logon to FlexSDS NVMe over Fabrics volume.....	12
Create datastore on the NVMe RDMA disk.....	13
Enable the VMware iSER Adapter.....	16
Add iSER Adapter.....	16
Set up the iSER adapter.....	17
Create Datastore on the iSER disk.....	18
Enable iSCSI adapter.....	22
Set up iSCSI adapter.....	22
Create Datastore on the iSCSI Disk.....	23
Create Virtual Machine.....	27

Test Speed (optional)	31
Conclusion	34
Contact	35

Overview

FlexSDS is a high-performance storage solution designed for HPC, cloud computing, edge computing and databases. It offers high availability, scale-out, redundant and auto-recovery storage options through n-ways replication or EC (erasure coding). Users can create single-node, two-node HA, or multi-node storage clusters and export dynamic volumes to VMWare vSphere ESXi via NVMe-oF, iSER and iSCSI.

This chapter explains how to set up NVMe-oF, iSER and iSCSI adapters in ESXi and connect to FlexSDS's NVMe over Fabrics, iSER and iSCSI volumes. It also compares the performance of NVMe-oF, iSER and iSCSI interfaces, with NVMe over RDMA and iSER showing much higher performance than iSCSI and NVMe over TCP due to its use of RDMA transport, as that enables full stack of kernel bypass, zero copies of FlexSDS software.

This document provides step-by-step guide for build FlexSDS high availability and scale-out storage cluster for VMware vSphere, and exporting storage service via NVMe-oF, iSER or iSCSI.

Topology

FlexSDS to be a scale-out software defined storage, that supports three topology modes:

- Single nodes scale up SDS, this mode supports only in node data redundancy
- Dual-nodes HA SDS, this mode is the minimum requires to support HA service, that provides not only in-node data redundancy, but also cross-node data redundancy.
- Tree or more nodes scale-out clustered SDS, this mode is true scale-out/scale-up clustered storage, that supports all types of data redundancy.

Single node mode

When FlexSDS is working as single node mode, FlexSDS can't offer high availability feature since the only one storage node failed, whole storage service will be failed. Therefore this mode is not recommended in most production cases, but it could be used as proven of your solutions, and it could be a start to use

two nodes HA or more nodes cluster. User can easily to expand the storage service to Dual nodes HA or 3+ clustered by adding new server nodes.

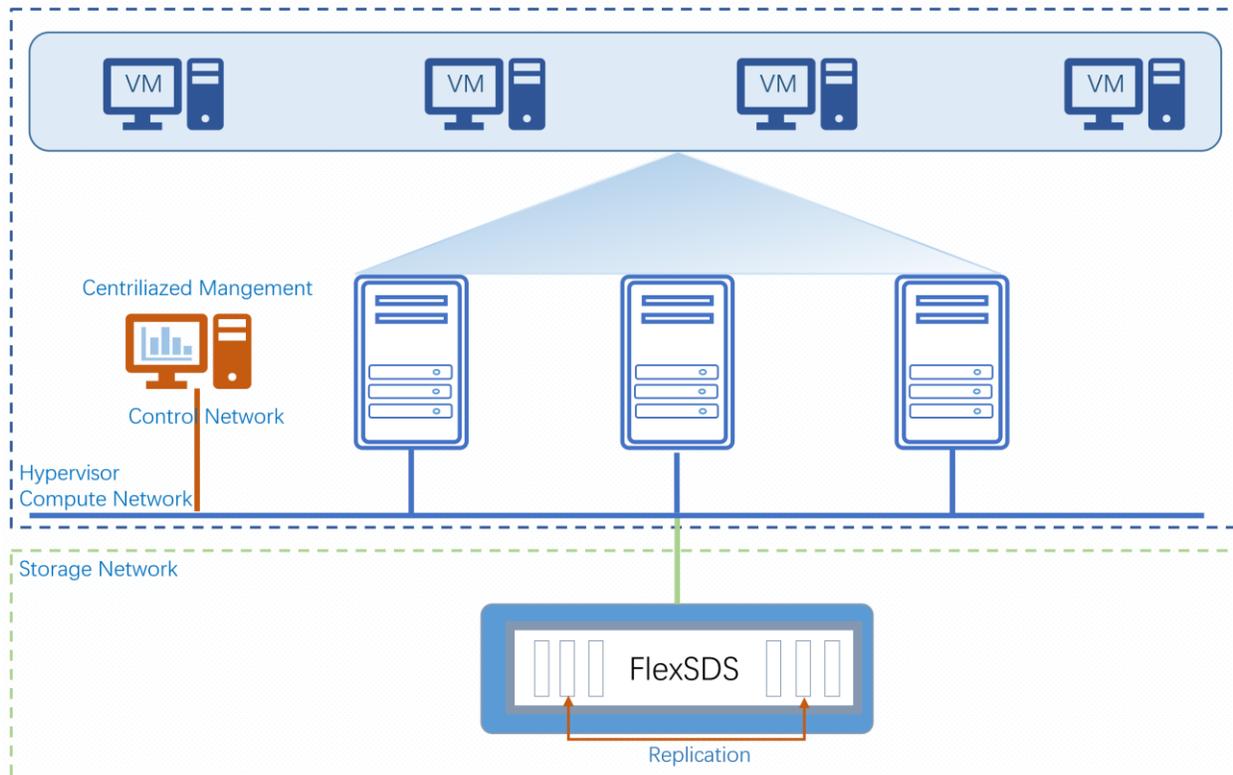


Figure 1. Single Node SDS

Dual nodes HA mode

FlexSDS doesn't really work in 2 node mode, as software-only solution can't 100% prevent brain-split, that still requires a third party node as arbitration node, the arbitration node could be a another server node inside the same network or a VM inside business server like ESX, qemu-kvm, etc. The arbitration node needs network function to communicate with the two nodes in HA, except this, no more other requires (CPU, storage etc).

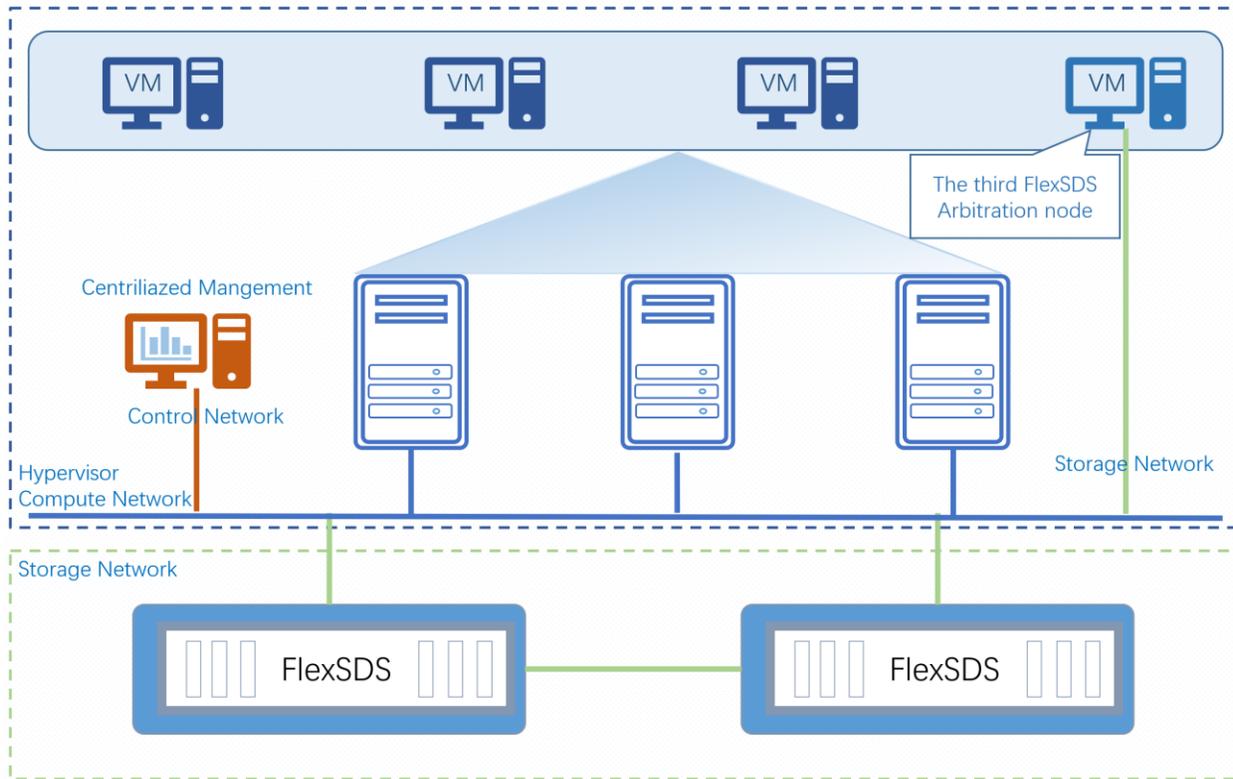


Figure 2. Two Nodes HA SDS

3(+) Nodes Scale-out SDS mode

As FlexSDS was designed as scale-out, distributed storage service, 3 or 3+ nodes are the most used working mode. User can dynamically add/remove node to expand/shrink the storage cluster.

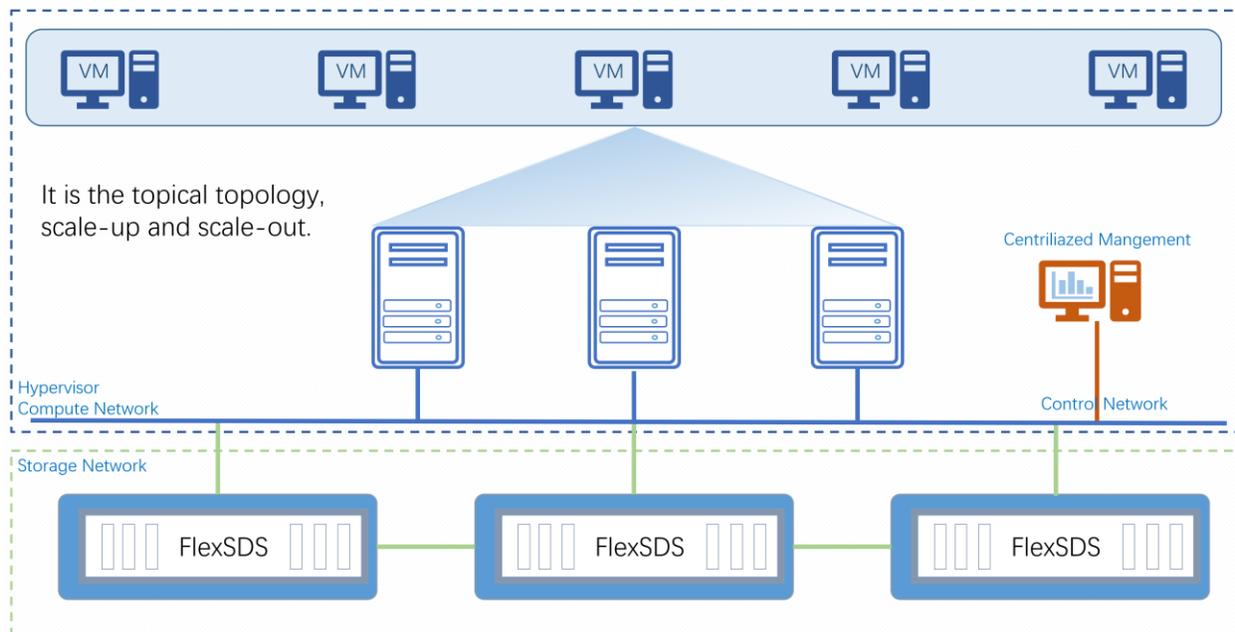


Figure 3. 3 Nodes Scale-out Clustered SDS

Install and Configure the FlexSDS Software

Please refer to the white paper “[Install and deploy FlexSDS scale-out software defined storage](#)”, to set up FlexSDS, follow these steps.

1. Install FlexSDS
2. Deploy the FlexSDS Cluster
3. Add Backends
4. Create Storage Pools (n-ways or EC, use single node 2 ways replication as an example).
5. Create a RAW volume with 1TB, or any desired size.
6. Add NVMe over Fabrics, iSER or iSCSI interface as needed.

FlexSDS is Linux based, the CentOS and Ubuntu Server are the recommended OS to install FlexSDS Scale-out Storage, install FlexSDS software in Linux is very easy.

Setup in the ESXi

Hardware Configuration.

This is for following performance test configuration, user can choose any x86 based servers.

- One Physical Servers for Storage (two or more are also acceptable, the more the higher maximum IOPS)
- 4x NVMe P4610 3.2TB (2 ways replication, Kernel bypass mode)
- One Physical server for ESXi
- Two Intel Xeon Gold 5215 2.50GHz
- 64GB DDR4 RAM
- Mellanox ConnectX-4 100GbE NIC, Direct link
- VMware ESXi, 7.0.3, 18644231
- CentOS 7 (64-bit) for Storage
- FlexSDS 2023 V1 (v5.0) 4 CPU Cores

Setup RoCE network in ESXi.

User should skip this chapter if they already have RoCE network configured, here are the steps for this scenario.

Install network HBA driver.

To install RoCE HBA driver, different hardware manufacturers offer different methods, please refer to the hardware manual and ESXi materials for detailed instructions.

Here we take Mellanox ConnectX-4 100G as an example, ESXi contains built-in drivers for it.

Prerequisites

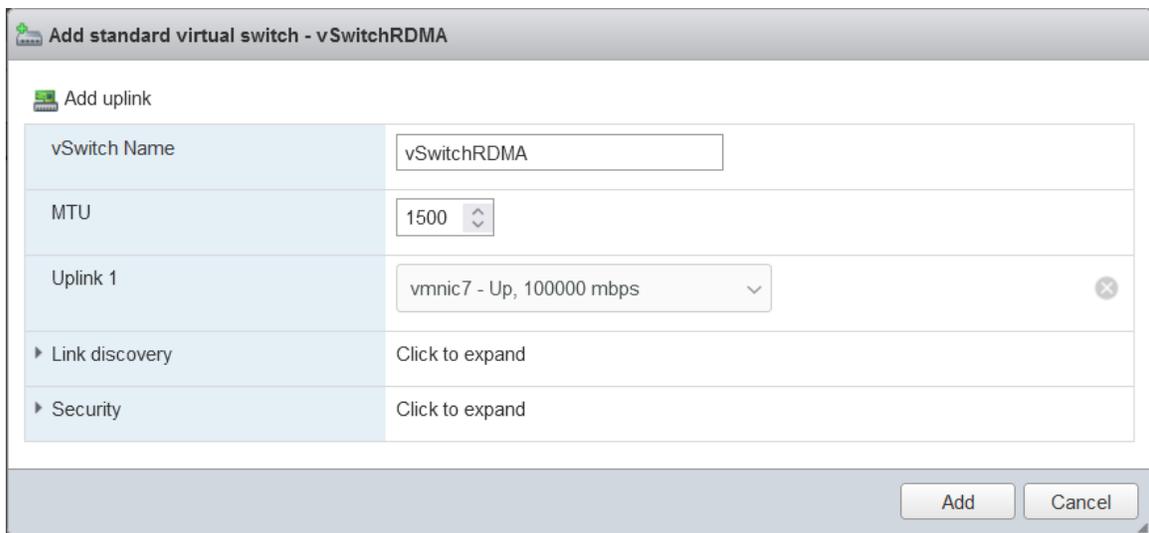
- RDMA capable adapter is necessary on both storage server and ESXi host.
- For RDMA capable adapters that support RDMA over Converged Ethernet (RoCE), determine the RoCE version that the adapter uses.
- Use the RDMA capable switch or direct link.

- Enable flow control on the ESXi host. To enable flow control for the host, use the `esxcli system module parameters` command. For details, see the VMware knowledge base article at <http://kb.vmware.com/kb/1013413>.
- Make sure to configure RDMA switch ports to create lossless connections between the initiator and target.

Setup Network in VMware ESXi

Navigate to Networking->Virtual Switches.

User should see the default vSwitch0, click the Add standard virtual switch link. The Add standard virtual switch link window shows



Add standard virtual switch - vSwitchRDMA	
Add uplink	
vSwitch Name	vSwitchRDMA
MTU	1500
Uplink 1	vmnic7 - Up, 100000 mbps
▶ Link discovery	Click to expand
▶ Security	Click to expand
Add Cancel	

Type vSwitch Name, for example: vSwitchRDMA.

Select Uplink 1, here we selected the one with 100G bandwidth.

Click the Add button to create vSwitch.

Navigate to Networking->VMkernel NICs.

Click the Add VMkernel NIC link, the Add VMkernel NIC window shows.

Port group	New port group
New port group	RDMA
Virtual switch	vSwitchRDMA
VLAN ID	0
MTU	1500
IP version	IPv4 only
IPv4 settings	
Configuration	<input type="radio"/> DHCP <input checked="" type="radio"/> Static
Address	192.168.20.100
Subnet mask	255.255.255.0
TCP/IP stack	Default TCP/IP stack

Create Cancel

Type name in the New port group field.

Select the newly created Virtual switch, vSwitchRDMA.

Configure the IPV4 address

Click the Create button to create the VMkernel NIC.

Enable NVMe over Fabrics adapter

In the vSphere Client, navigate to the ESXi host.

Click the Configure tab page.

Under Storage, click Storage Adapters, and click the Add Software Adapter icon.

Select the adapter type as required.

- NVMe over RDMA adapter
- NVMe over TCP adapter

Here are the steps to enable it via CLI

Connect to the ESXi host using SSH or the direct console.

Enter the following command to enable the NVMe RDMA adapter:

Check if nvme rdma module is loaded.

Issue the command to check if the nvmerdma system module is loaded.

```
#esxcli system module list |grep nvme
```

If it does not exist, use the command to load.

```
#esxcli system module load -m nvmerdma.
```

```
[root@esxi:~] esxcli system module list | grep nvme
vmknvme          true          true
nvme_pcie        true          true
[root@esxi:~] esxcli system module load -m nvmerdma
```

User may issue the command to enable nvme-tcp if needed.

```
#esxcli system module load -m nvmetcp.
```

Enable the nvme rdma adapter

Issue the command to find suitable RDMA nic to create the nvme-rdma adapter.

```
#esxcli rdma device list
```

Here we can see vmrdma0 is supported to run nvme-rdma adapter, then issue the command to enable it.

```
#esxcli nvme fabrics enable -d vmrdma0 -p RDMA
```

```
[root@esxi:~] esxcli rdma device list
Name      Driver      State      MTU      Speed      Paired Uplink  Description
-----
vmrdma0   nmlx5_rdma  Active     1024     100 Gbps    vmnic7         MT27700 Family [ConnectX-4]
[root@esxi:~] esxcli nvme fabrics enable -d vmrdma0 -p RDMA
true
```

To enable nvme-tcp, user should use NIC name.

```
#esxcli nvme fabrics enable -d vmnic7 -p TCP
```

```
[root@esxi:~] esxcli nvme fabrics enable -d vmnic7 -p TCP
true
```

Check if NVMe adapters were created correctly.

```
#esxcli nvme info get
```

and

```
#esxcli nvme adapter list.
```

User should see the host NQN name and adapter name for each NVMe adapters.

```
[root@esxi:~] esxcli nvme info get
Host NQN: nqn.2014-08.com.vmware:nvme:esxi
[root@esxi:~] esxcli nvme adapter list
Adapter Adapter Qualified Name      Transport Type  Driver      Associated Devices
-----
vmhba64  aqn:nvmerdma:b8-59-9f-05-b5-41      RDMA           nvmerdma    vmrdma0, vmnic7
vmhba65  aqn:nvmetcp:b8-59-9f-05-b5-41-T     TCP            nvmetcp     vmnic7
```

Logon to the FlexSDS NVMe over Fabrics volume.

Discover the NQN of NVMe over Fabrics volumes.

```
#nvmecli nvme fabrics discover -a vmhba64 -i 192.168.20.121 -p 4420
```

```
[root@esxi:~] esxcli nvme fabrics discover -a vmhba64 -i 192.168.20.121 -p 4420
Transport Type  Address Family  Subsystem Type  Controller ID  Admin Queue Max Size  Transport Address  Transport Service ID  Subsystem NQN
-----
RDMA            IPv4            NVM              65535          128                   192.168.20.121    4420                   nqn.2016-12.com.flexsds:all-flash-pool.nvmf
```

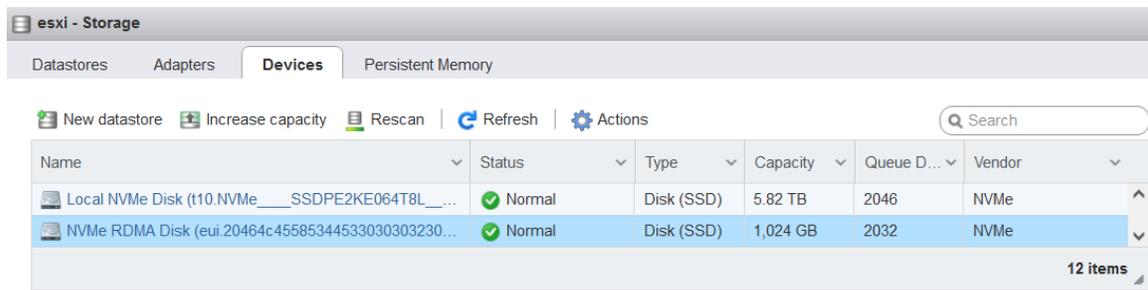
Connect to the NVMe over Fabrics volume

```
# esxcli nvme fabrics connect -a vmhba64 -i 192.168.20.121 -p 4420 -s nqn.2016-12.com.flexsds:all-flash-pool.nvmf
```

For supporting NVMe multi-path, user can add a second connection:

```
# esxcli nvme fabrics connect -a vmhba64 -i 192.168.20.122 -p 4420 -s nqn.2016-12.com.flexsds:all-flash-pool.nvmf
```

In ESXi WEB client, navigate to Storage->Devices, there should be a new disk like NVMe RDMA Disk appears.



The screenshot shows the ESXi Storage Devices page. At the top, there are tabs for 'Datastores', 'Adapters', 'Devices', and 'Persistent Memory'. Below the tabs, there are action buttons: 'New datastore', 'Increase capacity', 'Rescan', 'Refresh', and 'Actions'. A search bar is also present. The main content is a table with the following columns: Name, Status, Type, Capacity, Queue D..., and Vendor. There are two rows in the table:

Name	Status	Type	Capacity	Queue D...	Vendor
Local NVMe Disk (t10.NVMe____SSDPE2KE064T8L____...)	Normal	Disk (SSD)	5.82 TB	2046	NVMe
NVMe RDMA Disk (eui.20464c4558534453303030323030...)	Normal	Disk (SSD)	1,024 GB	2032	NVMe

At the bottom right of the table, it says '12 items'.



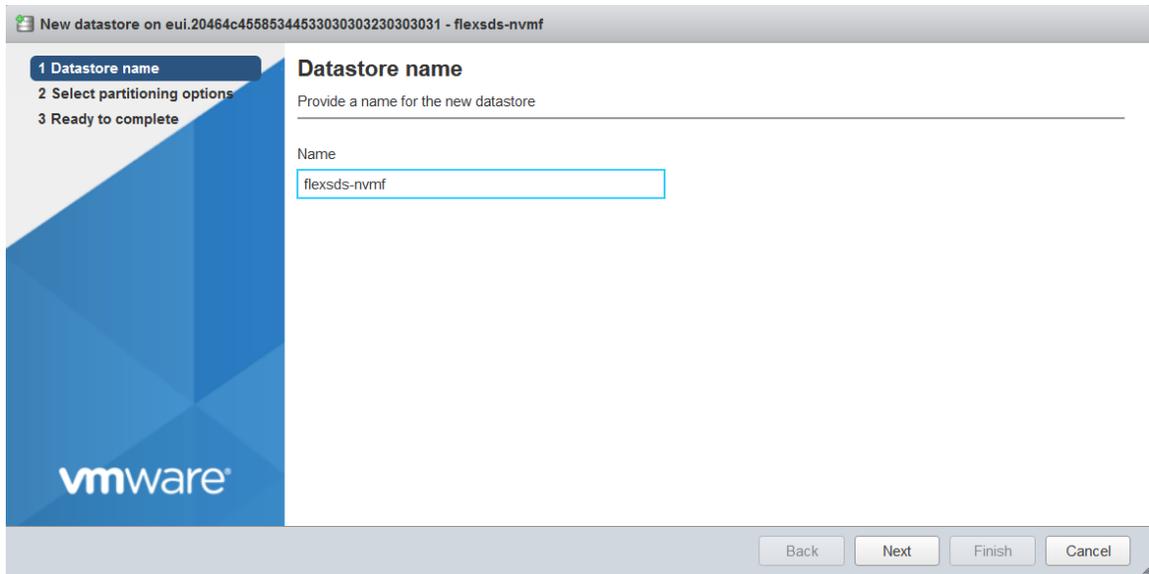
NVMe RDMA Disk (eui.20464c45585344533030303230303031)

Type: Disk
Model: FLEXSDS Controller
Path: /vmfs/devices/disks
/eui.20464c45585344533030303230303031
Capacity: 1,024 GB
Partition Format: Unknown
UUID: 0620464c45585344533030303230303031

Create datastore on the NVMe RDMA disk

click the New Datastore link.

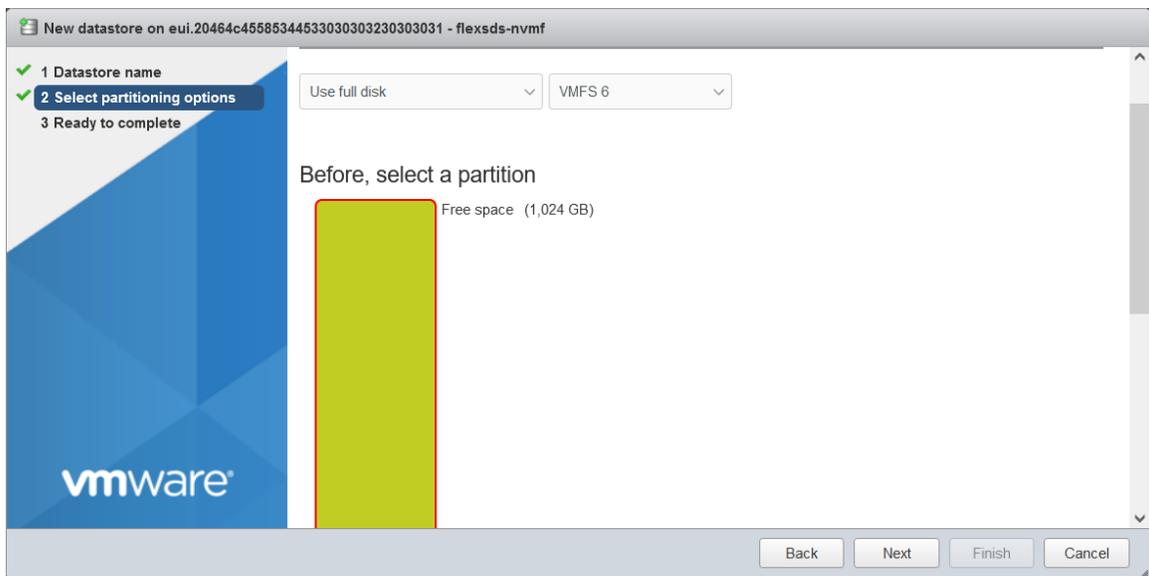
The New datastore – on eui.. window shows



Type in the Name, "flexsds-nvmf " as an example.

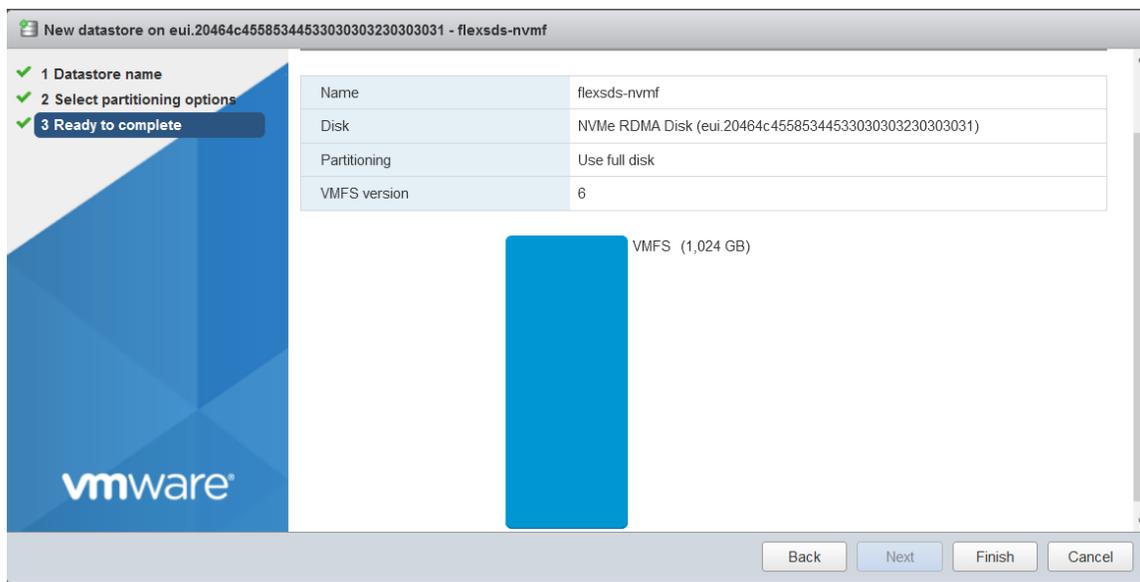
Click the Next button to continue.

Select partitioning options.



User can adjust partitions or leave it as there, click the Next button to continue.

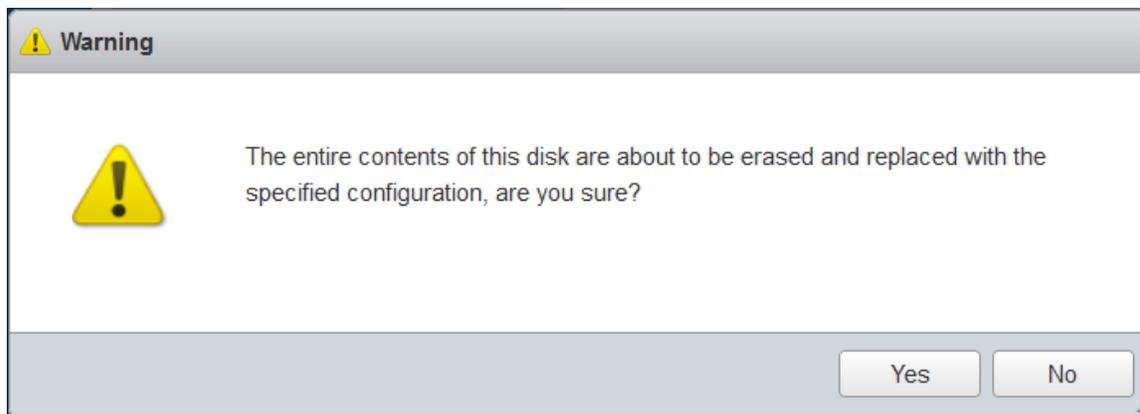
Ready to complete



User can review the settings, click Back if want to change anything.

Click the Finish button to create the new datastore on the FlexSDS NVMe over Fabrics volume.

Then a Warning dialog shows to ask user to confirm, click the Yes button to continue.



Navigate to the Datastores tab page.

After a while, new created datastore will be appear in the Datastores tab Page.

Name	Drive Type	Capacity	Provisio...	Free	Type	Thin pro...	Access
datastore1	SSD	5.81 TB	8.15 GB	5.8 TB	VMFS6	Supported	Single
flexsds-nvmf	SSD	1,023.75 GB	1.42 GB	1,022.33 GB	VMFS6	Supported	Single

flexsds-nvmf

Type: VMFS6
 Location: /vmfs/volumes/63e4a668-b12f714a-8b9e-3868dd0c58f8
 UUID: 63e4a668-b12f714a-8b9e-3868dd0c58f8
 Hosts: 1
 Virtual Machines: 0

STORAGE: USED: 1.42 GB FREE: 1,022.33 GB
 CAPACITY: 1,023.75 GB 0%

Enable the VMware iSER Adapter

Add iSER Adapter

Use the ESXi Shell or vSphere CLI to enable the VMware iSER storage adapter.

```
#esxcli rdma iser add
```

Verify that the iSER adapter has been added.

```
#esxcli iscsi adapter list
```

The output is similar to the following.

```
[[root@localhost:~] esxcli rdma iser add
[[root@localhost:~] esxcli iscsi adapter list
Adapter Driver State UID Description
-----
vmhba64 iser unbound iscsi.vmhba64 VMware iSCSI over RDMA (iSER) Adapter
vmhba65 iser unbound iscsi.vmhba65 VMware iSCSI over RDMA (iSER) Adapter
```

Adapter Driver State UID Description

vmhba64 iser unbound iscsi.vmhba64 VMware iSCSI over RDMA (iSER) Adapter

Specify the RoCE version that iSER uses to connect to the target.

Use the RoCE version of the RDMA capable adapter. The command you enter is similar to the following:

```
#esxcli rdma iser params set -a vmhba65 -r 2
```

This above step is depending on your RoCE network, both RoCE V1 and V2 are supported by FlexSDS.

When the command completes, a message similar to the following appears in the VMkernel log.

```
2023-01-18T10:30:47.324Z cpu30:2100415 opID=4bcbb158)iser: iser_set_roce: Setting roce type: 2  
for vmhba: vmhba65
```

If you do not specify the RoCE version, the host defaults to the highest RoCE version the RDMA capable adapter supports.

Set up the iSER adapter.

In the ESXi client, navigate to the Storage to review the list of adapters.

Select the "VMware iSCSI over RDMA (iSER) Adapter" and click the Configure iSCSI link, the Configure iSCSI – vmhba65 appears

Configure iSCSI - vmhba65

▶ Name & alias: iqn.1998-01.com.vmware.localhost:1144377393:65 (user-vmnic7)

▶ CHAP authentication: Do not use CHAP

▶ Mutual CHAP authentication: Do not use CHAP

▶ Advanced settings: Click to expand

Network port bindings

➕ Add port binding ➖ Remove port binding

VMkernel NIC	Port group	IPv4 address
vmk1	RDMA	192.168.20.200

Static targets

➕ Add static target ➖ Remove static target ✏ Edit settings

Target	Address	Port
No static targets		

Dynamic targets

➕ Add dynamic target ➖ Remove dynamic target ✏ Edit settings

Address	Port
192.168.20.120	3260

Click Add port binding, select the NIC has the port group named RDMA.

Click Add dynamic target link, and type storage server's IP address (RoCE network) in the text box.

Leave the default 3260 in the port box except you have changed iSCSI port in the storage settings.

Create Datastore on the iSER disk

Navigate to the Devices tab page, the new SCSI disk like FLEXSDS iSCSI Disk will be there, if it does not exist, click the Rescan link.

localhost.localdomain - Storage

Datstores Adapters **Devices** Persistent Memory

➕ New datastore ➕ Increase capacity 🔄 Rescan | 🔄 Refresh | ⚙ Actions

Name	Status	Type	Capacity	Queue ...	Vendor
 FLEXSDS iSCSI Disk (naa.600000000000000060d96...)	🟢 Normal	Disk	1,024 GB	113	FLEXSDS

12 items



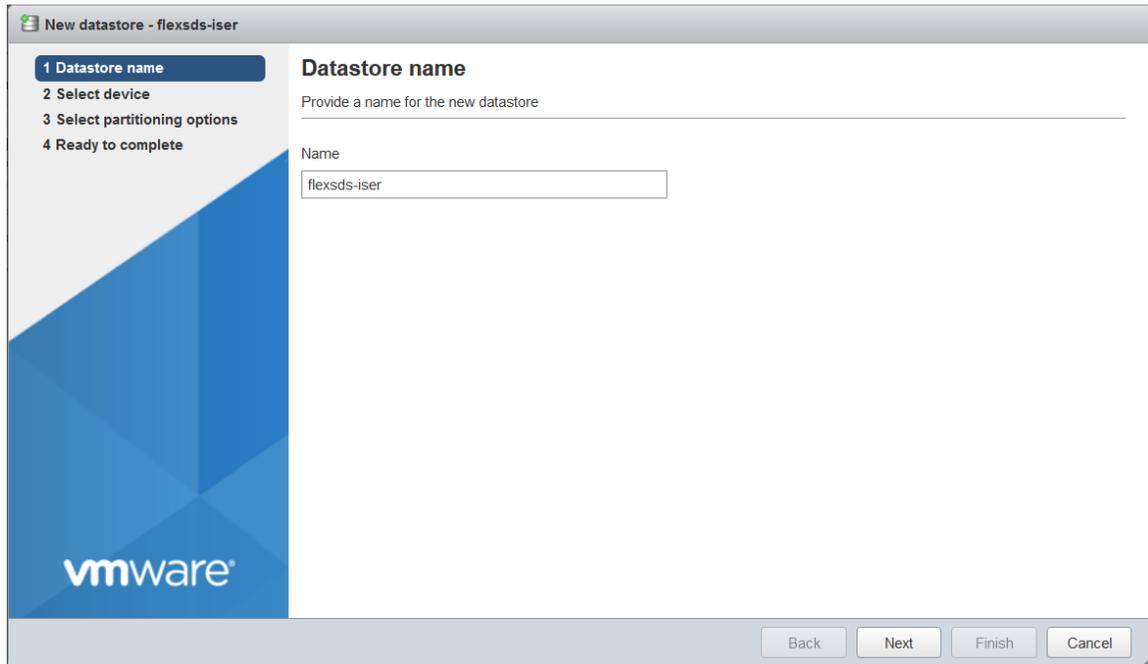
FLEXSDS iSCSI Disk

(naa.600000000000000060d96f553a4c04ef)

Type: Disk
 Model: FLEXSDS Controll
 Path: /vmfs/devices/disks/naa.600000000000000060d96f553a4c04ef
 Capacity: 1,024 GB
 Partition Format: gpt
 UUID: 020001030076c3000000000001464c45585344

After the FLEXSDS iSCSI Disk appears, click the New Datastore link.

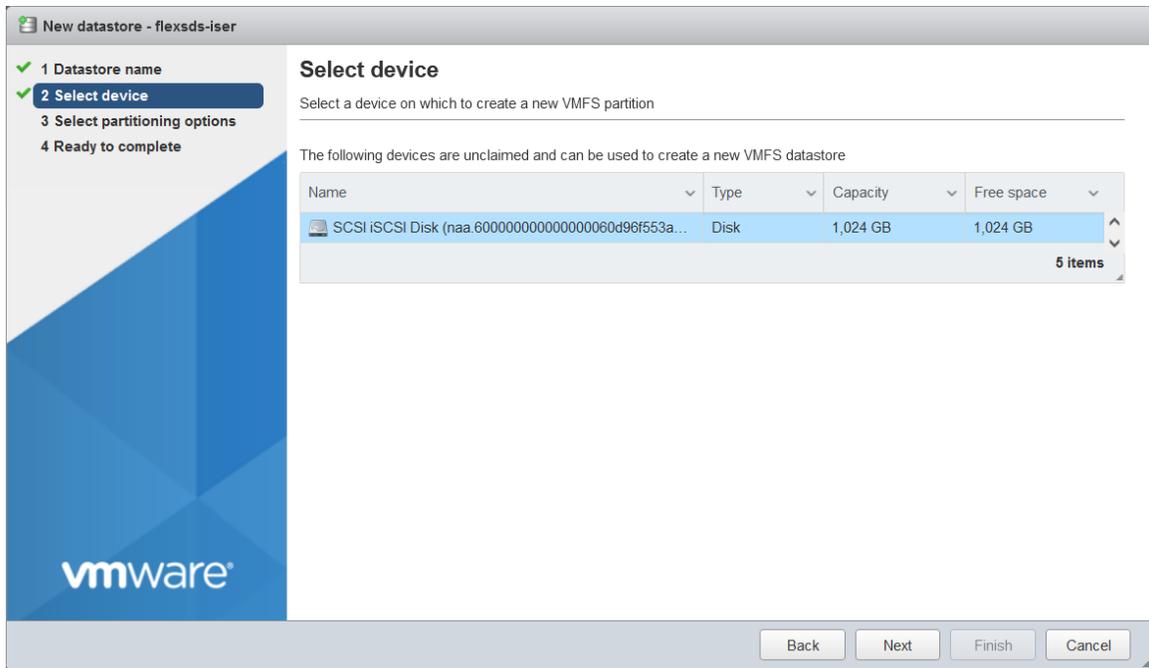
The New datastore – flexsds-iser window shows



Type in the Name, "flexsds-iser" as an example.

Click the Next button to continue.

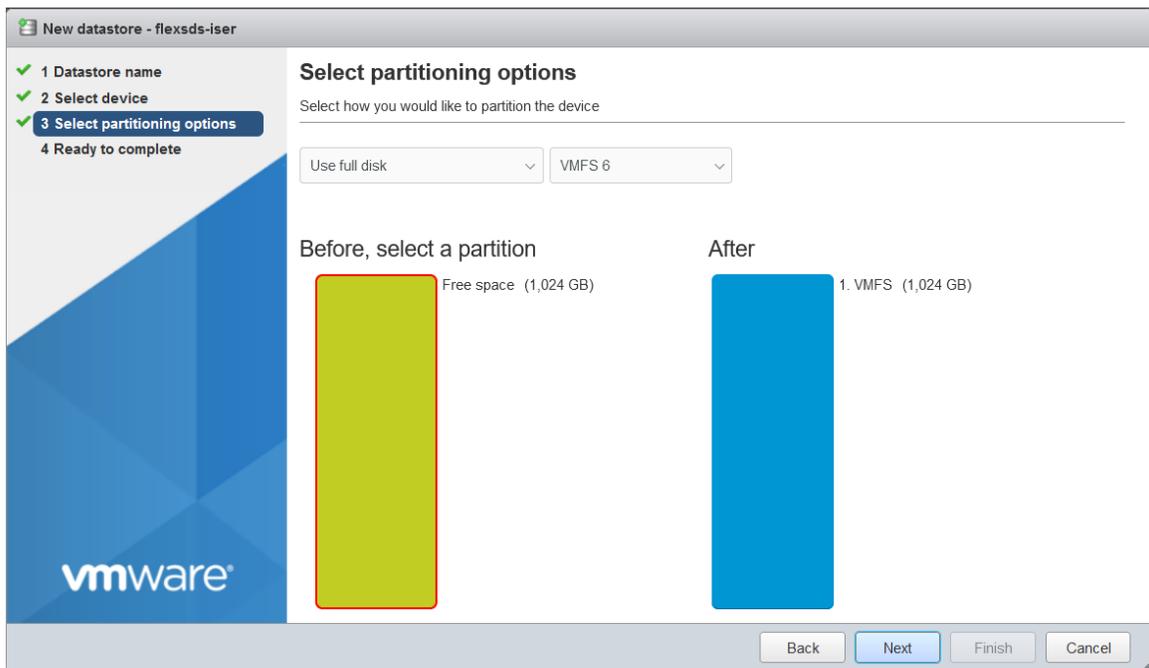
Select device



Select the SCSI iSCSI disk exported by FLEXSDS.

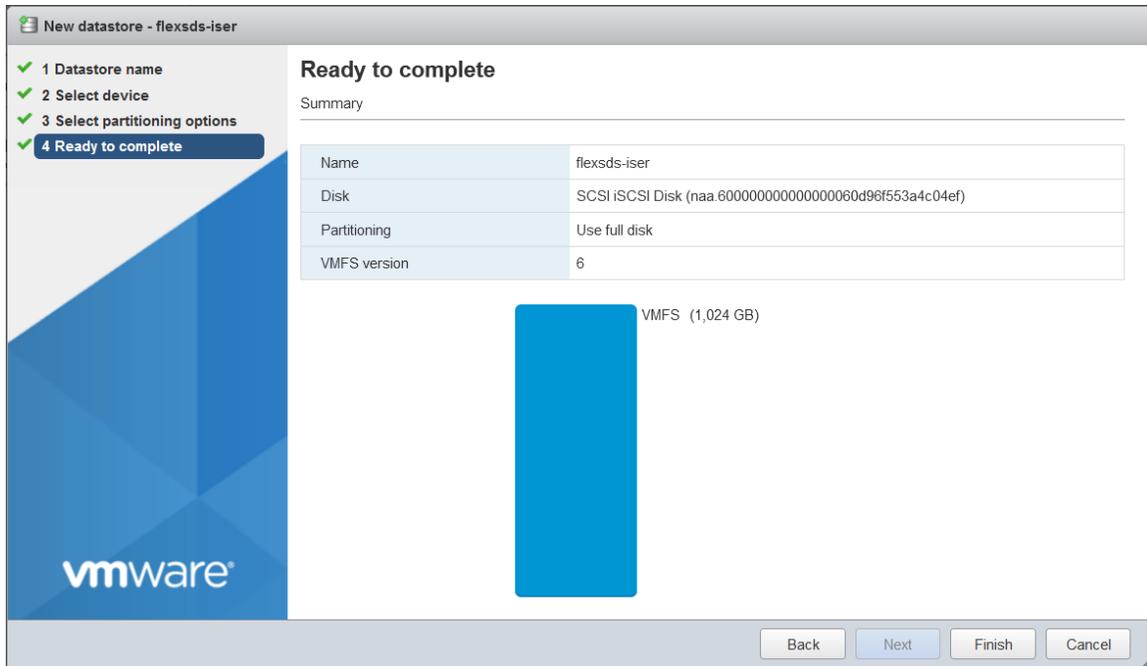
Click the Next button to continue.

Select partitioning options.



User can adjust partitions or leave it as there, click the Next button to continue.

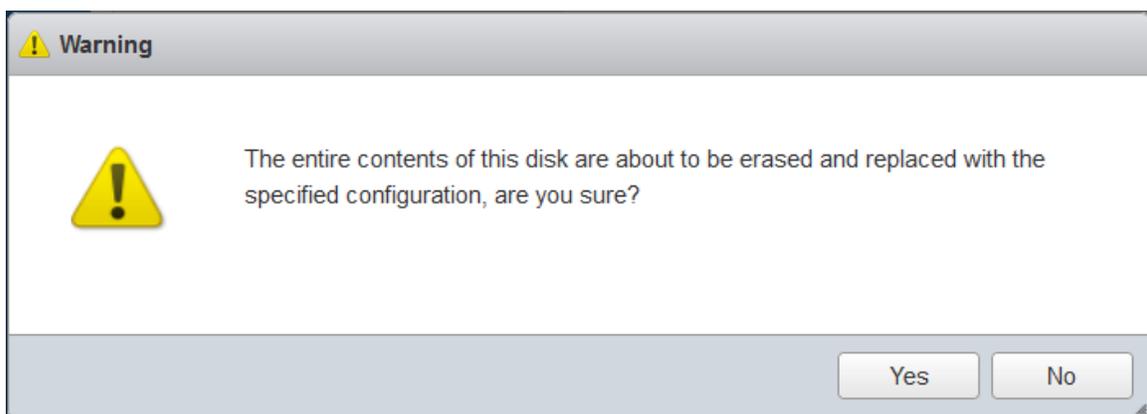
Ready to complete



User can review the settings, click Back if want to change anything.

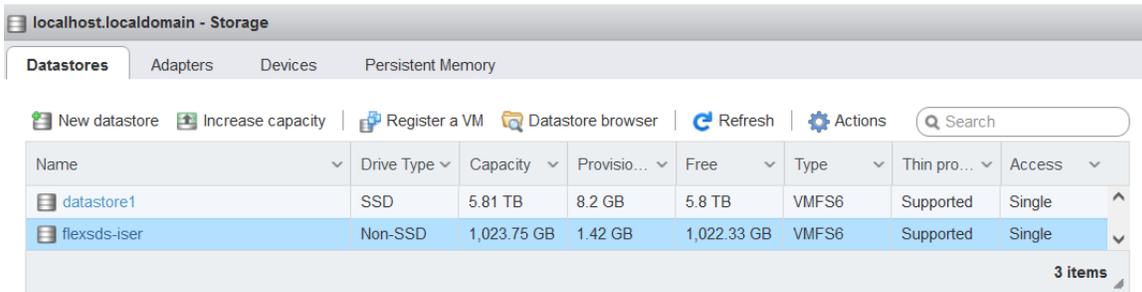
Click the Finish button to create the new datastore on the FlexSDS iSER volume.

Then a Warning dialog shows to ask user to confirm, click the Yes button to continue.



Navigate to the Datastores tab page.

After a while, new created datastore will be appear in the Datastores tab Page.



The screenshot shows the vSphere Storage interface for 'localhost.localdomain'. The 'Datastores' tab is active, displaying a table of datastores. The table has columns for Name, Drive Type, Capacity, Provisioning, Free space, Type, Thin provisioning, and Access. Two datastores are listed: 'datastore1' (SSD, 5.81 TB capacity, 8.2 GB provisioned, 5.8 TB free, VMFS6 type) and 'flexsds-iser' (Non-SSD, 1,023.75 GB capacity, 1.42 GB provisioned, 1,022.33 GB free, VMFS6 type). The 'flexsds-iser' datastore is selected. Below the table, there are icons for 'New datastore', 'Increase capacity', 'Register a VM', 'Datastore browser', 'Refresh', and 'Actions'. A search bar is also present.

Name	Drive Type	Capacity	Provisio...	Free	Type	Thin pro...	Access
datastore1	SSD	5.81 TB	8.2 GB	5.8 TB	VMFS6	Supported	Single
flexsds-iser	Non-SSD	1,023.75 GB	1.42 GB	1,022.33 GB	VMFS6	Supported	Single



flexsds-iser

Type: VMFS6
Location: /vmfs/volumes/63c8de86-466f0450-4167-3868dd0c5250
UUID: 63c8de86-466f0450-4167-3868dd0c5250
Hosts: 1
Virtual Machines: 0

STORAGE FREE: 1,022.33 GB 0%
USED: 1.42 GB CAPACITY: 1,023.75 GB

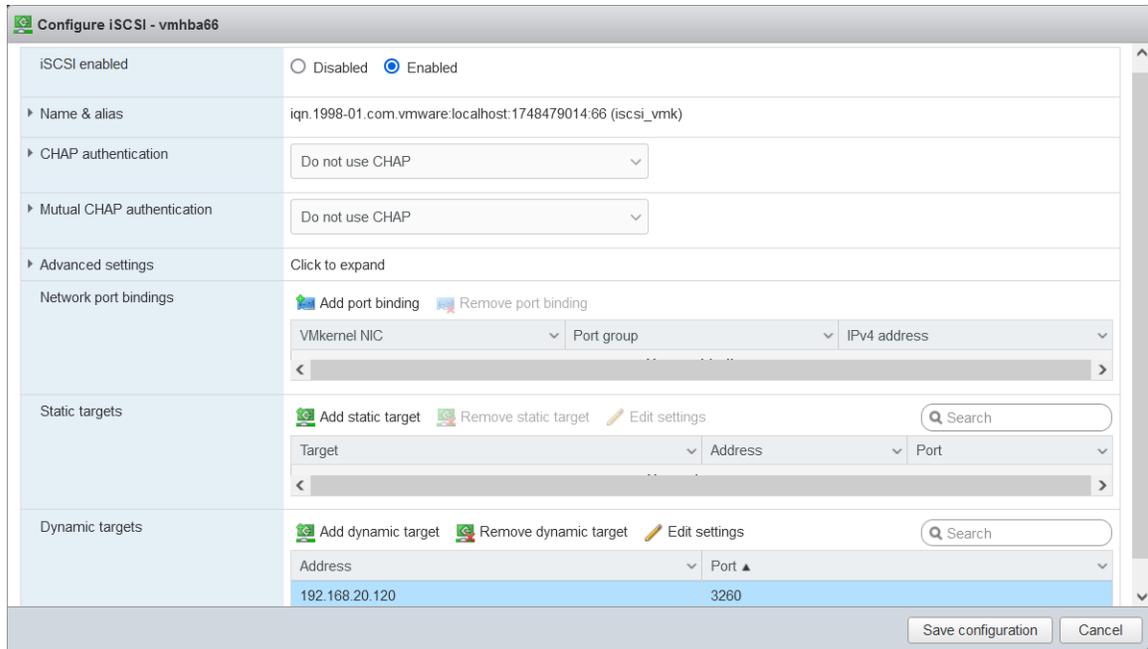
Enable iSCSI adapter.

This chapter we will enable iSCSI adapter and attach an iSCSI volume, this step is not required in high performance scenario, and here is mainly for performance comparison with the NVMe over RDMA and iSER volumes.

Set up iSCSI adapter

In the ESXi client, navigate to the Storage to review the list of adapters.

Select the "iSCSI Software Adapter" and click the Software iSCSI link, the Configure iSCSI – vmhba66 appears



Choose Enabled in the iSCSI Enabled option box.

Click Add port binding, select the NIC that will be used for iSCSI transport.

Click Add dynamic target link, and type storage server's IP address (the same RoCE network) in the text box.

Leave the default 3260 in the port box except you have changed iSCSI port in the storage settings.

Create Datastore on the iSCSI Disk

Navigate to the Devices tab page, the new SCSI disk like FLEXSDS iSCSI Disk will be there, if it does not exist, click the Rescan link.

localhost.localdomain - Storage

Datastores Adapters **Devices** Persistent Memory

New datastore Increase capacity Rescan Refresh Actions Search

Name	Status	Type	Capacity	Queue Depth	Vendor
FLEXSDS iSCSI Disk (naa.6000000000000000af16551fe223d241)	Normal	Disk	1,024 GB	128	FLEXSDS

12 items



Local NVMe Disk

(t10.NVMe____SSDPE2KE064T8L_____00013B397EE4D25C)

Type: Disk
 Model: SSDPE2KE064T8L
 Path: /vmfs/devices/disks/t10.NVMe____SSDPE2KE064T8L_____00013B397EE4D25C
 Capacity: 5.82 TB
 Partition Format: Unknown
 UUID: 055777bd5caafa3f8b97aceb672f41e9e9f1ac48addcaf6821be18fbae753b74f8

After the SCSI disk FLEXSDS iSCSI Disk appears, click the New Datastore link.

New datastore - flexsds-iscsi

- 1 Datastore name**
- 2 Select device
- 3 Select partitioning options
- 4 Ready to complete

Datastore name
 Provide a name for the new datastore

Name

vmware

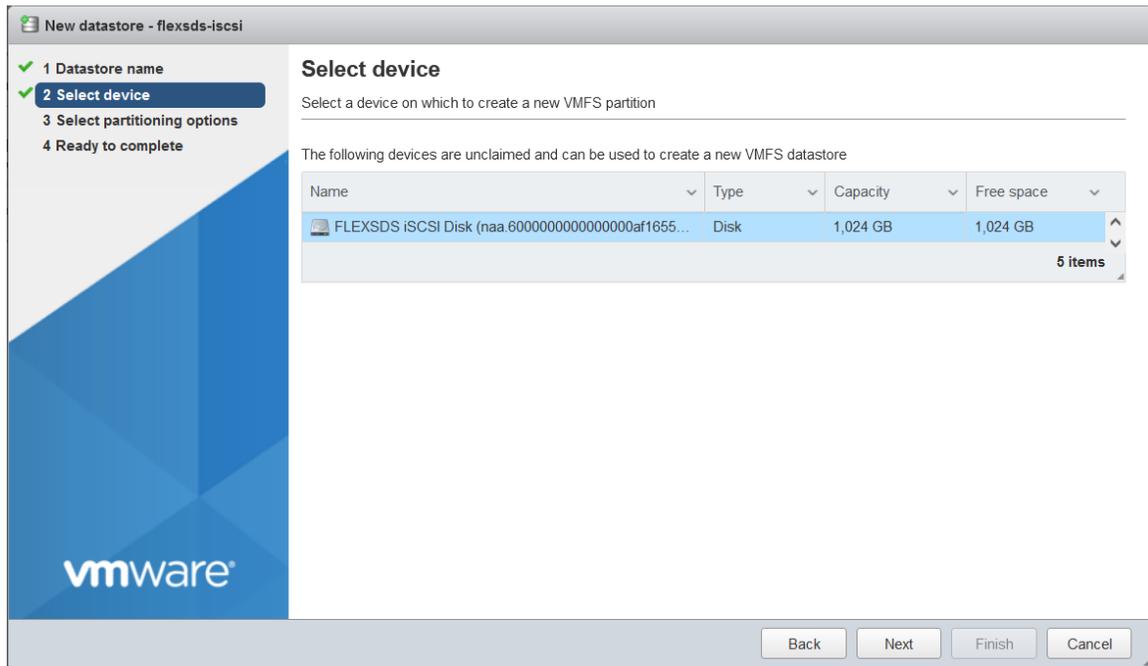
Back Next Finish Cancel

The New datastore – flexsds-iscsi window shows

Type in the Name, "flexsds-iscsi" as an example.

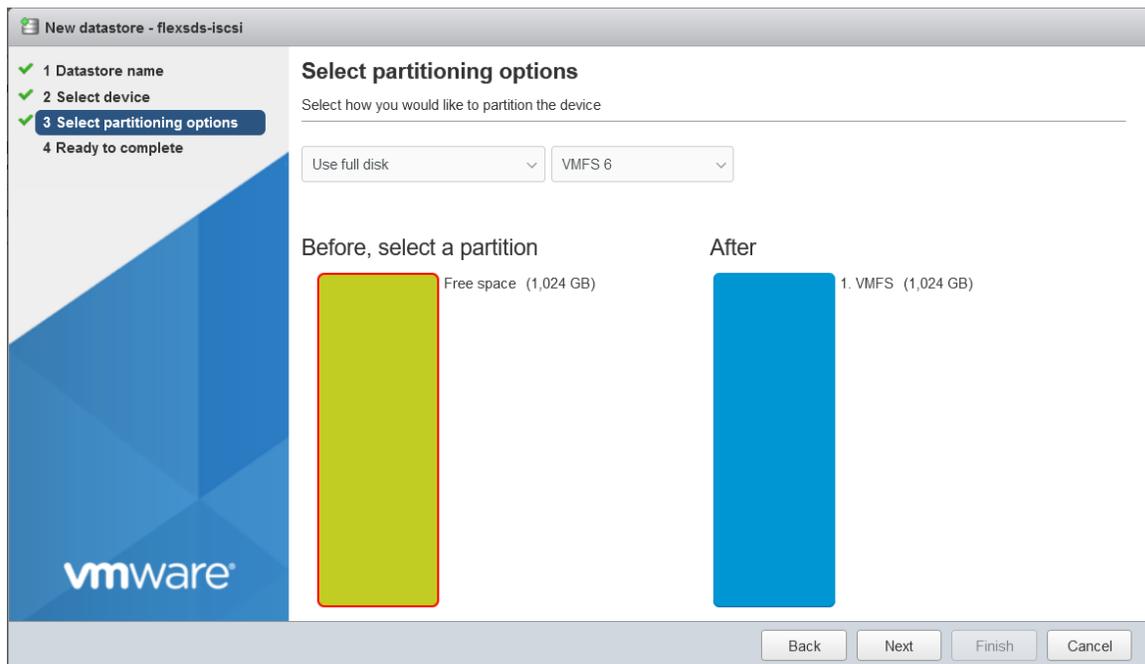
Click the Next button to continue.

Select device



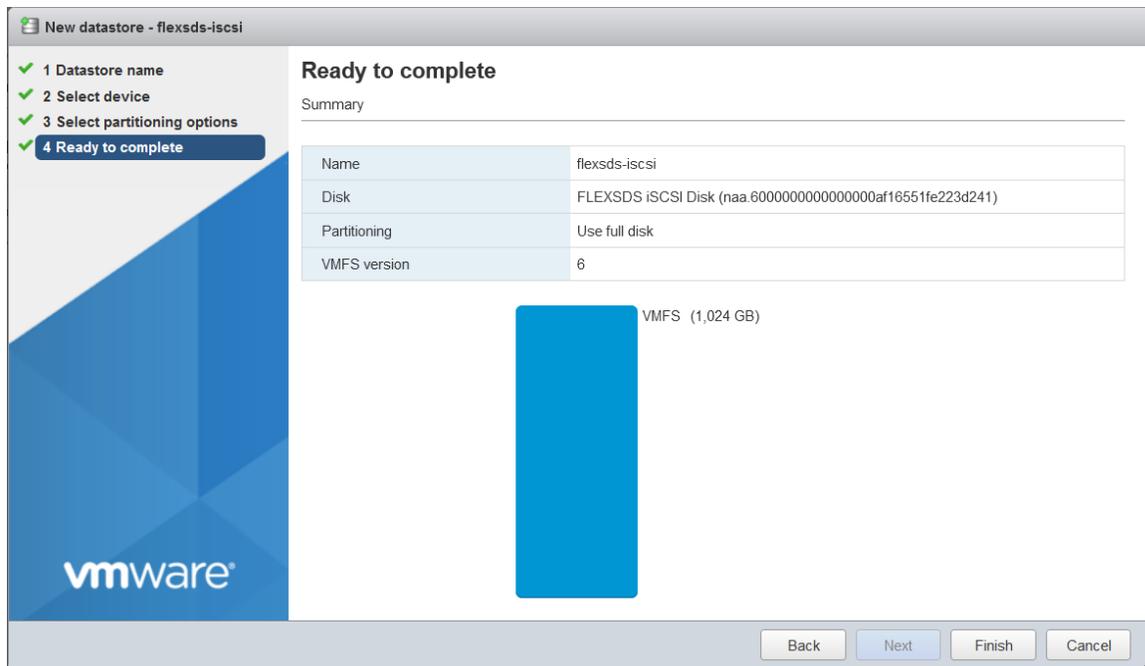
Select the Disk, click the Next button to continue.

Select partitioning options.



User can adjust partitions or leave it as there, click the Next button to continue.

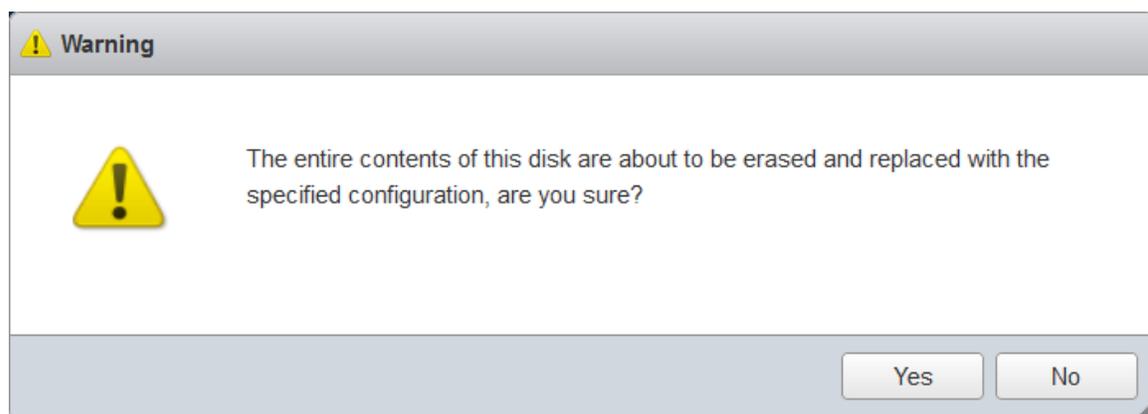
Ready to complete



User can review the settings, click Back if want to change anything.

Click the Finish button to create the new datastore on the flexsds-iscsi volume.

Then a Warning dialog shows to ask user to confirm, click the Yes button to continue.



Navigate to the Datastores tab page.

After a while, new created datastore will be appear in the Datastores tab Page.

localhost.localdomain - Storage

Datstores Adapters Devices Persistent Memory

New datastore
 Increase capacity
 Register a VM
 Datastore browser
 Refresh
 Actions

Name	Drive Type	Capacity	Provisio...	Free	Type	Thin pro...	Access
datastore1	SSD	5.81 TB	8.2 GB	5.8 TB	VMFS6	Supported	Single
flexsds-iscsi	Non-SSD	1,023.75 GB	1.42 GB	1,022.33 GB	VMFS6	Supported	Single

3 items



flexsds-iscsi

Type: VMFS6
 Location: /vmfs/volumes/63c902f3-f1ce3146-5494-3868dd0c5250
 UUID: 63c902f3-f1ce3146-5494-3868dd0c5250
 Hosts: 1
 Virtual Machines: 0

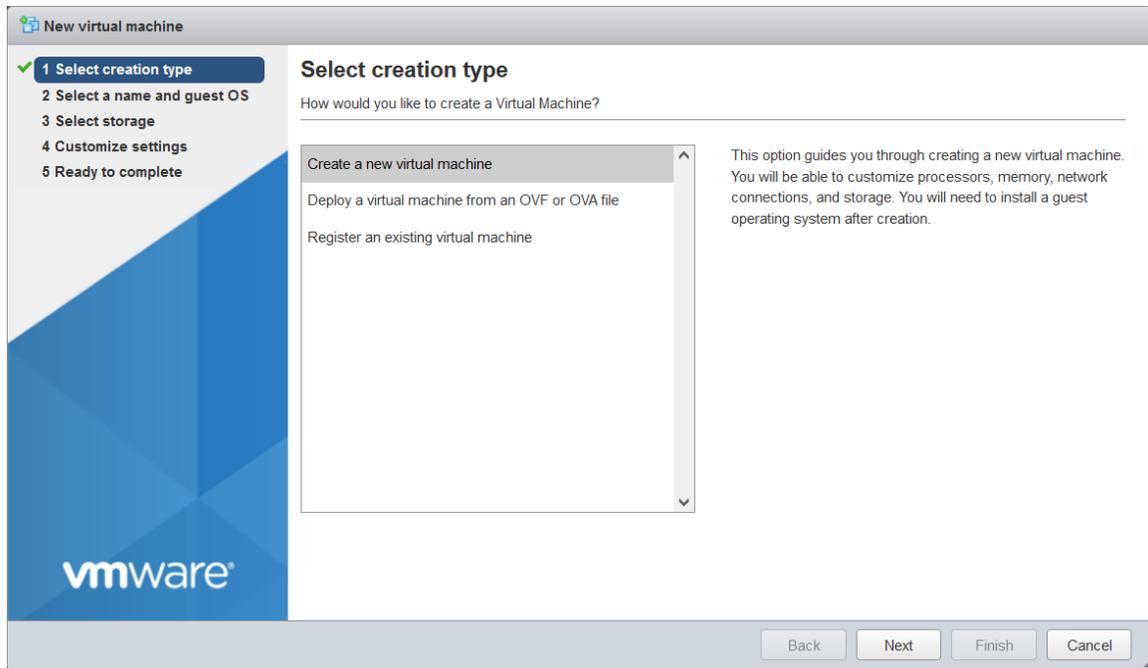
STORAGE FREE: 1,022.33 GB
 USED: 1.42 GB CAPACITY: 1,023.75 GB

Create Virtual Machine

Navigate to the Virtual Machines page

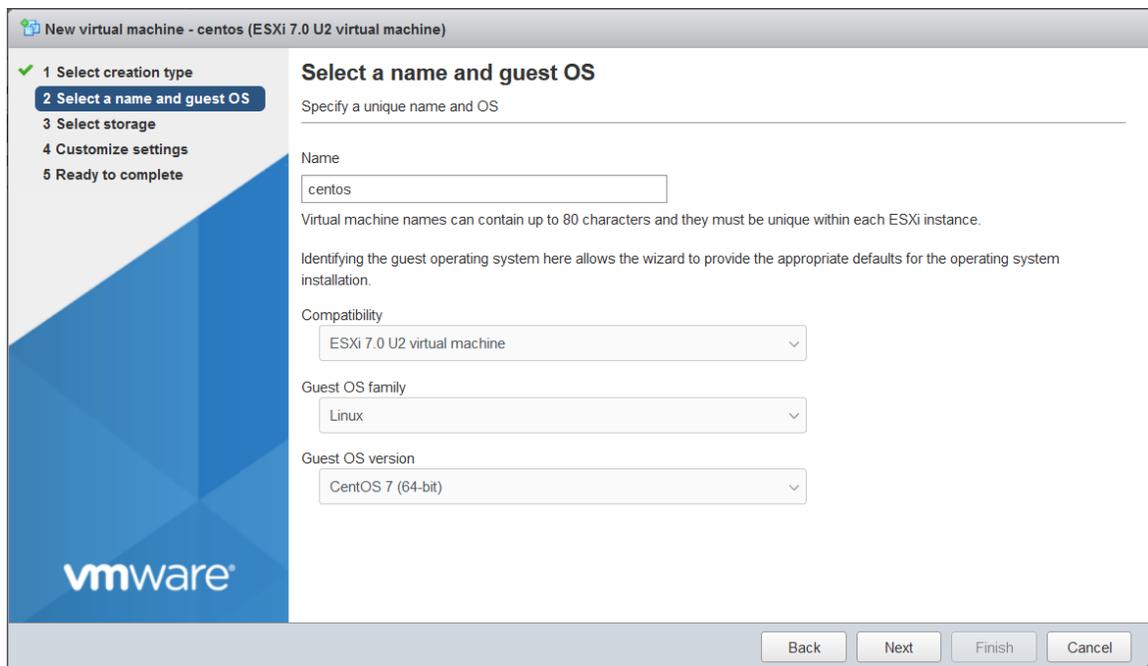
Click the Create / Register VM link

the New virtual machine window shows



Select Create a new virtual machine, click the Next button to continue.

Select a name and guest OS



Type Name in the virtual machine

User can set virtual machine type as he/she needs, we take CentOS 7 64-bit as an example.

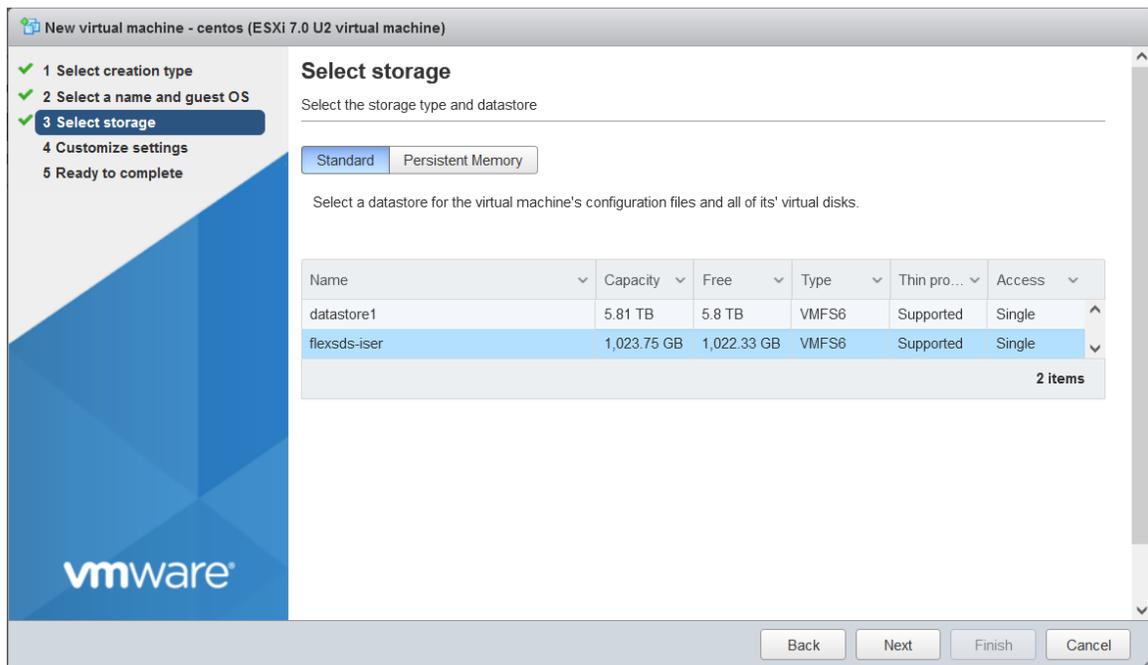
Choose ESXi 7.0 U2 virtual machine in Compatibility

Choose Linux in Guest OS family

Choose CentOS 7 (64-bit) in Guest OS version.

Click the Next button to continue.

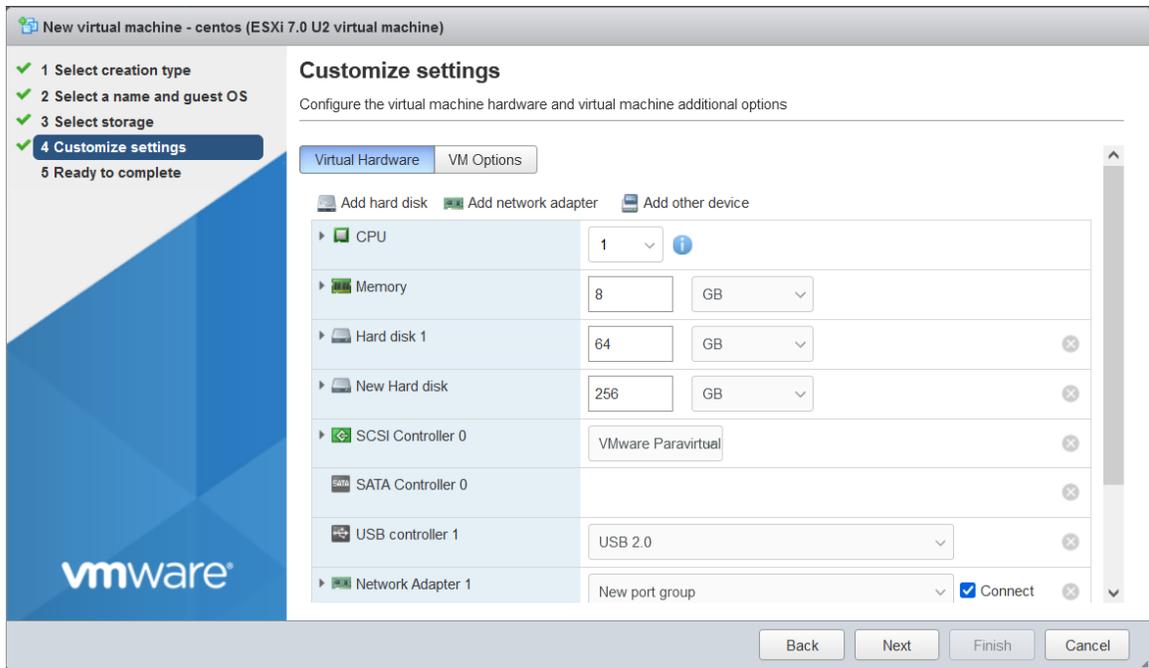
Select storage



Select the new create datastore, here could be flexsds-nvmf, flexsds-iser or flex-iscsi, that depends on that type of datastore created in the above steps.

Click the Next button to continue.

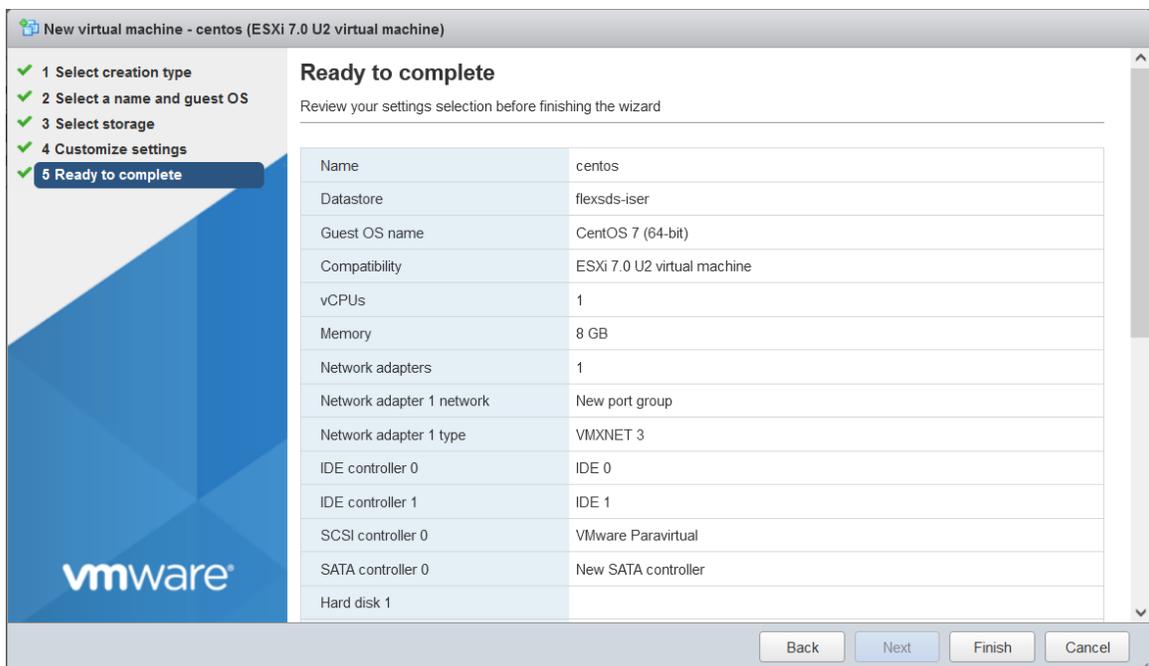
Customize settings.



User can customize virtual machine's settings, here we attached CentOS 7 ISO for later OS installation, and for performance testing later, here we attached another disk on flexsds-nvmf, flexsds-iser datastore and flexsds-iscsi datastore for performance comparing.

Then click the Next button to continue.

Ready to complete

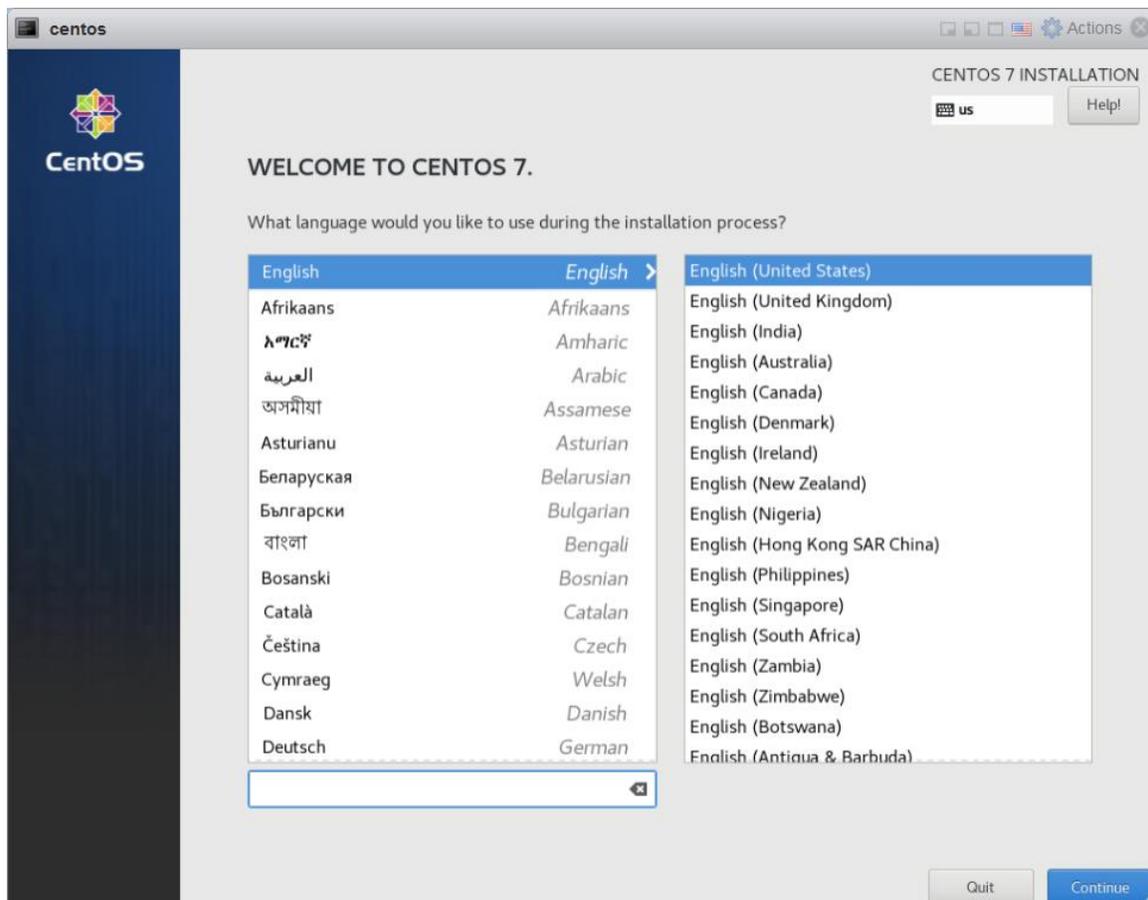


User can review the virtual machine's settings, if need anything to change, click the Back button otherwise click the Finish button to complete the Virtual Machine creating.

Select the new create Virtual Machine.

Click the Power On button.

Then click the Console link to manage the virtual machine



User can manage and do initial setup to the virtual machine just like a physical machine.

Test Speed (optional)

To test storage performance of virtual machine's disk on VMFS, we created three scenarios, 1 is connecting to FlexSDS via NVMe-oF (RDMA), the second is connecting via iSER, the third is via iSCSI. The virtual machine has two disks in the same datastore: /dev/sda for OS and /dev/sdb for this testing.

```
[root@localhost ~]# lsblk
NAME                MAJ:MIN RM  SIZE RO TYPE MOUNTPOINT
sda                  8:0    0   64G  0 disk
├─sda1                8:1    0   200M  0 part /boot/efi
├─sda2                8:2    0    1G    0 part /boot
└─sda3                8:3    0  62.8G  0 part
   ├─centos-root      253:0    0  37.9G  0 lvm  /
   ├─centos-swap      253:1    0   6.4G  0 lvm  [SWAP]
   └─centos-home      253:2    0  18.5G  0 lvm  /home
sdb                  8:16    0  256G  0 disk
sr0                  11:0    1  1024M  0 rom
```

Login to the virtual machine via ssh and install fio.

```
#yum install -y fio
```

Fill up fio test configuration file (test.fio as an example) with the content below:

```
[global]

runtime=36000

time_based

group_reporting

direct=1

size=200G

bs=4k

ioengine=libaio

rw=rw

numjobs=8
```

iodepth=64

rwmixread=0

[block]

filename=/dev/sdb

Issue the following command to start test:

#fio test.fio

```
[root@localhost ~]# fio test.fio
block: (g=0): rw=rw, bs=(R) 4096B-4096B, (W) 4096B-4096B, (T) 4096B-4096B, ioengine=libaio, iodepth=64
...
fio-3.7
Starting 8 processes
[ ] obs: 8 (f=8): [W(8)][0.5%][r=0KiB/s,w=853MiB/s][r=0,w=218k IOPS][eta 09h:57m:13s]
```

Here is the detailed comparing table between different test cases and user will see there are much performance improve from iSCSI to NVMe over RDMA or iSER.

NVMe-oF (RDMA)		
Operation	IOPS	Band width
4K Read	272K	1061Mib/s
4K Write	215K	838MiB/s
4K Read 70%	290.8K	1137Mib/s
256K Read	42K	11010Mib/s
256K Write	19.2K	5033Mib/s

iSER		
Operation	IOPS	Band width
4K Read	251K	981Mib/s
4K Write	218K	853MiB/s
4K Read 70%	245.2K	958Mib/s
256K Read	36.7K	9620Mib/s
256K Write	16.3K	4068Mib/s

iSCSI		
Operation	IOPS	Band width
4K Read	64.2K	251Mib/s
4K Write	66.8K	261MiB/s

4K Read 70%	63.9K	249.3Mib/s
256K Read	5.75K	1438Mib/s
256K Write	3.103K	776Mib/s

Conclusion

Using NVMe over RDMA and iSER interface will gain much more performance than iSCSI, NVMe over RDMA is a little better than iSER, and although there is performance cost over VMFS and virtual machine, it almost reaches to hardware limits while testing I/O with 256k, as write operation will write two NVMe's at the same time, that drop write performance 50%, that already fully utilized NVMe hardware performance.

Contact

Support: support@flexsds.com
Sales: sales@flexsds.com
Home Page: <http://www.flexsds.com/>
Product Page: <https://www.flexsds.com/scale-out-storage/>
Purchase <https://www.flexsds.com/pricing/>
Knowledge Base: <https://www.flexsds.com/support/kb/>

FlexSDS Software Limited.

www.flexsds.com

Copyright © FlexSDS Software Limited 2016-2023. All right reserved.