iSER over InfiniBand Setup

e-series

NetApp
March 11, 2022

Table of Contents

iSER over InfiniBand Setup

Verify the Linux configuration is supported

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

Steps

1. Go to the NetApp Interoperability Matrix Tool.
2. Click on the **Solution Search** tile.
3. In the **Protocols › SAN Host** area, click the **Add** button next to **E-Series SAN Host**.
4. Click **View Refine Search Criteria**.

   The Refine Search Criteria section is displayed. In this section you may select the protocol that applies, as well as other criteria for the configuration such as Operating System, NetApp OS, and Host Multipath driver.

5. Select the criteria you know you want for your configuration, and then see what compatible configuration elements apply.
6. As necessary, make the updates for your operating system and protocol that are prescribed in the tool.

   Detailed information for your chosen configuration is accessible on the View Supported Configurations page by clicking the right page arrow.

Configure IP addresses using DHCP

To configure communications between the management station and the storage array, use Dynamic Host Configuration Protocol (DHCP) to provide IP addresses.

What you'll need

A DHCP server installed and configured on the same subnet as the storage management ports.

About this task

Each storage array has either one controller (simplex) or two controllers (duplex), and each controller has two storage management ports. Each management port will be assigned an IP address.

The following instructions refer to a storage array with two controllers (a duplex configuration).

Steps

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.

   > Do not use management port 2 on either controller. Port 2 is reserved for use by NetApp technical personnel.

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.

Configure subnet manager

A subnet manager must be running in your environment on your switch or on your hosts. If you are running it host-side, use the following procedure to set it up.

Steps

1. Install the `opensm` package on any hosts that will be running the subnet manager.
2. Use the `ibstat -p` command to find `GUID0` and `GUID1` of the HBA ports. For example:

```
# ibstat -p
 0x248a070300a80a80
 0x248a070300a80a81
```

3. Enable Subnet Manager on each port of the connected HCA on the host:

   SLES example

   - Add the following two lines to /etc/rc.d/rc.after. Substitute the values you found in step 2 for `GUID0` and `GUID1`. For `P0` and `P1`, use the subnet manager priorities, with 1 being the lowest and 15 the highest:

```
opensm -B -g GUID0 -p P0 -f /var/log/opensm-ib0.log
opensm -B -g GUID1 -p P1 -f /var/log/opensm-ib1.log
```

     An example of the command with value substitutions:

```
# cat /etc/rc.d/rc.local
 opensm -B -g 0x248a070300a80a80 -p 15 -f /var/log/opensm-ib0.log
 opensm -B -g 0x248a070300a80a81 -p 1 -f /var/log/opensm-ib1.log
```

RHEL example

- Add the following two lines to /etc/rc.d/rc.local. Substitute the values you found in step 2 for `GUID0` and `GUID1`. For `P0` and `P1`, use the subnet manager priorities, with 1 being the lowest and 15 the highest:

```
opensm -B -g GUID0 -p P0 -f /var/log/opensm-ib0.log
opensm -B -g GUID1 -p P1 -f /var/log/opensm-ib1.log
```

An example of the command with value substitutions:

```
# cat /etc/rc.d/rc.local
 opensm -B -g 0x248a070300a80a80 -p 15 -f /var/log/opensm-ib0.log
 opensm -B -g 0x248a070300a80a81 -p 1 -f /var/log/opensm-ib1.log
```

Install and configure Linux Unified Host Utilities

The Linux Unified Host Utilities tools help you manage NetApp storage, including failover policies and physical paths.

Steps

1. Use the NetApp Interoperability Matrix Tool to determine the appropriate version of Unified Host Utilities to install.

   The versions are listed in a column within each supported configuration.

2. Download the Unified Host Utilities from NetApp Support.

> Alternatively, you can use the SANtricity SMdevices utility to perform the same functions as the Unified Host Utility tool. The SMdevices utility is 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 SANtricity software installation.

Install SANtricity Storage Manager for SMcli (SANtricity software version 11.53 or earlier)

If you are using SANtricity software 11.53 or earlier, you can install the SANtricity Storage Manager software on your management station to help manage the array.

SANtricity Storage Manager includes the command line interface (CLI) for additional management tasks, and also the Host Context Agent for pushing host configuration information to the storage array controllers through

the I/O path.

> If you are using SANtricity software 11.60 and newer, you do not need to follow these steps. The SANtricity Secure CLI (SMcli) is included in the SANtricity OS and downloadable through the SANtricity System Manager. For more information on how to download the SMcli through the SANtricity System Manager, refer to the *Download command line interface (CLI)* topic under the SANtricity System Manager Online Help.

What you'll need

- SANtricity software 11.53 or earlier.
- Correct administrator or superuser privileges.
- A system for the SANtricity Storage Manager client with the following minimum requirements:
  - **RAM**: 2 GB for Java Runtime Engine
  - **Disk space**: 5 GB
  - **OS/Architecture**: For guidance on determining the supported operating system versions and architectures, go to NetApp Support. From the **Downloads** tab, go to **Downloads › E-Series SANtricity Storage Manager**.

About this task

This task describes how to install SANtricity Storage Manager on both the Windows and Linux OS platforms, because both Windows and Linux are common management station platforms when Linux is used for the data host.

Steps

1. Download the SANtricity software release at NetApp Support. From the **Downloads** tab, go to **Downloads › E-Series SANtricity Storage Manager**.
2. Run the SANtricity installer.

Windows	Linux
Double-click the SMIA*.exe installation package to start the installation.	a. Go to the directory where the SMIA*.bin installation package is located. b. If the temp mount point does not have execute permissions, set the `IATEMPDIR` variable. Example: `IATEMPDIR=/root ./SMIA-LINUXX64-11.25.0A00.0002.bin` c. Run the `chmod +x SMIA*.bin` command to grant execute permission to the file. d. Run the `./SMIA*.bin` command to start the installer.

3. Use the installation wizard to install the software on the management station.

Access SANtricity System Manager and use the Setup wizard

To configure your storage array, you can use the Setup wizard in SANtricity System Manager.

SANtricity System Manager is a web-based interface embedded on each controller. To access the user interface, you point a browser to the controller's IP address. A setup wizard helps you get started with system configuration.

What you'll need

- Out-of-band management.
- A management station for accessing SANtricity System Manager that includes one of the following browsers:

Browser	Minimum version
Google Chrome	79
Microsoft Internet Explorer	11
Microsoft Edge	79
Mozilla Firefox	70
Safari	12

About this task

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.

Steps

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

   `IPAddress` is the address for one of the storage array controllers.

   The first time SANtricity 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. For more information about the four local user roles, see the online help available in the SANtricity System Manager user interface.

2. Enter the System Manager password for the admin role in the Set Administrator Password and Confirm

Password fields, and then click **Set Password**.

The Setup wizard launches if there are no pools, volumes groups, workloads, or notifications configured.

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 SANtricity System Manager.

Configure the multipath software

To provide a redundant path to the storage array, you can configure multipath software.

What you'll need

You must install 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`.

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

About this task

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

Steps

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. Save your kernel version by running the `uname -r` command.

```
# uname -r
3.10.0-327.el7.x86_64
```

You will use this information when you assign volumes to the host.

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

If you are using….	Do this…
RHEL 7.x and 8.x systems:	`systemctl enable multipathd`
SLES 12.x and 15.x systems:	`systemctl enable multipathd`

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

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

7. 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`.

8. 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.
9. Reboot the host.

Set 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 with a first non-white-space character of # is considered a comment line.

For SANtricity operating system 8.30 and newer, NetApp recommends using the default settings as provided.

The multipath.conf files 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`

Configure network connections

If your configuration uses the iSER over InfiniBand protocol, perform the steps in this section to configure network connections.

Steps

1. From System Manager, go to **Settings › System › Configure iSER over Infiniband Ports**. Refer to the System Manager online help for further instructions.

   Put the array iSCSI addresses on the same subnet as the host port(s) you will use to create iSCSI sessions. For addresses, see your iSER worksheet.

2. Record the IQN.

   This information might be necessary when you create iSER sessions from operating systems that do not support send targets discovery. Enter this information in the iSER worksheet.

Configure networking for storage attached hosts

If your configuration uses the iSER over InfiniBand protocol, perform the steps in this section.

The InfiniBand OFED driver stack supports running both iSER and SRP simultaneously on the same ports, so no additional hardware is required.

What you'll need

A NetApp recommended OFED installed on the system. For more information, see the NetApp Interoperability Matrix Tool.

Steps

1. Enable and start iSCSI services on the host(s):

   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 and 15 (SLES 12 and SLES 15)

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

2. Configure IPoIB network interfaces:
   a. Identify the InfiniBand ports that will be used. Document the HW Address (MAC address) of each port.
   b. Configure persistent names for the InfiniBand network interface devices.
   c. Configure the IP address and network information for the IPoIB interfaces identified.

      The specific interface configuration required might vary depending on the operating system used. Consult your vendor's operating system documentation for specific information on implementation.

   d. Start the IB network interfaces by restarting the networking service or by manually restarting each interface. For example:

      ```
      systemctl restart network
      ```

   e. Verify connectivity to the target ports. From the host, ping the IP addresses you configured when you configured network connections.
3. Restart services to load the iSER module.
4. Edit the iSCSI settings in /etc/iscsi/iscsid.conf.

```
node.startup = automatic
replacement_timeout = 20
```

5. Create iSCSI session configurations:
   a. Create iface configuration files for each InfiniBand interface.

      The directory location for the iSCSI iface files is operating system dependent. This example is for using Red Hat Enterprise Linux:

      ```
      iscsiadm -m iface -I iser > /var/lib/iscsi/ifaces/iface-ib0
      iscsiadm -m iface -I iser > /var/lib/iscsi/ifaces/iface-ib1
      ```

   b. Edit each iface file to set the interface name and initiator IQN. Set the following parameters appropriately for each iface file:

Option	Value
iface.net_ifacename	The interface device name (ex. ib0).

Option	Value
iface.initiatorname	The host initiator IQN documented in the worksheet.

c. Create iSCSI sessions to the target.

The preferred method to create the sessions is to use the SendTargets discovery method. However, this method does not work on some operating system releases.

Use **Method 2** for RHEL 6.x or SLES 11.3 or later.

- **Method 1 - SendTargets discovery:** Use the SendTargets discovery mechanism to one of the target portal IP addresses. This will create sessions for each of the target portals.

```
iscsiadm -m discovery -t st -p 192.168.130.101 -I iser
```

- **Method 2 - Manual creation:** For each target portal IP address, create a session using the appropriate host interface iface configuration. In this example, interface ib0 is on subnet A and interface ib1 is on subnet B. For these variables, substitute the appropriate value from the worksheet:
  - <Target IQN> = storage array Target IQN
  - <Target Port IP> = IP address configured on the specified target port

```
# Controller A Port 1
iscsiadm -m node --target <Target IQN\> -I iface-ib0 -p <Target Port
IP\> -l -o new
# Controller B Port 1
iscsiadm -m node --target <Target IQN\> -I iface-ib0 -p <Target Port
IP\> -l -o new
# Controller A Port 2
iscsiadm -m node --target <Target IQN\> -I iface-ib1 -p <Target Port
IP\> -l -o new
# Controller B Port 2
iscsiadm -m node --target <Target IQN\> -I iface-ib1 -p <Target Port
IP\> -l -o new
```

6. Log in to iSCSI sessions.

For each session, run the iscsiadm command to log in to the session.

```
# Controller A Port 1
iscsiadm -m node --target <Target IQN\> -I iface-ib0 -p <Target Port
IP\> -l
# Controller B Port 1
iscsiadm -m node --target <Target IQN\> -I iface-ib0 -p <Target Port
IP\> -l
# Controller A Port 2
iscsiadm -m node --target <Target IQN\> -I iface-ib1 -p <Target Port
IP\> -l
# Controller B Port 2
iscsiadm -m node --target <Target IQN\> -I iface-ib1 -p <Target Port
IP\> -l
```

7. Verify the iSER/iSCSI sessions.
   a. Check the iscsi session status from the host:

      ```
      iscsiadm -m session
      ```

   b. Check the iscsi session status from the array. From SANtricity System Manager, navigate to **Storage Array** > **iSER** > **View/End Sessions**.

When the OFED/RDMA service starts, the iSER kernel module(s) loads by default when the iSCSI services are running. To complete the iSER connection setup, the iSER module(s) should be loaded. Currently this requires a host reboot.

Create partitions and filesystems

Because 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.

What you'll need

- A LUN that is discovered by the host.
- A list of available disks. (To see available disks, run the `ls` command in the /dev/mapper folder.)

About this task

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.

Steps

1. Retrieve the SCSI ID of the mapped disk by issuing the `sanlun lun show -p` 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.

```
# sanlun lun show -p

                E-Series Array: ictm1619s01c01-
SRP(60080e50002908b40000000054efb9d2)
                   Volume Name:
                Preferred Owner: Controller in Slot B
                  Current Owner: Controller in Slot B
                           Mode: RDAC (Active/Active)
                       UTM LUN: None
                           LUN: 116
                      LUN Size:
                       Product: E-Series
                   Host Device:
mpathr(360080e50004300ac000007575568851d)
               Multipath Policy: round-robin 0
             Multipath Provider: Native
--------- ---------- ------- ------------
----------------------------------------------
host      controller                      controller
path      path       /dev/   host         target
state     type       node    adapter      port
--------- ---------- ------- ------------
----------------------------------------------
up        secondary  sdcx    host14       A1
up        secondary  sdat    host10       A2
up        secondary  sdbv    host13       B1
```

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.

What you'll need

An initialized volume that is formatted with a file system.

Steps

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.

After you finish

Remove the file and folder that you copied.

Record your iSER over IB configuration

You can generate and print a PDF of this page, and then use the following worksheet to record iSER over Infiniband storage configuration information. You need this information to perform provisioning tasks.

Host identifiers

The software initiator IQN is determined during the task, Configure networking for storage attached hosts.

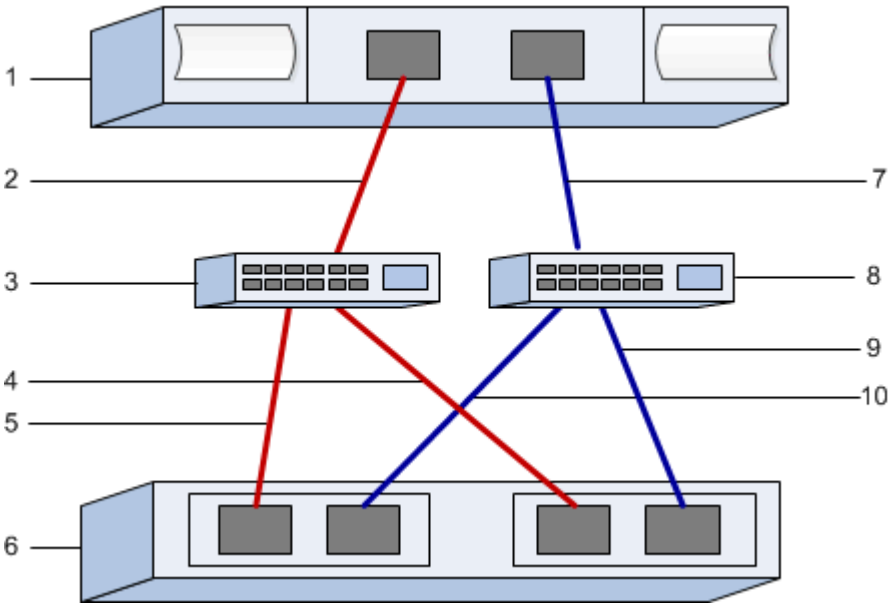
Locate and document the initiator IQN from each host. For software initiators, the IQN is typically found in the /etc/iscsi/initiatorname.iscsi file.

Callout No.	Host port connections	Software initiator IQN
1	Host (initiator) 1	
n/a		
n/a		

Callout No.	Host port connections	Software initiator IQN
n/a		
n/a		

Recommended configuration

Recommended configurations consist of two host (initiator) ports and four target ports.



Target IQN

Document the target IQN for the storage array. You will use this information in Configure networking for storage attached hosts.

Find the Storage Array IQN name using SANtricity: **Storage Array** > **iSER** > **Manage Settings**. This information might be necessary when you create iSER sessions from operating systems that do not support send targets discovery.

Callout No.	Array name	Target IQN
6	Array controller (target)	

Network configuration

Document the network configuration that will be used for the hosts and storage on the InfiniBand fabric. These instructions assume that two subnets will be used for full redundancy.

Your network administrator can provide the following information. You use this information in the topic, Configure networking for storage attached hosts.

Subnet A

Define the subnet to be used.

Network Address	Netmask

Document the IQNs to be used by the array ports and each host port.

Callout No.	Array controller (target) port connections	IQN
3	Switch	*not applicable*
5	Controller A, port 1	
4	Controller B, port 1	
2	Host 1, port 1	
	(Optional) Host 2, port 1	

Subnet B

Define the subnet to be used.

Network Address	Netmask

Document the IQNs to be used by the array ports and each host port.

Callout No.	Array controller (target) port connections	IQN
8	Switch	*not applicable*
10	Controller A, port 2	
9	Controller B, port 2	
7	Host 1, port 2	
	(Optional) Host 2, port 2	

Mapping host name

The mapping host name is created during the workflow.

Mapping host name	
Host OS type	