Lenovo

ThinkSystem® SAN OS 11.70.1

Installing and Configuring for Linux®

Express Guide

Note

Before using this information and the product it supports, be sure to read and understand the safetyinformation and the safety instructions, which are available at: http://thinksystem.lenovofiles.com/help/topic/safety_documentation/pdf_files.html

In addition, be sure that you are familiar with the terms and conditions of the Lenovo warranty for your server,which can be found at: http://datacentersupport.lenovo.com/warrantylookup

Fourth Edition (October 2021)

TABLE OF CONTENTS

Deciding whether to use this Express Guide 1

Understanding the workflow 3

Verifying the configuration is supported 4

Configuring management port IP addresses 10

Access ThinkSystem System Manager and use the Setup Wizard 13

Install ThinkSystem Host Utilities 15

Performing FC-specific tasks 16

Determining host WWPNs and making the recommended settings - FC 16

Configuring the switches - FC 16

Configure the multipath software 17

Create partitions and filesystems 18

Verify storage access on the host 19

FC worksheet 19

Performing iSCSI-specific tasks 21

Configuring the switches - iSCSI 21

Configuring networking - iSCSI 21

Configuring array-side networking - iSCSI 20

Configuring host-side networking - iSCSI 22

Verifying IP network connections—iSCSI 24

Configure the multipath software 24

Create partitions and filesystems 26

Verify storage access on the host 26

iSCSI worksheet 27

Performing SAS-specific tasks 28

Determining SAS host identifiers 28

Configure the multipath software 28

Create partitions and filesystems 29

Verify storage access on the host 30

SAS worksheet 30

Performing NVMe over RoCE-specific tasks 32

Verify the Linux configuration is supported 32

Configure the switch 32

Set up NVMe over RoCE on the host side 32

Configure storage array NVMe over RoCE connections 33

Discover and connect to the storage from the host 35

Set up failover on the host 36

Create filesystems 38

NVMe over RoCE worksheet for Linux 39

Performing NVMe over Fibre Channel tasks 42

Verify the Linux configuration is supported 42

Configure the switch 42

Set up NVMe over Fibre Channel on the host side 43
Display the volumes visible to the host 43
Set up failover on the host 44
Create filesystems 45
NVMe over Fibre Channel worksheet for Linux 45
Creating a workload 48
Create workloads 48
Create volumes 49
Step 1: Select host for a volume 49
Step 2: Select a workload for a volume 50
Step 3: Add or edit volumes 51
Step 4: Review volume configuration 55
Defining a host in ThinkSystem System Manager 57
Create host automatically 57
Create host manually 57
Create host cluster 59
Mapping a volume to a host 61
Discovering, Configuring, and Verifying storage on the host 62
Where to find additional information 63
Contacting Support 64
Notices 65
Trademarks 66

Deciding whether to use this Express Guide

The express method for installing your storage array and accessing ThinkSystem System Manager is appropriate for setting up a standalone Linux host to a DE Series storage system. It is designed to get the storage system up and running as quickly as possible with minimal decision points.

The express method includes the following steps:

1. Setting up one of the following communication environments:
    - Fibre Channel (FC)
    - iSCSI
    - SAS
    - NVMe over RoCE
    - NVMe over Fibre Channel
2. Creating logical volumes on the storage array.
3. Making the volume LUNs available to the data host.

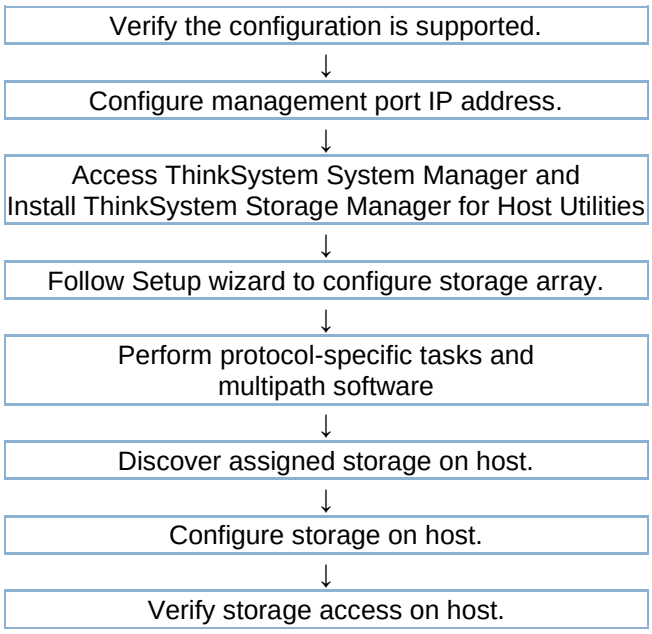
This guide is based on the following assumptions:

Component	Assumptions
Hardware	• You have used the Installation and Setup Instructions included with the controller shelves to install the hardware. • You have connected cables between the optional drive shelves and the array controllers. • You have applied power to the storage array. • You have installed all other hardware (for example, management station, switches) and made the necessary connections. • If you are using NVMe over Fabrics, each DE6000H or DE6000F controller contains at least 64 GB of RAM.
Host	• You have made a connection between the storage array and the data host. • You have installed the host operating system. • You are not using Windows as a virtualized guest. • You are not configuring the data (I/O attached) host to boot from SAN. • If you are NVMe over Fabrics, you have installed the latest compatible Linux version as listed under the Lenovo Storage Interoperation Center.
Storage management station	• You are using a 1 Gbps or faster management network.

Component	Assumptions
	• You are using a separate station for management rather than the data (I/O attached) host. • You are using out-of-band management, in which a storage management station sends commands to the storage array through the Ethernet connections to the controller. • You have attached the management station to the same subnet as the storage management ports.
IP addressing	• You have installed and configured a DHCP server. • You have **not** yet made an Ethernet connection between the management station and the storage array.
Storage provisioning	• You will not use shared volumes. • You will create pools rather than volume groups.
Protocol: FC	• You have made all host-side FC connections and activated switch zoning. • You are using Lenovo-supported FC HBAs and switches. • You are using FC HBA driver versions as listed on Lenovo Storage Interoperation Center (LSIC).
Protocol: iSCSI	• You are using Ethernet switches capable of transporting iSCSI traffic. • You have configured the Ethernet switches according to the vendor's recommendation for iSCSI.
Protocol: SAS	• You are using Lenovo-supported SAS HBAs. • You are using SAS HBA driver versions as listed on Lenovo Storage Interoperation Center (LSIC).
Protocol: NVMe over RoCE	• You have received the 100G host interface cards in a DE6000H and DE6000F storage system pre-configured with the NVMe over RoCE protocol • You are using RDMA enabled NIC(RNIC) driver versions as listed on Lenovo Storage Interoperation Center (LSIC).
Protocol: NVMe over Fibre Channel	• You have received the 32G host interface cards in a DE6000H and DE6000F storage system pre-configured with the NVMe over Fibre Channel protocol or the controllers were ordered with standard FC ports and need to be converted to NVMe-oF. • You are using FC-NVMe HBA driver versions as listed on Lenovo Storage Interoperation Center (LSIC).

Understanding the workflow

This workflow guides you through the express method for configuring your storage array and ThinkSystem System Manager to make storage available to a host.



Verifying the configuration is supported

To ensure reliable operation, you create an implementation plan and then verify that the entire configuration is supported.

1. Go to the Lenovo Storage Interoperation Center (LSIC).
2. Follow the Guidance for LSIC & Help.

   In this file, you may search for the product family that applies, as well as other criteria for the configuration such as Operating System, ThinkSystem SAN OS, and Host Multipath driver.
3. As necessary, make the updates for your operating system and protocol as listed in the table.

Operating system updates	Protocol	Protocol-related updates
You might need to install out-of-box drivers to ensure proper functionality and supportability. Each HBA vendor has specific methods for updating boot code and firmware. Refer to the support section of the vendor's website to obtain the instructions and software necessary to update the HBA boot code and firmware.	FC	Host bus adapter (HBA) driver, firmware, and bootcode
	iSCSI	Network interface card (NIC) driver, firmware and bootcode
	SAS	Host bus adapter (HBA) driver, firmware, and bootcode
	NVMe over RoCE	RDMA enabled NIC(RNIC) driver, firmware, and bootcode
	NVMe over Fibre Channel	FC-NVMe HBA driver, firmware, and bootcode

Configuring management port IP addresses

In this express method for configuring communications between the management station and the storage array, you use Dynamic Host Configuration Protocol (DHCP) to provide IP addresses. Each controller has two storage management ports, and each management port will be assigned an IP address.

Before you begin

You have installed and configured a DHCP server on the same subnet as the storage management ports.
The following instructions refer to a storage array with two controllers (a duplex configuration).

1. If you have not already done so, connect an Ethernet cable to the management station and to management port 1 on each controller (A and B).

   The DHCP server assigns an IP address to port 1 of each controller.

   Note: Do not use management port 2 on either controller. Port 2 is reserved for use by Lenovo technical personnel.

   Important: If you disconnect and reconnect the Ethernet cable, or if the storage array is power-cycled, DHCP assigns IP addresses again. This process occurs until static IP addresses are configured. It is recommended that you avoid disconnecting the cable or power-cycling the array.

   If the storage array cannot get DHCP-assigned IP addresses within 30 seconds, the following default IP addresses are set:

   - Controller A, port 1: 169.254.128.101
   - Controller B, port 1: 169.254.128.102
   - Subnet mask: 255.255.0.0

2. Locate the MAC address label on the back of each controller, and then provide your network administrator with the MAC address for port 1 of each controller.

   Your network administrator needs the MAC addresses to determine the IP address for each controller. You will need the IP addresses to connect to your storage system through your browser.

Access ThinkSystem System Manager and use the Setup Wizard

You use the Setup wizard in ThinkSystem System Manager to configure your storage array.

Before you begin

- You have ensured that the device from which you will access ThinkSystem System Manager contains one of the following browsers:

Browser	Minimum version
Google Chrome	47
Microsoft Internet Explorer	11
Microsoft Edge	EdgeHTML 12
Mozilla Firefox	31
Safari	9

- You are using out-of-band management.

If you are an iSCSI user, make sure you have closed the Setup wizard while configuring iSCSI.

The wizard automatically relaunches when you open System Manager or refresh your browser and *at least one* of the following conditions is met:

- No pools and volume groups are detected.
- No workloads are detected.
- No notifications are configured.

If the Setup wizard does not automatically appear, contact technical support.

1. From your browser, enter the following URL: https://<DomainNameOrIPAddress>

   IP Address is the address for one of the storage array controllers.

   The first time ThinkSystem System Manager is opened on an array that has not been configured, the Set Administrator Password prompt appears. Role-based access management configures four local roles: **admin**, **support**, **security**, and **monitor**. The latter three roles have random passwords that cannot be guessed. After you set a password for the **admin** role you can change all of the passwords using the **admin** credentials. See *ThinkSystem System Manager online help* for more information on the four local user roles.

2. Enter the System Manager password for the **admin** role in the Set Administrator Password and Confirm Password fields, and then select the **Set Password** button.

   When you open System Manager and no pools, volumes groups, workloads, or notifications have been configured, the Setup wizard launches.

3. Use the Setup wizard to perform the following tasks:

   - **Verify hardware (controllers and drives)** – Verify the number of controllers and drives in the storage array. Assign a name to the array.
   - **Verify hosts and operating systems** – Verify the host and operating system types that the storage array can access.

- **Accept pools** – Accept the recommended pool configuration for the express installation method. A pool is a logical group of drives.

- **Configure alerts** – Allow System Manager to receive automatic notifications when a problem occurs with the storage array.

- **Enable AutoSupport** – Automatically monitor the health of your storage array and have dispatches sent to technical support.

4. If you have not already created a volume, create one by going to **Storage** > **Volumes** > **Create** > **Volume**.

   For more information, see the online help for ThinkSystem System Manager.

Install ThinkSystem Host Utilities

Storage Manager (Host Utilities) can only be installed on host servers.

1. Download the ThinkSystem Host Utilities package from DE Series Product Support Site.
2. Install the ThinkSystem Host Unitilites binary.

Performing FC-specific tasks

For the Fibre Channel protocol, you configure the switches and determine the host port identifiers.

Determining host WWPNs and making the recommended settings - FC

You install an FC HBA utility so you can view the worldwide port name (WWPN) of each host port.

Guidelines for HBA utilities:

- Most HBA vendors offer an HBA utility. You will need the correct version of HBA for your host operating system and CPU. Examples of FC HBA utilities include:
    - Emulex OneCommand Manager for Emulex HBAs
    - QLogic QConverge Console for QLogic HBAs
- Host I/O ports might automatically register if the host context agent is installed.

1. Download the appropriate utility from your HBA vendor's web site.
2. Install the utility.
3. Select the appropriate settings in the HBA utility.

Configuring the switches - FC

Configuring (zoning) the Fibre Channel (FC) switches enables the hosts to connect to the storage array and limits the number of paths. You zone the switches using the management interface for the switches.

Before you begin

- You must have administrator credentials for the switches.
- You must have used your HBA utility to discover the WWPN of each host initiator port and of each controller target port connected to the switch.

For details about zoning your switches, see the switch vendor's documentation.

You must zone by WWPN, not by physical port. Each initiator port must be in a separate zone with all of its corresponding target ports.

1. Log in to the FC switch administration program, and then select the zoning configuration option.
2. Create a new zone that includes the first host initiator port and that also includes all of the target ports that connect to the same FC switch as the initiator.
3. Create additional zones for each FC host initiator port in the switch.
4. Save the zones, and then activate the new zoning configuration.

Configure the multipath software

Multipath software provides a redundant path to the storage array in case one of the physical paths is disrupted. The multipath software presents the operating system with a single virtual device that represents the active physical paths to the storage. The multipath software also manages the failover process that updates the virtual device. You use the device mapper multipath (DM-MP) tool for Linux installations.

Before you begin

You have installed the required packages on your system.

- For Red Hat (RHEL) hosts, verify the packages are installed by running **rpm -q device-mapper- multipath**.
- For SLES hosts, verify the packages are installed by running **rpm -q multipath-tools**.

By default, DM-MP is disabled in RHEL and SLES. Complete the following steps to enable DM-MP components on the host.

If you have not already installed the operating system, use the media supplied by your operating system vendor.

1. If a multipath.conf file is not already created, run the **# touch /etc/multipath.conf** command.
2. Use the default multipath settings by leaving the multipath.conf file blank.
3. Start the multipath service.
   # systemctl start multipathd
4. Configure multipath for startup persistence.
   # chkconfig multipathd on
5. Save your kernel version by running the uname -r command.
   # uname -r
6. Do one of the following to enable the multipathd daemon on boot.

If you are using....	Do this...
RHEL 6.x systems:	**chkconfig multipathd on**
RHEL 7.x and 8.x systems:	**systemctl enable multipathd**
SLES 12.x and 15.x systems:	**systemctl enable multipathd**

7. Rebuild the initramfs image or the initrd image under /boot directory:

If you are using....	Do this...
RHEL 6.x and 7.x systems:	**dracut --force --add multipath**
SLES 12.x and 15.x systems:	**dracut --force --add multipath**

8. Make sure that the newly created /boot/initrams-* image or /boot/initrd-* image is selected in the boot configuration file.For example, for grub it is /boot/grub/menu.lst and for grub2 it is /boot/ grub2/menu.cfg.
9. Use the "Create host manually" procedure in the online help to check whether the hosts are defined.

Verify that each host type is either Linux DM-MP (Kernel 3.10 or later) if you enable the Automatic Load Balancing feature, or Linux DM-MP (Kernel 3.9 or earlier) if you disable the Automatic Load Balancing feature. If necessary, change the selected host type to the appropriate setting.

10. Reboot the host.

Setting up the multipath.conf file

The multipath.conf file is the configuration file for the multipath daemon, multipathd. The multipath.conf file overrides the built-in configuration table for multipathd. Any line in the file whose first non-white-space character is # is considered a comment line. Empty lines are ignored.

Note: For ThinkSystem operating system 8.50 and newer, Lenovo recommends using the default settings as provided.

The multipath.conf are available in the following locations:

- For SLES, */usr/share/doc/packages/multipath-tools/multipath.conf.synthetic*
- For RHEL, */usr/share/doc/device-mapper-multipath-0.4.9/multipath.conf*

Create partitions and filesystems

A new LUN has no partition or file system when the Linux host first discovers it. You must format the LUN before it can be used. Optionally, you can create a file system on the LUN.

Before you begin

The host must have discovered the LUN. See common tasks on storage array in chapter Creating a workload, Create volumes, and Defining a host in ThinkSystem System Manager, and Mapping a volume to a host.

In the /dev/mapper folder, you have run the ls command to see the available disks.

You can initialize the disk as a basic disk with a GUID partition table (GPT) or Master boot record (MBR).

Format the LUN with a file system such as ext4. Some applications do not require this step.

1. Retrieve the SCSI ID of the mapped disk by issuing the multipath -ll command.The SCSI ID is a 33-character string of hexadecimal digits, beginning with the number 3. If user-friendly names are enabled, Device Mapper reports disks as mpath instead of by a SCSI ID.

   ```
   # multipath -ll
   mpathd(360080e5000321bb8000092b1535f887a) dm-2 LENOVO ,DE_Series
   size=1.0T features='3 queue_if_no_path pg_init_retries 50' hwhandler='1 alua' wp=rw
   |-+- policy='service-time 0' prio=50 status=active
   | |- 16:0:4:4 sde    69:144 active ready running
   | `- 15:0:5:4 sdf    65:176 active ready running
   `-+- policy='service-time 0' prio=10 status=enabled
   |- 16:0:5:4 sdg      70:80   active ready running
   `- 15:0:1:4 sdh      66:0       active ready running.
   ```

2. Create a new partition according to the method appropriate for your Linux OS release. Typically, characters identifying the partition of a disk are appended to the SCSI ID (the number 1 or p3 for instance).

   ```
   # parted -a optimal -s -- /dev/mapper/360080e5000321bb8000092b1535f887a mklabel gpt mkpart
   primary                                                ext4 0% 100%
   ```


3. Create a file system on the partition.The method for creating a file system varies depending on the file system chosen.

   # mkfs.ext4 /dev/mapper/360080e5000321bb8000092b1535f887a1

4. Create a folder to mount the new partition.

   # mkdir /mnt/ext4

5. Mount the partition.

   # mount /dev/mapper/360080e5000321bb8000092b1535f887a1 /mnt/ext4

Verify storage access on the host

Before using the volume, you verify that the host can write data to the volume and read it back.

Before you begin

You must have initialized the volume and formatted it with a file system.

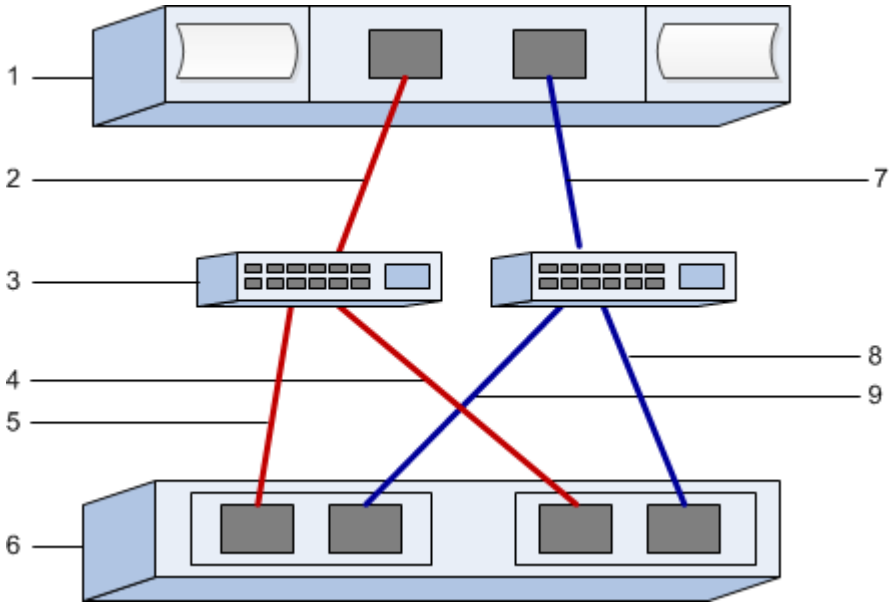
1. On the host, copy one or more files to the mount point of the disk
2. Copy the files back to a different folder on the original disk.
3. Run the diff command to compare the copied files to the originals.

Remove the file and folder that you copied.

FC worksheet

You can use this worksheet to record FC storage configuration information. You need this information to perform provisioning tasks.

The illustration shows a host connected to an DE Series storage array in two zones. One zone is indicated by the blue line; the other zone is indicated by the red line. Any single port has two paths to the storage (one to each controller).



Host indentifiers

Callout No.	Host (initiator) port connections	WWPN
1	Host	*not applicable*
2	Host port 0 to FC switch zone 0	
7	Host port 1 to FC switch zone 1	

Target identifiers

Callout No.	Host (initiator) port connections	WWPN
3	Switch	*not applicable*
6	Array controller (target)	*not applicable*
5	Controller A, port 1 to FC switch 1	
9	Controller A, port 2 to FC switch 2	
4	Controller B, port 1 to FC switch 1	
8	Controller B, port 2 to FC switch 2	

Mapping host

Mapping host name	
Host OS type	

Performing iSCSI-specific tasks

For the iSCSI protocol, you configure the switches and configure networking on the array side and the host side. Then you verify the IP network connections.

Configuring the switches - iSCSI

You configure the switches according to the vendor’s recommendations for iSCSI. These recommendations might include both configuration directives as well as code updates.

You must ensure the following:

- You have two separate networks for high availability. Make sure that you isolate your iSCSI traffic to separate network segments.
- You have enabled send and receive hardware flow control **end to end**.
- You have disabled priority flow control.
- If appropriate, you have enabled jumbo frames.

Note: Port channels/LACP is not supported on the controller's switch ports. Host-side LACP is not recommended; multipathing provides the same, and in some cases better, benefits.

Configuring networking - iSCSI

You can set up your iSCSI network in many ways, depending on your data storage requirements.

Consult your network administrator for tips on selecting the best configuration for your environment.

An effective strategy for configuring the iSCSI network with basic redundancy is to connect each host port and one port from each controller to separate switches and partition each set of host and controller ports on separate network segments using VLANs.

You must enable send and receive hardware flow control **end to end**. You must disable priority flow control.

If you are using jumbo frames within the IP SAN for performance reasons, make sure to configure the array, switches, and hosts to use jumbo frames. Consult your operating system and switch documentation for information on how to enable jumbo frames on the hosts and on the switches. To enable jumbo frames on the array, complete the steps in *Configuring array-side networking—iSCSI*.

Note: Many network switches must be configured above 9,000 bytes for IP overhead. Consult your switch documentation for more information.

Configuring array-side networking - iSCSI

You use the ThinkSystem System Manager GUI to configure iSCSI networking on the array side.

Before you begin

- You must know the IP address or domain name for one of the storage array controllers.
- You or your system administrator must have set up a password for the System Manager GUI, or you must have configured Role-Based Access Control (RBAC) or LDAP and a directory service for the appropriate security access to the storage array. See the *ThinkSystem System Manager online help* for more information about Access Management.

This task describes how to access the iSCSI port configuration from the Hardware page. You can also access the configuration from **System > Settings > Configure iSCSI ports**.

1. From your browser, enter the following URL: https://<DomainNameOrIPAddress>

   IP Address is the address for one of the storage array controllers.

   The first time ThinkSystem System Manager is opened on an array that has not been configured, the Set Administrator Password prompt appears. Role-based access management configures four local roles: **admin**, **support**, **security**, and **monitor**. The latter three roles have random passwords that cannot be guessed. After you set a password for the **admin** role you can change all of the passwords using the **admin** credentials. See *ThinkSystem System Manager online help* for more information on the four local user roles.

2. Enter the System Manager password for the **admin** role in the Set Administrator Password and Confirm Password fields, and then select the **Set Password** button.

   When you open System Manager and no pools, volumes groups, workloads, or notifications have been configured, the Setup wizard launches.

3. Close the Setup wizard.

   You will use the wizard later to complete additional setup tasks.

4. Select **Hardware**.

5. If the graphic shows the drives, click **Show back of shelf**.

   The graphic changes to show the controllers instead of the drives.

6. Click the controller with the iSCSI ports you want to configure.

   The controller's context menu appears.

7. Select **Configure iSCSI ports**.

   The Configure iSCSI Ports dialog box opens.

8. In the drop-down list, select the port you want to configure, and then click **Next**.

9. Select the configuration port settings, and then click **Next**.

To see all port settings, click the **Show more port settings** link on the right of the dialog box.

Port Setting	Description
Configured Ethernet port speed	Select the desired speed. The options that appear in the drop-down list depend on the maximum speed that your network can support (for example, 10 Gbps). **Note:** The optional iSCSI host interface cards in the DE6000H and DE6000F controllers do not auto-negotiate speeds. You must set the speed for each port to either 10 Gb or 25 Gb. All ports must be set to the same speed.
Enable IPv4 / Enable IPv6	Select one or both options to enable support for IPv4 and IPv6 networks.
TCP listening port (Available by clicking **Show more port settings**.)	If necessary, enter a new port number. The listening port is the TCP port number that the controller uses to listen for iSCSI logins from host iSCSI initiators. The default listening port is 3260. You must enter 3260 or a value between 49152 and 65535.
MTU size (Available by clicking **Show more port settings**.)	If necessary, enter a new size in bytes for the Maximum Transmission Unit (MTU). The default Maximum Transmission Unit (MTU) size is 1500 bytes per frame. You must enter a value between 1500 and 9000.
Enable ICMP PING responses	Select this option to enable the Internet Control Message Protocol (ICMP). The operating systems of networked computers use this protocol to send messages. These ICMP messages determine whether a host is reachable and how long it takes to get packets to and from that host.

If you selected **Enable IPv4**, a dialog box opens for selecting IPv4 settings after you click **Next**. If you selected **Enable IPv6**, a dialog box opens for selecting IPv6 settings after you click **Next**. If you selected both options, the dialog box for IPv4 settings opens first, and then after you click **Next**, the dialog box for IPv6 settings opens.

10. Configure the IPv4 and/or IPv6 settings, either automatically or manually. To see all port settings, click the **Show more settings** link on the right of the dialog box.

Port setting	Description
Automatically obtain configuration	Select this option to obtain the configuration automatically.
Manually specify static configuration	Select this option, and then enter a static address in the fields. For IPv4, include the network subnet mask and gateway. For IPv6, include the routable IP address and router IP address.
Enable VLAN support	**Important:** This option is only available in an iSCSI environment. Select this option to enable a VLAN and enter its ID. A VLAN is a logical network that behaves like it is physically separate from other

Port setting	Description
(Available by clicking **Show more settings**.)	physical and virtual local area networks (LANs) supported by the same switches, the same routers, or both.
Enable ethernet priority (Available by clicking **Show more settings**.)	**Important:** This option is only available in an iSCSI environment. Select this option to enable the parameter that determines the priority of accessing the network. Use the slider to select a priority between 1 and 7. In a shared local area network (LAN) environment, such as Ethernet, many stations might contend for access to the network. Access is on a first-come, first-served basis. Two stations might try to access the network at the same time, which causes both stations to back off and wait before trying again. This process is minimized for switched Ethernet, where only one station is connected to a switch port.

11. Click **Finish**.

12. Close System Manager.

Configuring host-side networking - iSCSI

You configure iSCSI networking on the host side by setting the number of node sessions per physical path, turning on the appropriate iSCSI services, configuring the network for the iSCSI ports, creating iSCSI face bindings, and establishing the iSCSI sessions between initiators and targets.

In most cases, you can use the inbox software-initiator for iSCSI CNA/NIC. You do not need to download the latest driver, firmware, and BIOS. Refer to the Interoperability Matrix document to determine code requirements.

Before you begin

- You have fully configured the switches that will be used to carry iSCSI storage traffic.
- You must have enabled send and receive hardware flow control **end to end** and disabled priority flow control.
- You have completed the array side iSCSI configuration.
- You must know the IP address of each port on the controller.

These instructions assume that two NIC ports will be used for iSCSI traffic.

1. Check the **node.session.nr_sessions** variable in the **/etc/iscsi/iscsid.conf** file to see the default number of sessions per physical path. If necessary, change the default number of sessions to one session.

   ```
   node.session.nr_sessions = 1
   ```

2. Change the **node.session.timeo.replacement_timeout** variable in the **/etc/iscsi/iscsid.conf** file to **20**, from a default value of 120.

   ```
   node.session.timeo.replacement_timeout=20
   ```


3. Make sure iscsid and (open-)iscsi services are on and enabled for boot.Red Hat Enterprise Linux 7 (RHEL 7) and Red Hat Enterprise Linux 7 and 8 (RHEL 7 and RHEL 8)

   # systemctl start iscsi

   # systemctl start iscsid

   # systemctl enable iscsi

   # systemctl enable iscsid

   SUSE Linux Enterprise Server 12 (SLES 12) and SUSE Linux Enterprise Server 12 and 15 (SLES 12 and SLES 15)

   # systemctl start iscsid.service # systemctl enable iscsid.service

   Optionally, you set node.startup = automatic in in /etc/iscsi/iscsid.conf before running any **iscsiadm** commands to have sessions persist after reboot:

4. Get the host IQN initiator name, which will be used to configure the host to an array.

   # cat /etc/iscsi/initiatorname.iscsi

5. Configure the network for iSCSI ports:
   Note: In addition to the public network port, iSCSI initiators should use two NICs or more on separate private segments or vLANs
   1. Determine the iSCSI port names using the # ifconfig -a command.
   2. Set the IP address for the iSCSI initiator ports. The initiator ports should be present on the same subnet as the iSCSI target ports.
      # vim /etc/sysconfig/network-scripts/ifcfg-<NIC port>

      Edit:
      BOOTPROTO=none
      ONBOOT=yes
      NM_CONTROLLED=no

      Add:
      IPADDR=192.168.xxx.xxx
      NETMASK=255.255.255.0

      Note: Be sure to set the address for both iSCSI initiator ports.
   3. Restart network services.

      # systemctl restart network

   4. Make sure the Linux server can ping all of the iSCSI target ports.

6. Configure the iSCSI interfaces by creating two iSCSI iface bindings.
   # iscsiadm –m iface –I iface0 –o new
   # iscsiadm –m iface –I iface0 –o update –n iface.net_ifacename –v <NIC port1>
   # iscsiadm –m iface –I iface1 –o new
   # iscsiadm –m iface –I iface1 –o update –n iface.net_ifacename –v <NIC port2>

   Note: To list the interfaces, use iscsiadm –m iface
7. Establish the iSCSI sessions between initiators and targets (four total).

1. Discover iSCSI targets. Save the IQN (it will be the same with each discovery) in the worksheet for the next step.
   # iscsiadm –m discovery –t sendtargets –p 192.168.0.1:3260 –I iface0 –P 1
   Note: The IQN looks like the following:
   iqn.2002-09.lenovo:de-series.600a098000af40fe000000005b565ef8
2. Create the connection between the iSCSI initiators and iSCSI targets, using ifaces.
   # iscsiadm –m node –T iqn.2002-09.lenovo:de-series.600a098000af40fe000000005b565ef8 –p 192.168.0.1:3260 –I iface0 -I
3. List the iSCSI sessions established on the host.
   # iscsiadm -m session

Verifying IP network connections—iSCSI

You verify Internet Protocol (IP) network connections by using ping tests to ensure the host and array are able to communicate.

1. On the host, run one of the following commands, depending on whether jumbo frames are enabled:

   - If jumbo frames are not enabled, run this command:
     ping -I <hostIP> <targetIP>
   - If jumbo frames are enabled, run the ping command with a payload size of 8,972 bytes. The IP and ICMP combined headers are 28 bytes, which when added to the payload, equals 9,000 bytes. The -s switch sets the packet size bit. The -d switch sets the debug option. These options allow jumbo frames of 9,000 bytes to be successfully transmitted between the iSCSI initiator and the target.
     ping -I <hostIP> -s 8972 -d <targetIP>

2. In this example, the iSCSI target IP address is 192.0.2.8.

   C:\> ping -I 192.0.2.100 -s 8972 -d 192.0.2.8

   Reply from 192.0.2.8: bytes=8972 time=2ms TTL=64
   Reply from 192.0.2.8: bytes=8972 time=2ms TTL=64
   Reply from 192.0.2.8: bytes=8972 time=2ms TTL=64
   Reply from 192.0.2.8: bytes=8972 time=2ms TTL=64

   Ping statistics for 192.0.2.8:

   Packets: Sent = 4, Received = 4, Lost = 0 (0% loss), Approximate round trip times in milli-seconds:

   Minimum = 2ms, Maximum = 2ms, Average = 2ms

Configure the multipath software

Multipath software provides a redundant path to the storage array in case one of the physical paths is disrupted. The multipath software presents the operating system with a single virtual device that represents the active physical paths to the storage. The multipath software also manages the failover process that updates the virtual device. You use the device mapper multipath (DM-MP) tool for Linux installations.

Before you begin

You have installed the required packages on your system.

- For Red Hat (RHEL) hosts, verify the packages are installed by running **rpm -q device-mapper- multipath**.
- For SLES hosts, verify the packages are installed by running **rpm -q multipath-tools**.

By default, DM-MP is disabled in RHEL and SLES. Complete the following steps to enable DM-MP components on the host.

If you have not already installed the operating system, use the media supplied by your operating system vendor.

1. If a multipath.conf file is not already created, run the **# touch /etc/multipath.conf** command.
2. Use the default multipath settings by leaving the multipath.conf file blank.
3. Start the multipath service.
   # systemctl start multipathd
4. Configure multipath for startup persistence.
   # chkconfig multipathd on
5. Save your kernel version by running the uname -r command.
   # uname -r
6. Do one of the following to enable the multipathd daemon on boot.

If you are using....	Do this...
RHEL 6.x systems:	**chkconfig multipathd on**
RHEL 7.x systems:	**systemctl enable multipathd**
SLES 12.x and 15.x systems:	**systemctl enable multipathd**

7. Rebuild the initramfs image or the initrd image under /boot directory:

If you are using....	Do this...
RHEL 6.x and 7.x systems:	**dracut --force --add multipath**
SLES 12.x and 15.x systems:	**dracut --force --add multipath**

8. Make sure that the newly created /boot/initrams-* image or /boot/initrd-* image is selected in the boot configuration file.For example, for grub it is /boot/grub/menu.lst and for grub2 it is /boot/ grub2/menu.cfg.
9. Use the "Create host manually" procedure in the online help to check whether the hosts are defined. Verify that each host type is either Linux DM-MP (Kernel 3.10 or later) if you enable the Automatic Load Balancing feature, or Linux DM-MP (Kernel 3.9 or earlier) if you disable the Automatic Load Balancing feature. If necessary, change the selected host type to the appropriate setting.
10. Reboot the host.

Setting up the multipath.conf file

The multipath.conf file is the configuration file for the multipath daemon, multipathd. The multipath.conf file overrides the built-in configuration table for multipathd. Any line in the file whose first non-white-space character is # is considered a comment line. Empty lines are ignored.

Note: For ThinkSystem operating system 8.50 and newer, Lenovo recommends using the default settings as provided.

The multipath.conf are available in the following locations:

- For SLES, */usr/share/doc/packages/multipath-tools/multipath.conf.synthetic*
- For RHEL, */usr/share/doc/device-mapper-multipath-0.4.9/multipath.conf*

Create partitions and filesystems

A new LUN has no partition or file system when the Linux host first discovers it. You must format the LUN before it can be used. Optionally, you can create a file system on the LUN.

Before you begin

The host must have discovered the LUN. See common tasks on storage array in chapter Creating a workload, Create volumes, and Defining a host in ThinkSystem System Manager, and Mapping a volume to a host.

In the /dev/mapper folder, you have run the ls command to see the available disks.

You can initialize the disk as a basic disk with a GUID partition table (GPT) or Master boot record (MBR).

Format the LUN with a file system such as ext4. Some applications do not require this step.

1. Retrieve the SCSI ID of the mapped disk by issuing the multipath -ll command.The SCSI ID is a 33-character string of hexadecimal digits, beginning with the number 3. If user-friendly names are enabled, Device Mapper reports disks as mpath instead of by a SCSI ID.
   ```
   # multipath -ll
   mpathd(360080e5000321bb8000092b1535f887a) dm-2 LENOVO ,DE_Series
   size=1.0T features='3 queue_if_no_path pg_init_retries 50' hwhandler='1 alua' wp=rw
   |-+- policy='service-time 0' prio=50 status=active
   | |- 16:0:4:4 sde    69:144 active ready running
   | `- 15:0:5:4 sdf    65:176 active ready running
   `-+- policy='service-time 0' prio=10 status=enabled
   |- 16:0:5:4 sdg      70:80   active ready running
   `- 15:0:1:4 sdh      66:0       active ready running.
   ```
2. Create a new partition according to the method appropriate for your Linux OS release. Typically, characters identifying the partition of a disk are appended to the SCSI ID (the number 1 or p3 for instance).
   ```
   # parted -a optimal -s -- /dev/mapper/360080e5000321bb8000092b1535f887a mklabel gpt mkpart
   primary ext4 0% 100%
   ```
3. Create a file system on the partition.The method for creating a file system varies depending on the file system chosen.
   ```
   # mkfs.ext4 /dev/mapper/360080e5000321bb8000092b1535f887a1
   ```
4. Create a folder to mount the new partition.
   ```
   # mkdir /mnt/ext4
   ```
5. Mount the partition.
   ```
   # mount /dev/mapper/360080e5000321bb8000092b1535f887a1 /mnt/ext4
   ```

Verify storage access on the host

Before using the volume, you verify that the host can write data to the volume and read it back.

Before you begin

You must have initialized the volume and formatted it with a file system.

4. On the host, copy one or more files to the mount point of the disk
5. Copy the files back to a different folder on the original disk.
6. Run the diff command to compare the copied files to the originals.

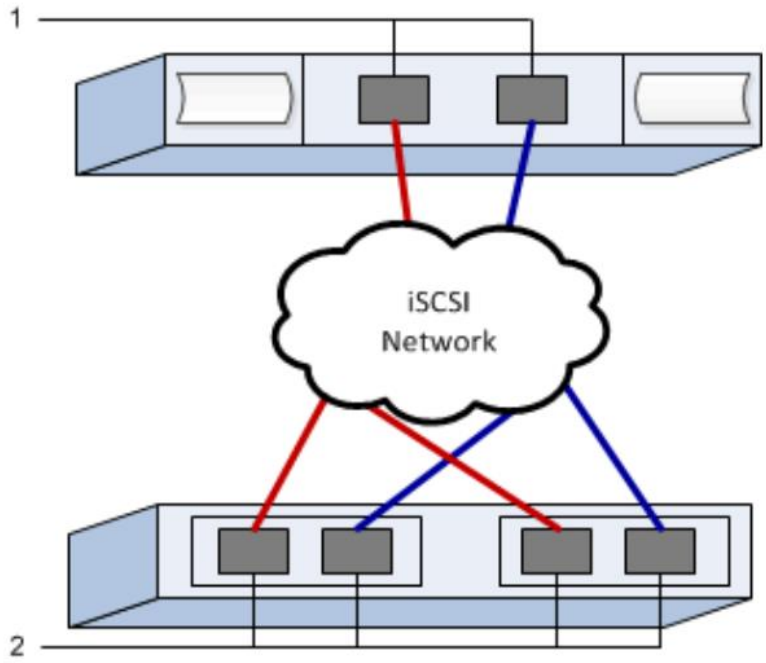
Remove the file and folder that you copied.

iSCSI worksheet

You can use this worksheet to record iSCSI storage configuration information. You need this information to perform provisioning tasks.

Recommended configuration

Recommended configurations consist of two initiator ports and four target ports with one or more VLANs.



Target IQN

Callout No.	Target port connection	IQN
2	Target port	

Mappings host name

Callout No.	Host information	Name and type
1	Mappings host name	
	Host OS type	

Performing SAS-specific tasks

Determining SAS host identifiers

For the SAS protocol, you find the SAS addresses using the HBA utility, then use the HBA BIOS to make the appropriate configuration settings.

Guidelines for HBA utilities:

- Most HBA vendors offer an HBA utility. Depending on your host operating system and CPU, use either the LSI-sas2flash(6G) or sas3flash(12G) utility.
- Host I/O ports might automatically register if the host context agent is installed.

1. Download the LSI-sas2flash(6G) or sas3flash(12G) utility from your HBA vendor's web site.
2. Install the utility.
3. Use the HBA BIOS to select the appropriate settings for your configuration.

Configure the multipath software

Multipath software provides a redundant path to the storage array in case one of the physical paths is disrupted. The multipath software presents the operating system with a single virtual device that represents the active physical paths to the storage. The multipath software also manages the failover process that updates the virtual device. You use the device mapper multipath (DM-MP) tool for Linux installations.

Before you begin

You have installed the required packages on your system.

- For Red Hat (RHEL) hosts, verify the packages are installed by running **rpm -q device-mapper- multipath**.
- For SLES hosts, verify the packages are installed by running **rpm -q multipath-tools**.

By default, DM-MP is disabled in RHEL and SLES. Complete the following steps to enable DM-MP components on the host.

If you have not already installed the operating system, use the media supplied by your operating system vendor.

1. If a multipath.conf file is not already created, run the **# touch /etc/multipath.conf** command.
2. Use the default multipath settings by leaving the multipath.conf file blank.
3. Start the multipath service.
   # systemctl start multipathd
4. Configure multipath for startup persistence.
   # chkconfig multipathd on
5. Save your kernel version by running the uname -r command.
   # uname -r

6. Do one of the following to enable the multipathd daemon on boot.

If you are using....	Do this...
RHEL 6.x systems:	**chkconfig multipathd on**
RHEL 7.x and 8.x systems:	**systemctl enable multipathd**
SLES 12.x and 15.x systems:	**systemctl enable multipathd**

7. Rebuild the initramfs image or the initrd image under /boot directory:

If you are using....	Do this...
RHEL 6.x, 7.x and 8.x systems:	**dracut --force --add multipath**
SLES 12.x and 15.x systems:	**dracut --force --add multipath**

8. Make sure that the newly created /boot/initrams-* image or /boot/initrd-* image is selected in the boot configuration file.For example, for grub it is /boot/grub/menu.lst and for grub2 it is /boot/ grub2/menu.cfg.
9. Use the "Create host manually" procedure in the online help to check whether the hosts are defined. Verify that each host type is either Linux DM-MP (Kernel 3.10 or later) if you enable the Automatic Load Balancing feature, or Linux DM-MP (Kernel 3.9 or earlier) if you disable the Automatic Load Balancing feature. If necessary, change the selected host type to the appropriate setting.
10. Reboot the host.

Setting up the multipath.conf file

The multipath.conf file is the configuration file for the multipath daemon, multipathd. The multipath.conf file overrides the built-in configuration table for multipathd. Any line in the file whose first non-white-space character is # is considered a comment line. Empty lines are ignored.

Note: For ThinkSystem operating system 8.50 and newer, Lenovo recommends using the default settings as provided.

The multipath.conf are available in the following locations:

- For SLES, */usr/share/doc/packages/multipath-tools/multipath.conf.synthetic*
- For RHEL, */usr/share/doc/device-mapper-multipath-0.4.9/multipath.conf*

Create partitions and filesystems

A new LUN has no partition or file system when the Linux host first discovers it. You must format the LUN before it can be used. Optionally, you can create a file system on the LUN.

Before you begin

The host must have discovered the LUN. See common tasks on storage array in chapter Creating a workload, Create volumes, and Defining a host in ThinkSystem System Manager, and Mapping a volume to a host.

In the /dev/mapper folder, you have run the ls command to see the available disks.

You can initialize the disk as a basic disk with a GUID partition table (GPT) or Master boot record (MBR).

Format the LUN with a file system such as ext4. Some applications do not require this step.

1. Retrieve the SCSI ID of the mapped disk by issuing the multipath -ll command.The SCSI ID is a 33-character string of hexadecimal digits, beginning with the number 3. If user-friendly names are enabled, Device Mapper reports disks as mpath instead of by a SCSI ID.
   ```
   # multipath -ll
   mpathd(360080e5000321bb8000092b1535f887a) dm-2 LENOVO ,DE_Series
   size=1.0T features='3 queue_if_no_path pg_init_retries 50' hwhandler='1 alua' wp=rw
   |-+- policy='service-time 0' prio=50 status=active
   | |- 16:0:4:4 sde    69:144 active ready running
   | `- 15:0:5:4 sdf    65:176 active ready running
   `-+- policy='service-time 0' prio=10 status=enabled
   |- 16:0:5:4 sdg      70:80   active ready running
   `- 15:0:1:4 sdh      66:0       active ready running.
   ```

2. Create a new partition according to the method appropriate for your Linux OS release. Typically, characters identifying the partition of a disk are appended to the SCSI ID (the number 1 or p3 for instance).
   ```
   # parted -a optimal -s -- /dev/mapper/360080e5000321bb8000092b1535f887a mklabel gpt mkpart
   primary ext4 0% 100%
   ```

3. Create a file system on the partition.The method for creating a file system varies depending on the file system chosen.
   ```
   # mkfs.ext4 /dev/mapper/360080e5000321bb8000092b1535f887a1
   ```

4. Create a folder to mount the new partition.
   ```
   # mkdir /mnt/ext4
   ```

5. Mount the partition.
   ```
   # mount /dev/mapper/360080e5000321bb8000092b1535f887a1 /mnt/ext4
   ```

Verify storage access on the host

Before using the volume, you verify that the host can write data to the volume and read it back.

Before you begin

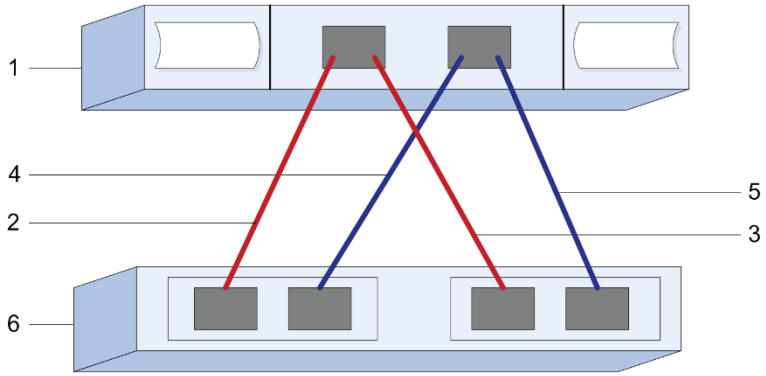
You must have initialized the volume and formatted it with a file system.

7. On the host, copy one or more files to the mount point of the disk
8. Copy the files back to a different folder on the original disk.
9. Run the diff command to compare the copied files to the originals.

Remove the file and folder that you copied.

SAS worksheet

You can use this worksheet to record SAS storage configuration information. You need this information to perform provisioning tasks.



Host Identifiers

Callout No.	Host (initiator) port connections	SAS address
1	Host	*not applicable*
2	Host (initiator) port 1 connected to Controller A, port 1	
3	Host (initiator) port 1 connected to Controller B, port 1	
4	Host (initiator) port 2 connected to Controller A, port 1	
5	Host (initiator) port 2 connected to Controller B, port 1	

Target Identifiers

Recommended configurations consist of two target ports.

Mappings Host

Mappings Host Name	
Host OS Type	

Performing NVMe over RoCE-specific tasks

You can use NVMe with the RDMA over Converged Ethernet (RoCE) network protocol.

Verify the Linux configuration is supported

To ensure reliable operation, you create an implementation plan and then use the Lenovo Interoperability Matrix to verify that the entire configuration is supported.

1. Go to Lenovo Storage Interoperation Center (LSIC).
2. Appropriate settings for your configuration are listed

NVMe over RoCE restrictions

Controller restrictions

- NVME over RoCE can be configured for the DE6000H or DE6000F 64GB controllers. The controllers must have 100GB host ports.

Switch restrictions

Attention: **RISK OF DATA LOSS**. You must enable Priority Flow Control or Global Pause Control on the switch to eliminate the risk of data loss in an NVMe over RoCE environment.

Host, host protocol, and host operating system restrictions

- The host must be running SUSE Linux Enterprise Server 12 SP5 or later system. See the Lenovo Storage Interoperation Center (LSIC) for a complete list of requirements.
- For a list of supported host channel adapters see the Lenovo Storage Interoperation Center (LSIC).
- In-band CLI management via 11.50.3 SMcli is not supported in NVMe-oF modes.

Storage and disaster recovery restrictions

- Asynchronous and synchronous mirroring are not supported.
- Thin provisioning (the creation of thin volumes) is not supported

Configure the switch

You configure the switches according to the vendor's recommendations for NVMe over RoCE. These recommendations might include both configuration directives as well as code updates.

Attention: **RISK OF DATA LOSS**. You must enable Priority Flow Control or Global Pause Control on the switch to eliminate the risk of data loss in an NVMe over RoCE environment.

Enable Ethernet pause frame flow control end to end as the best practice configuration.

Consult your network administrator for tips on selecting the best configuration for your environment.

Set up NVMe over RoCE on the host side

NVMe initiator configuration in an NVMe-RoCE environment includes installing and configuring the **rdma-core** and **nvme-cli** packages, configuring initiator IP addresses, and setting up the NVMe-oF layer on the host.

1. Install the rdma and nvme-cli packages:
   # zypper install rdma-core
   # zypper install nvme-cli.

2. Setup IPv4 IP addresses on the ethernet ports used to connect NVMe over RoCE. For each network interface, create a configuration script that contains the different variables for that interface.

   The variables used in this step are based on server hardware and the network environment. The variables include the IPADDR and GATEWAY. These are example instructions for the latest SUSE Linux Enterprise Server 12 service pack:

   Create the example file/etc/sysconfig/network/ifcfg-eth4 as follows:
   BOOTPROTO='static'
   BROADCAST=
   ETHTOOL_OPTIONS=
   IPADDR='192.168.1.87/24'
   GATEWAY='192.168.1.1'
   MTU=
   NAME='MT27800 Family [ConnectX-5]'
   NETWORK=
   REMOTE_IPADDR=
   STARTMODE='auto'

   Create the example file/etc/sysconfig/network/ifcfg-eth5 as follows:
   BOOTPROTO='static'
   BROADCAST=
   ETHTOOL_OPTIONS=
   IPADDR='192.168.2.87/24'
   GATEWAY='192.168.2.1'
   MTU=
   NAME='MT27800
   Family [ConnectX-5]'
   NETWORK=
   REMOTE_IPADDR=
   STARTMODE='auto'

3. Enable the network interfaces:
   # ifup eth4
   # ifup eth5.

4. Set up the NVMe-oF layer on the host

   Create the following file under /etc/modules-load.d/ to load the nvme-rdma kernel module and make sure the kernel module will always be on, even after a reboot:
   # cat /etc/modules-load.d/nvme-rdma.conf
   nvme-rdma

Configure storage array NVMe over RoCE connections

If your controller includes a connection for NVMe over RoCE (RDMA over Converged Ethernet), you can configure the NVMe port settings from the Hardware page or the System page in ThinkSystem System Manager.

Attention: **RISK OF DATA LOSS**. You must enable Priority Flow Control or Global Pause Control on the switch to eliminate the risk of data loss in an NVMe over RoCE environment.

- Your controller must include an NVMe over RoCE host port; otherwise, the NVMe over RoCE settings are not available in ThinkSystem System Manager.
- You must know the IP address of the host connection.

You can access the NVMe over RoCE configuration from the Hardware page or from **Settings** > **System**. This task describes how to configure the ports from the Hardware page.

Note: The NVMe over RoCE settings and functions appear only if your storage array's controller includes an NVMe over RoCE port.

1. Select **Hardware**.
2. Click the controller with the NVMe over RoCE port you want to configure.
   The controller's context menu appears.
3. Select **Configure NVMe over RoCE ports**.
   The Configure NVMe over RoCE ports dialog box opens.
4. In the drop-down list, select the port you want to configure, and then click **Next**.
5. Select the port configuration settings you want to use, and then click **Next**. To see all port settings, click the Show more port settings link on the right of the dialog box

Port Setting	Description
Configured ethernet port speed	Select the desired speed. The options that appear in the drop-down list depend on the maximum speed that your network can support (for example, 10 Gbps). Possible values include: • Auto-negotiate • 10 Gbps • 25 Gbps • 40 Gbps • 50 Gbps • 100 Gbps Note: The configured NVMe over RoCE port speed should match the speed capability of the SFP on the selected port. All ports must be set to the same speed.
Enable IPv4 and/or Enable IPv6	Select one or both options to enable support for IPv4 and IPv6 networks.
MTU size (Available by clicking Showmore port settings.)	If necessary, enter a new size in bytes for the maximum transmissionunit (MTU). The default MTU size is 1500 bytes per frame. You must enter a valuebetween 1500 and 4200.

If you selected Enable IPv4, a dialog box opens for selecting IPv4 settings after you click Next. If you selected Enable IPv6, a dialog box opens for selecting IPv6 settings after you click Next. If you selected both options, the dialog box for IPv4 settings opens first, and then after you click Next, the dialog box for IPv6 settings opens.

6. Configure the IPv4 and/or IPv6 settings, either automatically or manually. To see all port settings, click the Show more settings link on the right of the dialog box

Port setting	Description

Automatically obtain configuration from DHCPserver	Select this option to obtain the configuration automatically.
Manually specify staticconfiguration	Select this option, and then enter a static address in the fields. For IPv4, include the network subnet mask and gateway. For IPv6, include the routable IP addresses and router IP address. Note: If there is only one routable IP address, set the remaining address to 0:0:0:0:0:0:0:0.

7. Click **Finish**.

Discover and connect to the storage from the host

Before making definitions of each host in ThinkSystem System Manager, you must discover the target controller ports from the host, and then establish NVMe connections.

1. Discover available subsystems on the NVMe-oF target for all paths using the following command:
   nvme discover -t rdma -a target_ip_address

   In this command, *target_ip_address* is the IP address of the target port.

   Note: The nvme discover command discovers all controller ports in the subsystem, regardless ofhost access.

   ```
   # nvme discover -t rdma -a 192.168.1.77
   Discovery Log Number of Records 2, Generation counter 0
   =====Discovery Log Entry 0======
   trtype: rdma
   adrfam: ipv4
   subtype: nvme subsystem
   treq:          not specified
   portid: 0
   trsvcid: 4420
   subnqn: nqn.1992-08.com.netapp:5700.600a098000a527a7000000005ab3af94 traddr:
   192.168.1.77
   rdma_prtype: roce
   rdma_qptype: connected
   rdma_cms: rdma-cm
   rdma_pkey: 0x0000

   =====Discovery Log Entry 1======
   trtype: rdma
   adrfam: ipv4
   subtype: nvme subsystem
   treq:          not specified
   portid: 1
   trsvcid: 4420
   subnqn: nqn.1992-08.com.netapp:5700.600a098000a527a7000000005ab3af94 traddr:
   192.168.2.77
   rdma_prtype: roce
   rdma_qptype: connected
   ```

2. Repeat step 1 for any other connections

   Connect to the discovered subsystem on the first path using the command: *nvme connect -t rdma -n discovered_sub_nqn -a target_ip_address -Q queue_depth_setting -l controller_loss_timeout_period*.

3. Repeat step 3 to connect the discovered subsystem on the second path

   Important: Connections are not established for any discovered port inaccessible

by the host.

Important: If you specify a port number using this command, the connection fails. The default port is the only port set up for connections.

Important: The recommended queue depth setting is 1024. Override the default setting of 128 with 1024 using the -Q 1024 command line option, as shown in the following example.

Important: The recommended controller loss timeout period in seconds is 60 minutes (3600 seconds). Override the default setting of 600 seconds with 3600 seconds using the -l 3600 command line option, as shown in the following example.

```
# nvme connect -t rdma -a 192.168.1.77 -n nqn.1992-08.com.netapp:5700.
600a098000a527a7000000005ab3af94 -Q 1024 -l 3600
# nvme connect -t rdma -a 192.168.2.77 -n nqn.1992-08.com.netapp:5700.
600a098000a527a7000000005ab3af94 -Q 1024 -l 3600
```

4. Repeat step 1 for any other connections.
5. Auto connect after system reboot setup.
    1. Create file nvmf-autoconnect.service under /usr/lib/systemd/system, if the file doesn't exist.
    2. Fulfill following service content as below:

```
[Unit]
Description=Connect NVMe-oF subsystems automatically during boot
ConditionPathExists=/etc/nvme/discovery.conf
After=network.target Before=remote-fs-pre.target

[Service] Type=oneshot
ExecStart=/usr/sbin/nvme connect-all

[Install] WantedBy=default.target
```

Set up failover on the host

Multipath software provides a redundant path to the storage array in case one of the physical paths is disrupted. There are currently two methods of multipathing available for NVMe, and which you will be using is going to be dependent on which OS version you are running. For SLES 12 SP5 and later, device mapper multipath (DMMP) will be used.

Configuring the SLES 12 SP5 and later host to run failover

The SUSE Linux Enterprise Server hosts require additional configuration changes to run failover.

- You have installed the required packages on your system.
- For SLES 12 SP5 and later hosts, verify the packages are installed by running **rpm -q multipath-tools**

By default, DM-MP is disabled in RHEL and SLES. Complete the following steps to enable DM-MP components on the host.

1. Add the NVMe DE Series device entry to the devices section of the **/etc/multipath.conf** file, as shown in the following example:

```
devices {
  device {
    vendor "NVME"
    product "NetApp E-Series" path_grouping_policy group_by_prio failback immediate
    no_path_retry 30
```



```
  }
}
```

2. Configure multipathd to start at system boot

   # systemctl enable multipathd

3. Start multipathd if it is not currently running.

   # systemctl start multipathd

4. if it is not currently running.

5. Verify the status of multipathd to make sure it is active and running:

   # systemctl status multipathd

Accessing NVMe Volumes

You can configure the I/O directed to the device target based on your Linux version.

Before you begin

The host must have discovered the namespace. See common tasks on storage array in chapter Creating a workload, Create volumes, and Defining a host in ThinkSystem System Manager, and Mapping a volume to a host.

Accessing NVMe volumes for virtual device targets (DM-MP devices)

For SLES 12, I/O is directed to virtual device targets by the Linux host. DM-MP manages the physical paths underlying these virtual targets. Make sure you are running I/O only to the virtual devices created by DM-MP and not to the physical device paths. If you are running I/O to the physical paths, DM-MP cannot manage a failover event and the I/O fails.

You can access these block devices through the dm device or the symlink in /dev/mapper, for example:

```
/dev/dm-1
/dev/mapper/eui.00001bc7593b7f5f00a0980000af4462
```

Example

The following example output from the nvme list command shows the host node name and its correlation with the namespace ID.

```
NODE    SN         MODEL  NAMESPACE
/dev/nvme1n1 021648023072 Lenovo DE-Series 10
/dev/nvme1n2 021648023072 Lenovo DE-Series 11
/dev/nvme1n3 021648023072 Lenovo DE-Series 12
/dev/nvme1n4 021648023072 Lenovo DE-Series 13
/dev/nvme2n1 021648023151 Lenovo DE-Series 10
/dev/nvme2n2 021648023151 Lenovo DE-Series 11
/dev/nvme2n3 021648023151 Lenovo DE-Series 12
/dev/nvme2n4 021648023151 Lenovo DE-Series 13
```


Column	Description
Node	The node name includes two parts: • The notation nvme1 represents controller A and nvme2 represents controller B. • The notation n1, n2, and so on represent the namespace identifier from the host perspective. These identifiers are repeated in the table, once for controller A and once forcontroller B.
Namespace	The Namespace column lists the namespace ID (NSID), which is the identifier from the storagearray perspective.

In the following multipath -ll output, the optimized paths are shown with a prio value of 50, while the non- optimized paths are shown with a prio value of 10.

The Linux operating system routes I/O to the path group that is shown as status=active, while the path groups listed as status=enabled are available for failover.

```
eui.00001bc7593b7f500a0980000af4462 dm-0 NVME,Lenovo DE-Series size=15G features='1
queue_if_no_path' hwhandler='0' wp=rw
|-+- policy='service-time 0' prio=50 status=active
| `- #:#:#:# nvme1n1 259:5 active ready running
`-+- policy='service-time 0' prio=10 status=enabled
`- #:#:#:# nvme2n1 259:9 active ready running

eui.00001bc7593b7f5f00a0980000af4462 dm-0 NVME,Lenovo DE-Series size=15G features='1
queue_if_no_path' hwhandler='0' wp=rw
|-+- policy='service-time 0' prio=0 status=enabled
| `- #:#:#:# nvme1n1 259:5 failed faulty running
`-+- policy='service-time 0' prio=10 status=active
`- #:#:#:# nvme2n1 259:9 active ready running
```

Line item	Description
policy='service-time0' prio=50 status= active	This line and the following line show that nvme1n1, which is the namespace with an NSID of10, is optimized on the path with a prio value of 50 and a status value of active. This namespace is owned by controller A.
policy='service-time0' prio=10 status= enabled	This line shows the failover path for namespace 10, with a prio value of 10 and a status valueof enabled. I/O is not being directed to the namespace on this path at the moment. This namespace is owned by controller B.
policy='service-time0' prio=0 status= enabled	This example shows multipath -ll output from a different point in time, while controller A is rebooting. The path to namespace 10 is shown as failed faulty running with a prio value of 0and a status value of enabled.
policy='service-time0' prio=10 status= active	Note that the active path refers to nvme2, so the I/O is being directed on this path to controllerB.

Create filesystems

You create a file system on the namespace or native nvme device and mount the filesystem.

Create filesystems (SLES 12)

For SLES 12, you create a file system on the namespace and mount the filesystem.

1. Run the multipath -ll command to get a list of /dev/mapper/dm devices.

   ```
   # multipath -ll
   ```

2. The result

of this command shows two devices, dm-19 and dm-16

```
eui.00001ffe5a94ff8500a0980000af4444 dm-19 NVME,Lenovo DE-Series size=10G
features='1 queue_if_no_path' hwhandler='0' wp=rw
|-+- policy='service-time 0' prio=8 status=active
| |- #:#:#:# nvme0n19 259:19 active ready running
| `- #:#:#:# nvme1n19 259:115 active ready running
`-+- policy='service-time 0' prio=2 status=enabled
|- #:#:#:# nvme2n19 259:51 active ready running
`- #:#:#:# nvme3n19 259:83 active ready running
eui.00001fd25a94fef000a0980000af4444 dm-16 NVME,Lenovo DE-Series size=16G
features='1 queue_if_no_path' hwhandler='0' wp=rw
|-+- policy='service-time 0' prio=8 status=active
| |- #:#:#:# nvme0n16 259:16 active ready running
| `- #:#:#:# nvme1n16 259:112 active ready running
`-+- policy='service-time 0' prio=2 status=enabled
|- #:#:#:# nvme2n16 259:48 active ready running
`- #:#:#:# nvme3n16 259:80 active ready running
```

3. Create a file system on the partition for each /dev/mapper/dm device.The method for creating a file system varies depending on the file system chosen. In this example, we are creating an ext4 file system.

   ```
   # mkfs.ext4 /dev/mapper/dm-19 mke2fs 1.42.11 (09-Jul-2014)
   Creating filesystem with 2620928 4k blocks and 655360 inodes Filesystem UUID: 97f987e9-
   47b8-47f7-b434-bf3ebbe826d0 Superblock backups stored on blocks:
   32768, 98304, 163840, 229376, 294912, 819200, 884736, 1605632
   Allocating group tables: done Writing inode tables: done
   Creating journal (32768 blocks): done
   Writing superblocks and filesystem accounting information: done g
   ```

4. Create a folder to mount the new device.

   ```
   # mkdir /mnt/ext4
   ```

5. Mount the device.

   ```
   # mount /dev/mapper/eui.00001ffe5a94ff8500a0980000af4444 /mnt/ext4
   ```

Verify storage access on the host

Before using the namespace, you verify that the host can write data to the namespace and read it back.

1. On the host, copy one or more files to the mount point of the disk
2. Copy the files back to a different folder on the original disk.
3. Run the **diff** command to compare the copied files to the originals.

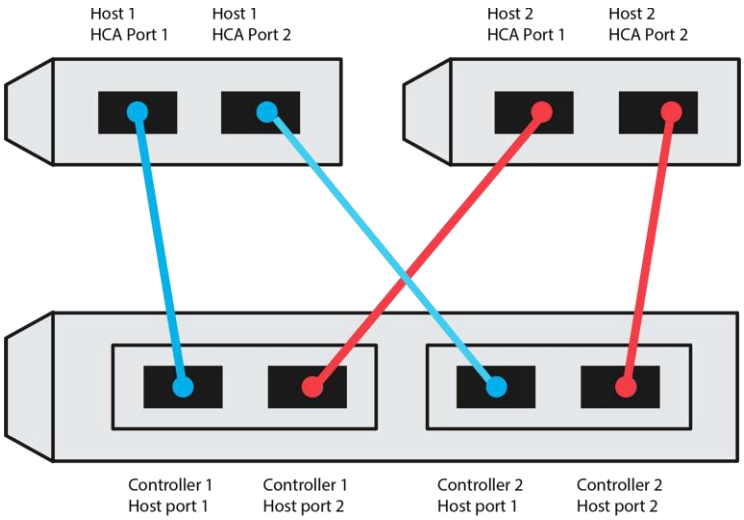
Remove the file and folder that you copied.

NVMe over RoCE worksheet for Linux

You can use this worksheet to record NVMe over RoCE storage configuration information. You need this information to perform provisioning tasks.

Direct connect topology

In a direct connect topology, one or more hosts are directly connected to the subsystem. In the ThinkSystem SAN OS 11.60.2 release, we support a single connection from each host to a subsystem controller, as shown below. In this configuration, one HCA (host channel adapter) port from each host should be on the same subnet as the DE Series controller port it is connected to, but on a different subnet from the other HCA port.

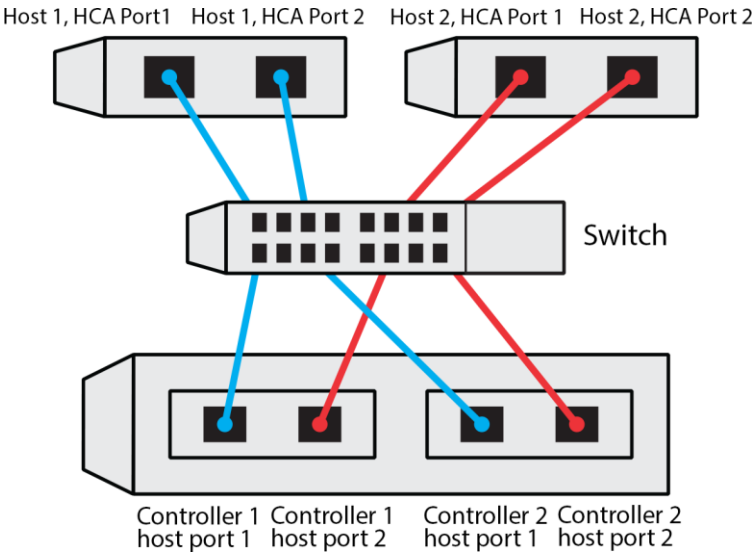


An example configuration that satisfies the requirements consists of four network subnets as follows:

- Subnet 1: Host 1 HCA Port 1 and Controller 1 Host port 1.
- Subnet 2: Host 1 HCA Port 2 and Controller 2 Host port 1
- Subnet 3: Host 2 HCA Port 1 and Controller 1 Host port 2
- Subnet 4: Host 2 HCA Port 2 and Controller 2 Host port 2

Switch connect topology

In a fabric topology, one or more switches are used. For a list of supported switches, go to the Lenovo Storage Interoperation Center (LSIC) and look for proper configuration.



NVMe over RoCE: Host Identifiers

Locate and document the initiator NQN from each host.

Host port connections	Initiator NQN
Host (initiator) 1	Host
Host (initiator) 2	

NVMe over RoCE: Target NQN

Recommended configurations consist of two target ports.

Array Name	Target NQN
Array controller (target)	

Mappings Host

Mappings Host Name	
Host OS Type	

Performing NVMe over Fibre Channel tasks

You can use NVMe with the Fibre Channel protocol.

Verify the Linux configuration is supported

To ensure reliable operation, you create an implementation plan and then use the Lenovo Interoperability Matrix to verify that the entire configuration is supported.

3. Go to Lenovo Storage Interoperation Center (LSIC).
4. Appropriate settings for your configuration are listed

NVMe over Fibre Channel restrictions

Controller restrictions

- NVME over Fibre Channel can be configured for the DE6000H or DE6000F 64GB controllers. The controllers must have 100GB host ports.

Switch restrictions

Attention: **RISK OF DATA LOSS**. You must enable Priority Flow Control or Global Pause Control on the switch to eliminate the risk of data loss in an NVMe over RoCE environment.

Host, host protocol, and host operating system restrictions

- The host must be running SUSE Linux Enterprise Server 12 SP5 or later system. See the Lenovo Storage Interoperation Center (LSIC) for a complete list of requirements.
- For a list of supported host channel adapters see the Lenovo Storage Interoperation Center (LSIC).
- In-band CLI management via 11.50.3 SMcli is not supported in NVMe-oF modes.

Storage and disaster recovery restrictions

- Asynchronous and synchronous mirroring are not supported.
- Thin provisioning (the creation of thin volumes) is not supported

Configure the switch

Configuring (zoning) the Fibre Channel (FC) switches enables the hosts to connect to the storage array and limits the number of paths. You zone the switches using the management interface for the switches.

Before you begin

- You must have administrator credentials for the switches.
- You must have used your HBA utility to discover the WWPN of each host initiator port and of each controller target port connected to the switch.

For details about zoning your switches, see the switch vendor's documentation.

You must zone by WWPN, not by physical port. Each initiator port must be in a separate zone with all of its corresponding target ports.

1. Log in to the FC switch administration program, and then select the zoning configuration option.

2. Create a new zone that includes the first host initiator port and that also includes all of the target ports that connect to the same FC switch as the initiator.
3. Create additional zones for each FC host initiator port in the switch.
4. Save the zones, and then activate the new zoning configuration.

Set up NVMe over Fibre Channel on the host side

NVMe initiator configuration in a Fibre Channel environment includes installing and configuring the nvme-cli package, and enabling the NVMe/FC initiator on the host.

These are the instructions for SUSE Linux Enterprise Server 15 SP1 and 32Gb FC HBAs.

1. Install the **nvme-cli** package:

   For SLES15 SP1:

   ```
   # zypper install nvme-cli
   ```

2. Enable and start the **nvmefc-boot-connections** service.

   ```
   # systemctl enable nvmefc-boot-connections.service
   # systemctl start nvmefc-boot-connections.service
   ```

3. Set **lpfc_enable_fc4_type** to **3** to enable SLES15 SP1 as an NVMe/FC initiator.

   ```
   # cat /etc/modprobe.d/lpfc.conf
   options lpfc lpfc_enable_fc4_type=3
   ```

4. Re-build the initrd to get the Emulex change and the boot parameter change.

   ```
   # dracut –force
   ```

5. Reboot the host to reconfigure the lpfc driver.

   ```
   # reboot
   ```

   The host is rebooted and the NVMe/FC initiator is enabled on the host.

   Note: After completing the host side setup, configuration of the NVMe over Fibre Channel ports occur automatically.

Display the volumes visible to the host

Before you begin

The host must have discovered the namespace. See common tasks on storage array in chapter Creating a workload, Create volumes, and Defining a host in ThinkSystem System Manager, and Mapping a volume to a host.

The **SMdevices** tool, part of the nvme-cli package, allows you to view the volumes currently visible on the host. This tool is an alternative to the nvme list command.

1. To view information about each NVMe path to a DE Series volume, use the nvme netapp smdevices [-o <format>] command. The output <format> can be normal (the default if -o is not used), column, or json.

   ```
   # nvme netapp smdevices
   /dev/nvme1n1, Array Name ICTM0706SYS04, Volume Name NVMe2, NSID 1, Volume ID
   000015bd5903df4a00a0980000af4462, Controller A, Access State unknown, 2.15GB
   ```


/dev/nvme1n2, Array Name ICTM0706SYS04, Volume Name NVMe3, NSID 2, Volume ID 000015c05903e24000a0980000af4462, Controller A, Access State unknown, 2.15GB
/dev/nvme1n3, Array Name ICTM0706SYS04, Volume Name NVMe4, NSID 4, Volume ID 00001bb0593a46f400a0980000af4462, Controller A, Access State unknown, 2.15GB
/dev/nvme1n4, Array Name ICTM0706SYS04, Volume Name NVMe6, NSID 6, Volume ID 00001696593b424b00a0980000af4112, Controller A, Access State unknown, 2.15GB
/dev/nvme2n1, Array Name ICTM0706SYS04, Volume Name NVMe2, NSID 1, Volume ID 000015bd5903df4a00a0980000af4462, Controller B, Access State unknown, 2.15GB
/dev/nvme2n2, Array Name ICTM0706SYS04, Volume Name NVMe3, NSID 2, Volume ID 000015c05903e24000a0980000af4462, Controller B, Access State unknown, 2.15GB
/dev/nvme2n3, Array Name ICTM0706SYS04, Volume Name NVMe4, NSID 4, Volume ID 00001bb0593a46f400a0980000af4462, Controller B, Access State unknown, 2.15GB
/dev/nvme2n4, Array Name ICTM0706SYS04, Volume Name NVMe6, NSID 6, Volume ID 00001696593b424b00a0980000af4112, Controller B, Access State unknown, 2.15GB

Set up failover on the host

Accessing NVMe volumes for physical NVMe device targets (SLES 15)

For SLES 15 SP1, I/O is directed to the physical NVMe device targets by the Linux host. A native NVMe multipathing solution manages the physical paths underlying the single apparent physical device displayed by the host.

Note: It is best practice to use the links in /dev/disk/by-id/ rather than /dev/nvme0n1, for example: # ls /dev/disk/by-id/ -l lrwxrwxrwx 1 root root 13 Oct 18 15:14 nvme-eui.0000320f5cad32cf00a0980000af4112 -> ../../nvme0n1

Physical NVMe devices are I/O targets

Run I/O to the physical nvme device path. There should only be one of these devices present for each namespace using the following format:

```
/dev/nvme[subsys#]n[id#]
```

All paths are virtualized using the native multipathing solution underneath this device. You can view your paths by running:

```
# nvme list-subsys
nvme-subsys0 - NQN=nqn.1992-08.com.netapp:5700.600a098000d709d6000000005e27796e
\
+- nvme0 fc traddr=nn-0x200200a098d709d6:pn-0x204200a098d709d6 host_traddr=\ nn-
0x200000109b211680:pn-0x100000109b211680 live
+- nvme1 fc traddr=nn-0x200200a098d709d6:pn-0x204300a098d709d6 host_traddr=\ nn-
0x200000109b21167f:pn-0x100000109b21167f live
```

If you specify a namespace device when using the nvme list-subsys command, it provides additional information about the paths to that namespace:

```
# nvme list-subsys /dev/nvme0n1
nvme-subsys0 - NQN=nqn.1992-08.com.netapp:5700.600a098000d709d6000000005e27796e
\
+- nvme0 fc traddr=nn-0x200200a098d709d6:pn-0x204200a098d709d6 host_traddr=\ nn-
0x200000109b211680:pn-0x100000109b211680 live
+- nvme1 fc traddr=nn-0x200200a098d709d6:pn-0x204300a098d709d6 host_traddr=\ nn-
0x200000109b21167f:pn-0x100000109b21167f live
```


There are also hooks into the multipath commands to allow you to view your path information for native failover through them as well:

```
# multipath -ll
eui.000007e15e903fac00a0980000d663f2 [nvme]:nvme0n1 NVMe,Lenovo DE-Series,98620002
size=207618048 features='n/a' hwhandler='ANA' wp=rw
|-+- policy='n/a' prio=n/a status=n/a\
| `- 0:10:1          nvme0c10n1      0:0 n/a   n/a       live
`-+- policy='n/a' prio=n/a status=n/a\
`- 0:32778:1 nvme0c32778n1 0:0 n/a          n/a       live
```

Create filesystems

Create filesystems (SLES 15)

For SLES 15 SP1, you create a filesystem on the native nvme device and mount the filesystem.

1. Run the **multipath -ll** command to get a list of /dev/nvme devices.The result of this command shows device nvme0n1:

   ```
   # multipath -ll
   eui.000007e15e903fac00a0980000d663f2 [nvme]:nvme0n1 NVMe,Lenovo DE-Series,98620002
   size=207618048 features='n/a' hwhandler='ANA' wp=rw
   |-+- policy='n/a' prio=n/a status=n/a\
   | `- 0:10:1    nvme0c10n1      0:0 n/a   n/a       live
   `-+- policy='n/a' prio=n/a status=n/a\
   `- 0:32778:1 nvme0c32778n1 0:0 n/a    n/a       live
   ```

2. Create a file system on the partition for each /dev/nvme0n# device. The method for creating a file system varies depending on the file system chosen. This example shows creating an ext4 file system.

   ```
   # mkfs.ext4 /dev/disk/by-id/nvme-eui.000082dd5c05d39300a0980000a52225 mke2fs 1.42.11
   (22-Oct-2019) Creating filesystem with 2620928 4k blocks and 655360 inodes Filesystem UUID:
   97f987e9-47b8-47f7-b434-bf3ebbe826d0 Superblock backups stored on blocks: 32768, 98304,
   163840, 229376, 294912, 819200, 884736, 1605632
   Allocating group tables: done
   Writing inode tables: done
   Creating journal (32768 blocks): done
   Writing superblocks and filesystem accounting information: done
   ```

3. Create a folder to mount the new device.# mkdir /mnt/ext4

4. Mount the device

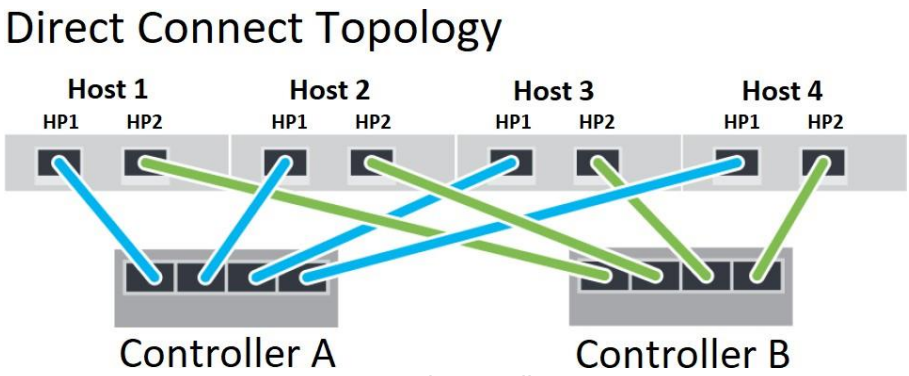
   ```
   # mount /dev/disk/by-id/nvme-eui.000082dd5c05d39300a0980000a52225 /mnt/ext4
   ```

NVMe over Fibre Channel worksheet for Linux

You can use this worksheet to record NVMe over Fibre Channel storage configuration information. You need this information to perform provisioning tasks.

Direct connect topology

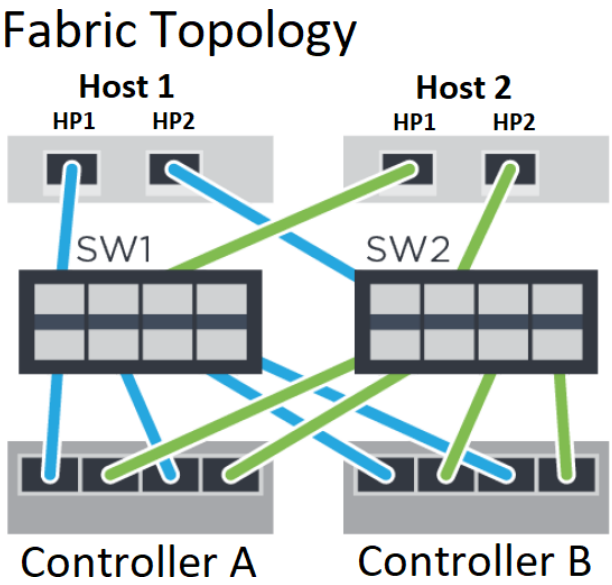
In a direct connect topology, one or more hosts are directly connected to the controller.



- Host 1 HBA Port 1 and Controller A Host port 1
- Host 1 HBA Port 2 and Controller B Host port 1
- Host 2 HBA Port 1 and Controller A Host port 2
- Host 2 HBA Port 2 and Controller B Host port 2
- Host 3 HBA Port 1 and Controller A Host port 3
- Host 3 HBA Port 2 and Controller B Host port 3
- Host 4 HBA Port 1 and Controller A Host port 4
- Host 4 HBA Port 2 and Controller B Host port 4

Switch connect topology

In a fabric topology, one or more switches are used. For a list of supported switches, go to the Lenovo Storage Interoperation Center (LSIC) and look for proper configuration.



NVMe over Fibre Channel: Host identifiers

Locate and document the initiator NQN from each host.

Host port connections	Software NQN
Host (initiator) 1	Host
Host (initiator) 2	

NVMe over Fibre Channel: Target NQN

Recommended configurations consist of two target ports.

Array Name	Target NQN
Array controller (target)	

Mappings Host

Mappings Host Name	
Host OS Type	

Creating a workload

You create storage by first creating a workload for a specific application type. Next, you add storage capacity to the workload by creating volumes with similar underlying volume characteristics.

Create workloads

You can create workloads for any type of application.

About this task

A workload is a storage object that supports an application. You can define one or more workloads, or instances, per application. For some applications, the system configures the workload to contain volumes with similar underlying volume characteristics. These volume characteristics are optimized based on the type of application the workload supports.

Keep these guidelines in mind:

- *When using an application-specific workload* , the system recommends an optimized volume configuration to minimize contention between application workload I/O and other traffic from your application instance. You can review the recommended volume configuration, and then edit, add, or delete the system-recommended volumes and characteristics using the **Add/Edit Volumes** dialog box.
- *When using other application types* , you manually specify the volume configuration using the **Add/Edit Volumes** dialog box.

1. Select **Storage** > **Volumes** .
2. Select **Create** > **Workload** .

   The **Create Application Workload** dialog box appears.
3. Use the drop-down list to select the type of application that you want to create the workload for and then type a workload name.
4. Click **Create** .

You are ready to add storage capacity to the workload you created. Use the **Create Volume** option to create one or more volumes for an application, and to allocate specific amounts of capacity to each volume.

Create volumes

You create volumes to add storage capacity to an application-specific workload, and to make the created volumes visible to a specific host or host cluster. In addition, the volume creation sequence provides options to allocate specific amounts of capacity to each volume you want to create.

Most application types default to a user-defined volume configuration. Some application types have a smart configuration applied at volume creation. For example, if you are creating volumes for Microsoft Exchange application, you are asked how many mailboxes you need, what your average mailbox capacity requirements are, and how many copies of the database you want. System Manager uses this information to create an optimal volume configuration for you, which can be edited as needed.

The process to create a volume is a multi-step procedure.

Note: If you want to mirror a volume, first create the volumes that you want to mirror, and then use the **Storage** > **Volumes** > **Copy Services** > **Mirror a volume asynchronously** option.

1. Step 1: Select host for a volume
   You create volumes to add storage capacity to an application-specific workload, and to make the created volumes visible to a specific host or host cluster. In addition, the volume creation sequence provides options to allocate specific amounts of capacity to each volume you want to create.

2. Step 2: Select a workload for a volume
   Select a workload to customize the storage array configuration for a specific application, such as Microsoft SQL Server, Microsoft Exchange, Video Surveillance applications, or VMware. You can select "Other application" if the application you intend to use on this storage array is not listed.

3. Step 3: Add or edit volumes
   System Manager may suggest a volume configuration based on the application or workload you selected. This volume configuration is optimized based on the type of application the workload supports. You can accept the recommended volume configuration, or you can edit it as needed. If you selected one of the "Other" applications, you must manually specify the volumes and characteristics you want to create.

4. Step 4: Review volume configuration
   Review a summary of the volumes you intend to create and make any necessary changes.

Step 1: Select host for a volume

You create volumes to add storage capacity to an application-specific workload, and to make the created volumes visible to a specific host or host cluster. In addition, the volume creation sequence provides options to allocate specific amounts of capacity to each volume you want to create.

- Valid hosts or host clusters exist under the **Hosts** tile.
- Host port identifiers have been defined for the host.
- Before creating a DA-enabled volume, the host connection you are planning to use must support DA. If any of the host connections on the controllers in your storage array do not support DA, the associated hosts cannot access data on DA-enabled volumes. ThinkSystem DE Series storage only supports DA between the controller and the drives.

Keep these guidelines in mind when you assign volumes:

- A host's operating system can have specific limits on how many volumes the host can access. Keep this limitation in mind when you create volumes for use by a particular host.
- You can define one assignment for each volume in the storage array.
- Assigned volumes are shared between controllers in the storage array.
- The same logical unit number (LUN) cannot be used twice by a host or a host cluster to access a volume. You must use a unique LUN.
- If you want to speed the process for creating volumes, you can skip the host assignment step so that newly created volumes are initialized offline.

Note: Assigning a volume to a host will fail if you try to assign a volume to a host cluster that conflicts with an established assignment for a host in the host clusters.

1. Select **Storage** > **Volumes**.
2. Select **Create** > **Volume**.

   The **Create Volumes** dialog box appears.
3. From the drop-down list, select a specific host or host cluster to which you want to assign volumes, or choose to assign the host or host cluster at a later time.
4. To continue the volume creation sequence for the selected host or host cluster, click **Next**, and go to Step 2: Select a workload for a volume.

   The **Select Workload** dialog box appears

Step 2: Select a workload for a volume

Select a workload to customize the storage array configuration for a specific application, such as Microsoft SQL Server, Microsoft Exchange, Video Surveillance applications, or VMware. You can select "Other application" if the application you intend to use on this storage array is not listed.

This task describes how to create volumes for an existing workload.

- *When you are creating volumes using an application-specific workload*, the system may recommend an optimized volume configuration to minimize contention between application workload I/O and other traffic from your application instance. You can review the recommended volume configuration and edit, add, or delete the system-recommended volumes and characteristics using the **Add/Edit Volumes** dialog box.
- *When you are creating volumes using "Other"* applications (or applications without specific volume creation support), you manually specify the volume configuration using the **Add/Edit Volumes** dialog box.

1. Do one of the following:
   - Select the **Create volumes for an existing workload** option to create volumes for an existing workload.
   - Select the **Create a new workload** option to define a new workload for a supported application or for "Other" applications.

- From the drop-down list, select the name of the application you want to create the new workload for.

  Select one of the "Other" entries if the application you intend to use on this storage array is not listed.

- Enter a name for the workload you want to create.

2. Click **Next**.
3. If your workload is associated with a supported application type, enter the information requested; otherwise, go to Step 3: Add or edit volumes.

Step 3: Add or edit volumes

System Manager may suggest a volume configuration based on the application or workload you selected. This volume configuration is optimized based on the type of application the workload supports. You can accept the recommended volume configuration or you can edit it as needed. If you selected one of the "Other" applications, you must manually specify the volumes and characteristics you want to create.

Before you begin

- The pools or volume groups must have sufficient free capacity.
- The maximum number of volumes allowed in a volume group is 256.
- The maximum number of volumes allowed in a pool depends on the storage system model:
  - 2,048 volumes (DE6000H, DE6000F series)
  - 512 volumes (DE2000H, DE4000H, DE4000F series)
- To create a Data Assurance (DA)-enabled volume, the host connection you are planning to use must support DA.

  Selecting a DA capable pool or volume group

  If you want to create a DA-enabled volume, select a pool or volume group that is DA capable (look for **Yes** next to "DA" in the pool and volume group candidates table).

  DA capabilities are presented at the pool and volume group level in System Manager. DA protection checks for and corrects errors that might occur as data is transferred through the controllers down to the drives. Selecting a DA-capable pool or volume group for the new volume ensures that any errors are detected and corrected.

  If any of the host connections on the controllers in your storage array do not support DA, the associated hosts cannot access data on DA-enabled volumes. ThinkSystem DE Series only supports DA between the controller and the drives.
- To create a secure-enabled volume, a security key must be created for the storage array.

  Selecting a secure-capable pool or volume group

  If you want to create a secure-enabled volume, select a pool or volume group that is secure capable (look for **Yes** next to "Secure-capable" in the pool and volume group candidates table).

Drive security capabilities are presented at the pool and volume group level in System Manager. Secure-capable drives prevent unauthorized access to the data on a drive that is physically removed from the storage array. A secure-enabled drive encrypts data during writes and decrypts data during reads using a unique *encryption key*.

A pool or volume group can contain both secure-capable and non-secure-capable drives, but all drives must be secure-capable to use their encryption capabilities.

About this task

You create volumes from pools or volume groups. The **Add/Edit Volumes** dialog box shows all eligible pools and volume groups on the storage array. For each eligible pool and volume group, the number of drives available and the total free capacity appears.

For some application-specific workloads, each eligible pool or volume group shows the proposed capacity based on the suggested volume configuration and shows the remaining free capacity in GiB. For other workloads, the proposed capacity appears as you add volumes to a pool or volume group and specify the reported capacity.

1. Choose one of these actions based on whether you selected Other or an application-specific workload:

   - **Other** – Click **Add new volume** in each pool or volume group that you want to use to create one or more volumes.

     Table 1. Field Details

Field	Description
Volume Name	A volume is assigned a default name by System Manager during the volume creation sequence. You can either accept the default name or provide a more descriptive one indicating the type of data stored in the volume.
Reported Capacity	Define the capacity of the new volume and the capacity units to use (MiB, GiB, or TiB). For **Thick volumes**, the minimum capacity is 1 MiB, and the maximum capacity is determined by the number and capacity of the drives in the pool or volume group. Keep in mind that storage capacity is also required for copy services (snapshot images, snapshot volumes, volume copies, and remote mirrors); therefore, do not allocate all of the capacity to standard volumes. Capacity in a pool is allocated in 4-GiB increments. Any capacity that is not a multiple of 4 GiB is allocated but not usable. To make sure that the entire capacity is usable, specify the capacity in 4-GiB increments. If unusable capacity exists, the only way to regain it is to increase the capacity of the volume.
Segment Size	Shows the setting for segment sizing, which only appears for volumes in a volume group. You can change the segment size to optimize performance.

Table 1. Field Details

Field	Description
	Allowed segment size transitions – System Manager determines the segment size transitions that are allowed. Segment sizes that are inappropriate transitions from the current segment size are unavailable on the drop-down list. Allowed transitions usually are double or half of the current segment size. For example, if the current volume segment size is 32 KiB, a new volume segment size of either 16 KiB or 64 KiB is allowed. **SSD Cache-enabled volumes** – You can specify a 4-KiB segment size for SSD Cache-enabled volumes. Make sure you select the 4-KiB segment size only for SSD Cache-enabled volumes that handle small-block I/O operations (for example, 16 KiB I/O block sizes or smaller). Performance might be impacted if you select 4 KiB as the segment size for SSD Cache-enabled volumes that handle large block sequential operations. **Amount of time to change segment size** – The amount of time to change a volume's segment size depends on these variables: ▪ The I/O load from the host ▪ The modification priority of the volume ▪ The number of drives in the volume group ▪ The number of drive channels ▪ The processing power of the storage array controllers When you change the segment size for a volume, I/O performance is affected, but your data remains available.
Secure-capable	**Yes** appears next to "Secure-capable" only if the drives in the pool or volume group are secure-capable. Drive Security prevents unauthorized access to the data on a drive that is physically removed from the storage array. This option is available only when the Drive Security feature has been enabled, and a security key is set up for the storage array. A pool or volume group can contain both secure-capable and non-secure-capable drives, but all drives must be secure-capable to use their encryption capabilities.
DA	**Yes** appears next to "DA" only if the drives in the pool or volume group support Data Assurance (DA). DA increases data integrity across the entire storage system. DA enables the storage array to check for errors that might

Table 1. Field Details

Field	Description
	occur when data is moved between the controllers and drives on a Storage Array.

- **Application-specific workload** – Either click **Next** to accept the system-recommended volumes and characteristics for the selected workload, or click **Edit Volumes** to change, add, or delete the system-recommended volumes and characteristics for the selected workload.

  Field Details

Field	Description
Volume Name	A volume is assigned a default name by System Manager during the volume creation sequence. You can either accept the default name or provide a more descriptive one indicating the type of data stored in the volume.
Reported Capacity	Define the capacity of the new volume and the capacity units to use (MiB, GiB, or TiB). For **Thick volumes**, the minimum capacity is 1 MiB, and the maximum capacity is determined by the number and capacity of the drives in the pool or volume group. Keep in mind that storage capacity is also required for copy services (snapshot images, snapshot volumes, volume copies, and remote mirrors); therefore, do not allocate all of the capacity to standard volumes. Capacity in a pool is allocated in 4-GiB increments. Any capacity that is not a multiple of 4 GiB is allocated but not usable. To make sure that the entire capacity is usable, specify the capacity in 4-GiB increments. If unusable capacity exists, the only way to regain it is to increase the capacity of the volume.
Volume Type	Volume type indicates the type of volume that was created for an application-specific workload.
Segment Size	Shows the setting for segment sizing, which only appears for volumes in a volume group. You can change the segment size to optimize performance. **Allowed segment size transitions** – System Manager determines the segment size transitions that are allowed. Segment sizes that are inappropriate transitions from the current segment size are unavailable on the drop-down list. Allowed transitions usually are double or half of the current segment size. For example, if the current volume segment size is 32 KiB, a new volume segment size of either 16 KiB or 64 KiB is allowed.

Field	Description
	SSD Cache-enabled volumes – You can specify a 4-KiB segment size for SSD Cache-enabled volumes. Make sure you select the 4-KiB segment size only for SSD Cache-enabled volumes that handle small-block I/O operations (for example, 16 KiB I/O block sizes or smaller). Performance might be impacted if you select 4 KiB as the segment size for SSD Cache-enabled volumes that handle large block sequential operations. **Amount of time to change segment size** – The amount of time to change a volume's segment size depends on these variables: ▪ The I/O load from the host ▪ The modification priority of the volume ▪ The number of drives in the volume group ▪ The number of drive channels ▪ The processing power of the storage array controllers When you change the segment size for a volume, I/O performance is affected, but your data remains available.
Secure-capable	**Yes** appears next to "Secure-capable" only if the drives in the pool or volume group are secure-capable. Drive security prevents unauthorized access to the data on a drive that is physically removed from the storage array. This option is available only when the drive security feature has been enabled, and a security key is set up for the storage array. A pool or volume group can contain both secure-capable and non-secure-capable drives, but all drives must be secure-capable to use their encryption capabilities.
DA	**Yes** appears next to "DA" only if the drives in the pool or volume group support Data Assurance (DA). DA increases data integrity across the entire storage system. DA enables the storage array to check for errors that might occur when data is moved between the controllers and drives on a Storage Array.

2. To continue the volume creation sequence for the selected application, click **Next**, and go to Step 4: Review volume configuration.

Step 4: Review volume configuration

Review a summary of the volumes you intend to create and make any necessary changes.

1. Review the volumes you want to create. Click **Back** to make any changes.
2. When you are satisfied with your volume configuration, click **Finish**.

System Manager creates the new volumes in the selected pools and volume groups, and then displays the new volumes in the **All Volumes** table.

- Perform any operating system modifications necessary on the application host so that the applications can use the volume.
- Run either the host-based hot_add utility or an operating system-specific utility (available from a third-party vendor), and then run the SMdevices utility to correlate volume names with host storage array names.

  The hot_add utility and the SMdevices utility are included as part of the SMutils package. The SMutils package is a collection of utilities to verify what the host sees from the storage array. It is included as part of the Storage Manager software installation.

Defining a host in ThinkSystem System Manager

You can create a host automatically or manually. To make it easier to give multiple hosts access to the same volumes, you can also create a host cluster.

Create host automatically

You can allow the Host Context Agent (HCA) to automatically detect the hosts, and then verify that the information is correct. Creating a host is one of the steps required to let the storage array know which hosts are attached to it and to allow I/O access to the volumes.

Before you begin

The Host Context Agent (HCA) is installed and running on every host connected to the storage array. Hosts with the HCA installed and connected to the storage array are created automatically. To install the HCA, install ThinkSystem Storage Manager on the host and select the Host option. The HCA is not available on all supported operating systems. If it is not available, you must create the host manually.

1. Select **Storage** > **Hosts** .

   The table lists the automatically-created hosts.

2. Verify that the information provided by the HCA is correct (name, host type, host port identifiers).

   If you need to change any of the information, select the host, and then click **View/Edit Settings** .

3. (Optional) If you want the automatically-created host to be in a cluster, create a host cluster and add the host or hosts.

What happens next?

After a host is created automatically, the system displays the following items in the Hosts tile table:

- The host name derived from the system name of the host.
- The host identifier ports that are associated with the host.
- The Host Operating System Type of the host.

Create host manually

For hosts that cannot be automatically discovered, you can manually create a host. Creating a host is one of the steps required to let the storage array know which hosts are attached to it and to allow I/O access to the volumes.

About this task

Keep these guidelines in mind when you create a host:

- You must define the host identifier ports that are associated with the host.
- Make sure that you provide the same name as the host's assigned system name.

- This operation does not succeed if the name you choose is already in use.
- The length of the name cannot exceed 30 characters.

1. Select **Storage** > **Hosts** .
2. Click **Create** > **Host** .

   The **Create Host** dialog box appears.

3. Select the settings for the host as appropriate.

Table 1. Field Details

Setting	Description
Name	Type a name for the new host.
Host operating system type	Select the operating system that is running on the new host from the drop-down list.
Host interface type	(Optional) If you have more than one type of host interface supported on your storage array, select the host interface type that you want to use.
Host ports	Do one of the following: • **Select I/O Interface** Generally, the host ports should have logged in and be available from the drop-down list. You can select the host port identifiers from the list. • **Manual add** If a host port identifier is not displayed in the list, it means that the host port has not logged in. An HBA utility or the iSCSI initiator utility may be used to find the host port identifiers and associate them with the host. You can manually enter the host port identifiers or copy/paste them from the utility (one at a time) into the **Host ports** field. You must select one host port identifier at a time to associate it with the host, but you can continue to select as many identifiers that are associated with the host. Each identifier is displayed in the **Host ports** field. If necessary, you also can remove an identifier by selecting the **X** next to it.
CHAP initiator	(Optional) If you selected or manually entered a host port with an iSCSI IQN, and if you want to require a host that tries to access the storage array to authenticate using Challenge Handshake Authentication Protocol (CHAP), select the **CHAP initiator** checkbox. For each iSCSI host port you selected or manually entered, do the following: • Enter the same CHAP secret that was set on each iSCSI host initiator for CHAP authentication. If you are using mutual CHAP authentication (two-way authentication that enables a host to validate itself to the storage array and for a storage array to validate itself to the host), you also must set the

Table 1. Field Details

Setting	Description
	CHAP secret for the storage array at initial setup or by changing settings. • Leave the field blank if you do not require host authentication. Currently, the only iSCSI authentication method used by System Manager is CHAP.

4. Click **Create** .

What happens next?

After the host is successfully created, the system creates a default name for each host port configured for the host (user label).

The default alias is < Hostname_Port Number >. For example, the default alias for the first port created for host IPT is IPT_1 .

Create host cluster

You create a host cluster when two or more hosts require I/O access to the same volumes.

About this task

Keep these guidelines in mind when you create a host cluster:

- This operation does not start unless there are two or more hosts available to create the cluster.
- Hosts in host clusters can have different operating systems (heterogeneous).
- NVMe hosts in host clusters cannot be mixed with non-NVMe hosts.
- This operation does not succeed if the name you choose is already in use.
- The length of the name cannot exceed 30 characters.

1. Select **Storage** > **Hosts** .
2. Select **Create** > **Host Cluster** .

   The **Create Host Cluster** dialog box appears.
3. Select the settings for the host cluster as appropriate.

   Table 1. Field Details

Setting	Description
Name	Type the name for the new host cluster.
Select hosts to share volume access	Select two or more hosts from the drop-down list. Only those hosts that are not already part of a host cluster appear in the list.

4. Click **Create** .

   If the selected hosts are attached to interface types that have different Data Assurance

(DA) capabilities, a dialog appears with the message that DA will be unavailable on the host cluster. This unavailability prevents DA-enabled volumes from being added to the host cluster. Select **Yes** to continue or **No** to cancel.

DA increases data integrity across the entire storage system. DA enables the storage array to check for errors that might occur when data is moved between the controllers and the drives. Using DA for the new volume ensures that any errors are detected.

What happens next?

The new host cluster appears in the table with the assigned hosts in the rows beneath.

Mapping a volume to a host

For a host or host cluster to send I/O to a volume, you must assign the volume to the host or host cluster.

You can select a host or host cluster when you create a volume or you can assign a volume to a host or host cluster later. A host cluster is a group of hosts. You create a host cluster to make it easy to assign the same volumes to multiple hosts.

Assigning volumes to hosts is flexible, allowing you to meet your particular storage needs.

- **Stand-alone host, not part of a host cluster** – You can assign a volume to an individual host. The volume can be accessed only by the one host.
- **Host cluster** – You can assign a volume to a host cluster. The volume can be accessed by all the hosts in the host cluster.
- **Host within a host cluster** – You can assign a volume to an individual host that is part of a host cluster. Even though the host is part of a host cluster, the volume can be accessed only by the individual host and not by any other hosts in the host cluster.

When volumes are created, logical unit numbers (LUNs) are assigned automatically. The LUN serves as the "address" between the host and the controller during I/O operations. You can change LUNs after the volume is created.

Discovering, Configuring, and Verifying storage on the host

Volumes on your storage system appear as disk LUNs to the Linux host when you use FC, iSCSI, or SAS. It shows as NVMe namespace when you use NVMe over RoCE or NVMe over Fibre Channel. When you add new volumes, you must manually rescan the associated LUNs or namespace to discover them. The host does not automatically discover new storage space. The discovering procedures are protocol specific, see related chapters in previous sections.

Where to find additional information

Use the resources listed here if you need additional information.

- ThinkSystem Storage DE Series on-line publications:
  - Hardware Installation and Maintenance Guide, Version 11.70.1
  - SAN Manager software, Version 5.1
  - System Manager software, Version 11.70.1
  - Embedded Command Line Interface, Version 11.70.1
- Lenovo Storage Interoperation Center
- Lenovo Press DE Series Storage

Contacting Support

You can contact Support to obtain help for your issue.

You can receive hardware service through a Lenovo Authorized Service Provider. To locate a service provider authorized by Lenovo to provide warranty service, go to https://datacentersupport.lenovo.com/serviceprovider and use filter searching for different countries. For Lenovo support telephone numbers, see https://datacentersupport.lenovo.com/supportphonelist for your region support details.

Notices

Lenovo may not offer the products, services, or features discussed in this document in all countries. Consult your local Lenovo representative for information on the products and services currently available in your area.

Any reference to a Lenovo product, program, or service is not intended to state or imply that only that Lenovo product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any Lenovo intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any other product, program, or service.

Lenovo may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document is not an offer and does not provide a license under any patents or patent applications. You can send inquiries in writing to the following:

Lenovo (United States), Inc.

8001 Development Drive

Morrisville, NC 27560

U.S.A.

Attention: Lenovo Director of Licensing

LENOVO PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some jurisdictions do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. Lenovo may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

The products described in this document are not intended for use in implantation or other life support applications where malfunction may result in injury or death to persons. The information contained in this document does not affect or change Lenovo product specifications or warranties. Nothing in this document shall operate as an express or implied license or indemnity under the intellectual property rights of Lenovo or third parties. All information contained in this document was obtained in specific environments and is presented as an illustration. The result obtained in other operating environments may vary.

Lenovo may use or distribute any of the information you supply in any way it believes appropriate without incurring any obligation to you.

Any references in this publication to non-Lenovo Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this Lenovo product, and use of those Web sites is at your own risk.

Any performance data contained herein was determined in a controlled environment. Therefore, the result obtained in other operating environments may vary significantly. Some measurements may have been made on development-level systems and there is no guarantee that these measurements will be the same on generally available systems. Furthermore, some measurements may have been estimated through extrapolation. Actual results may vary. Users of this document should verify the applicable data for their specific environment.

Trademarks

LENOVO, LENOVO logo, and THINKSYSTEM are trademarks of Lenovo. All other trademarks are the property of their respective owners. © 2021 Lenovo